

Article in press, accepted for publication 10/20/2012

Journal: NeuroImage
Journal Section: Cognitive

Title: The time course of shape discrimination in the human brain

Authors:

Justin M. Ales^a, L. Gregory Appelbaum^{b,c}, Benoit R. Cottureau^{a,d}, Anthony M. Norcia^a

^a Department of Psychology, Stanford University, Stanford, CA

^b Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC

^c Center for Cognitive Neuroscience, Duke University, Durham, NC

^d Université de Toulouse, Centre de Recherche Cerveau et Cognition (CerCo), Toulouse,

Corresponding author:

Justin M. Ales,
Stanford University
Department of Psychology
Jordan Hall, Building 01-420
Stanford, CA 94305
justin.ales@stanford.edu
650-725-2440

Keywords:

Perceptual decision, shape discrimination, Visual Evoked Potentials,

Abstract

The lateral occipital cortex (LOC) activates selectively to images of intact objects versus scrambled controls, is selective for the figure-ground relationship of a scene, and exhibits at least some degree of invariance for size and position. Because of these attributes, it is considered to be a crucial part of the object recognition pathway. Here we show that human LOC is critically involved in perceptual decisions about object shape. High-density EEG was recorded while subjects performed a threshold-level shape discrimination task on texture-defined figures segmented by either phase or orientation cues. The appearance or disappearance of a figure region from a uniform background generated robust visual evoked potentials throughout retinotopic cortex as determined by inverse modeling of the scalp voltage distribution. Contrasting responses from trials containing shape changes that were correctly detected (hits) with trials in which no change occurred (correct rejects) revealed stimulus-locked, target-selective activity in the occipital visual areas LOC and V4 preceding the subject's response. Activity that was locked to the subjects' reaction time was present in the LOC. Response-locked activity in the LOC was determined to be related to shape discrimination for several reasons: shape-selective responses were silenced when subjects viewed identical stimuli but their attention was directed away from the shapes to a demanding letter discrimination task; shape-selectivity was present across four different stimulus configurations used to define the figure; LOC responses correlated with participants' reaction times. These results indicate that decision-related activity is present in the LOC when subjects are engaged in threshold-level shape discriminations.

Introduction

Shape has been described as being “what is left when the (image) differences which can be attributed to translations, rotations, and dilatations have been quotiented out” (Kendall, 1984). To this definition, visual neuroscientists would include insensitivity to the cue that signals the boundary. This definition describes everything shape is not, but does not provide any guidance as to how shape should be represented.

Many experiments in macaque have demonstrated that areas V4 and IT exhibit selectivity for geometrical properties of object boundaries while also showing a degree of location, cue and size invariance (see Pasupathy, 2006; Tompa and Sary, 2010 for reviews). Single unit studies provide an accurate and detailed description of the presence and degree shape selectivity of single cells in identified cortical areas. However, only a few areas have been explored. None of the studies have recorded shape selectivity in multiple areas at the same time, and this complicates determinations of the temporal sequence of shape processing operations

With its high spatial resolution and large field of view, functional magnetic resonance imaging (fMRI) in humans has been used extensively to localize and characterize networks involved in object recognition. In these studies, area V4 and the lateral occipital complex (LOC) have been shown to be responsive to shapes (Malach et al., 1995; Kourtzi and Kanwisher, 2000; Haushofer et al., 2008; Vinberg and Grill-Spector, 2008), and to exhibit a degree of invariance over location, size and viewpoint (Konen and Kastner, 2008; Cichy et al., 2011). Attempts have been made to discern the order of information flow through these areas using techniques such as Dynamic Causal Modeling (DCM) (Summerfield and Koechlin, 2008). While a DCM of the BOLD signal can discover latent connectivity between cortical areas, given uncertainties about the relationship between the millisecond time-scale of neuronal activity and the much slower time-evolution of the BOLD signal, it does not assign an absolute time to the sequence of activations, only a coupling strength between cortical areas.

Here we are interested in the neural substrate underlying precise judgments of shape --- those that are at the threshold of visibility. Our goal was to find the earliest stage of the visual pathway that contains the requisite information needed to drive a threshold-level percept of a change in shape. Instead of probing a wide range of different stimulus shapes, locations and cues to study how shape is encoded in the cortex, we studied how small shape changes are processed. By using stimuli defined by second-order (orientation and alignment) texture boundaries, we have diminished the contribution from neurons responding to low-level aspects such as the luminance or contrast cues signaling the object boundary. By focusing on responses to threshold-level events, we further restrict the population of cells contributing to the measurements to the population that enables shape discriminations. Through the combination of these stimulus-control measures and a novel response-locked analysis, we show that LOC initiates fine shape discriminations. Following the LOC, later activity was observed in an extended cortical network that includes V4.

Materials and Methods

Participants

Twelve subjects (9 males, 3 females, mean age 35.8) with normal or corrected to normal vision participated. Their informed consent was obtained prior to experimentation under a protocol that was approved by the Institutional Review Board of the Smith-Kettlewell Eye Research Institute.

Stimuli

The stimuli are conceptually similar to dynamic random-dot stereograms or coherent motion displays in which the local elements constantly change at a fast rate, masking low-level cues that occur when global aspects of the display are updated at a slower rate. All of the conditions were constructed from the same local elements: one-dimensional noise patterns that updated at 36 Hz. To create different figure/ground segmentation states we changed the relative orientation of the 1-d noise patterns. The screen appears uniform when the figure and background regions have the same orientation. Altering the orientation of the figure-region to be rotated 90° relative to the background texture segments the figure regions on the basis of a difference in orientation. By flipping the figure regions 180° relative to the background texture, the figure regions are segmented due to a difference in element alignment. We will refer to these states as the 'orientation defined' and 'phase defined' states, respectively.

The four conditions used in this study are built from the above-described segmentation states and are shown schematically in the upper right panel of Figure 1. Across these four conditions, the stimuli varied with regard to both the segmentation cue and whether the figures appeared and disappeared from the background or were continuously segmented from the background. Condition 1 "Orientation Defined/Figure Appearance" alternated between a uniform field and a field containing orientation-defined shapes. Condition 2 "Phase Defined/Figure Appearance" alternated between a uniform field and a field containing phase-defined shapes. In Condition 3 "Changing Cue/Constant Figure" the shapes were always segmented but were alternately defined by a phase difference or an orientation difference between the figure and background regions (i.e. the elements within the figure regions rotated by 90 deg of orientation). Finally, in Condition 4 "Phase Defined/Constant Figure", alignment differences were always present giving the appearance of a continuously segmented field with no 1 Hz modulation of the global configuration.

Stimuli were generated on a Power Macintosh G4 running the in-house PowerDiva software suite and presented on a LaCie CRT monitor at a resolution of 800x600, and a 72 Hz vertical refresh rate and an average luminance of 81.4 cd/m². The luminance values in the textures were uniformly distributed over a range of 8.1 to 155, cd/m² yielding a Michelson contrast of 90%. The minimal width of a texture element was 6 arcmin. Both the background and figure regions were made of textures with the same average statistics and the stimulus subtended 23° both vertically and horizontally. The figure regions were 5° diameter discs

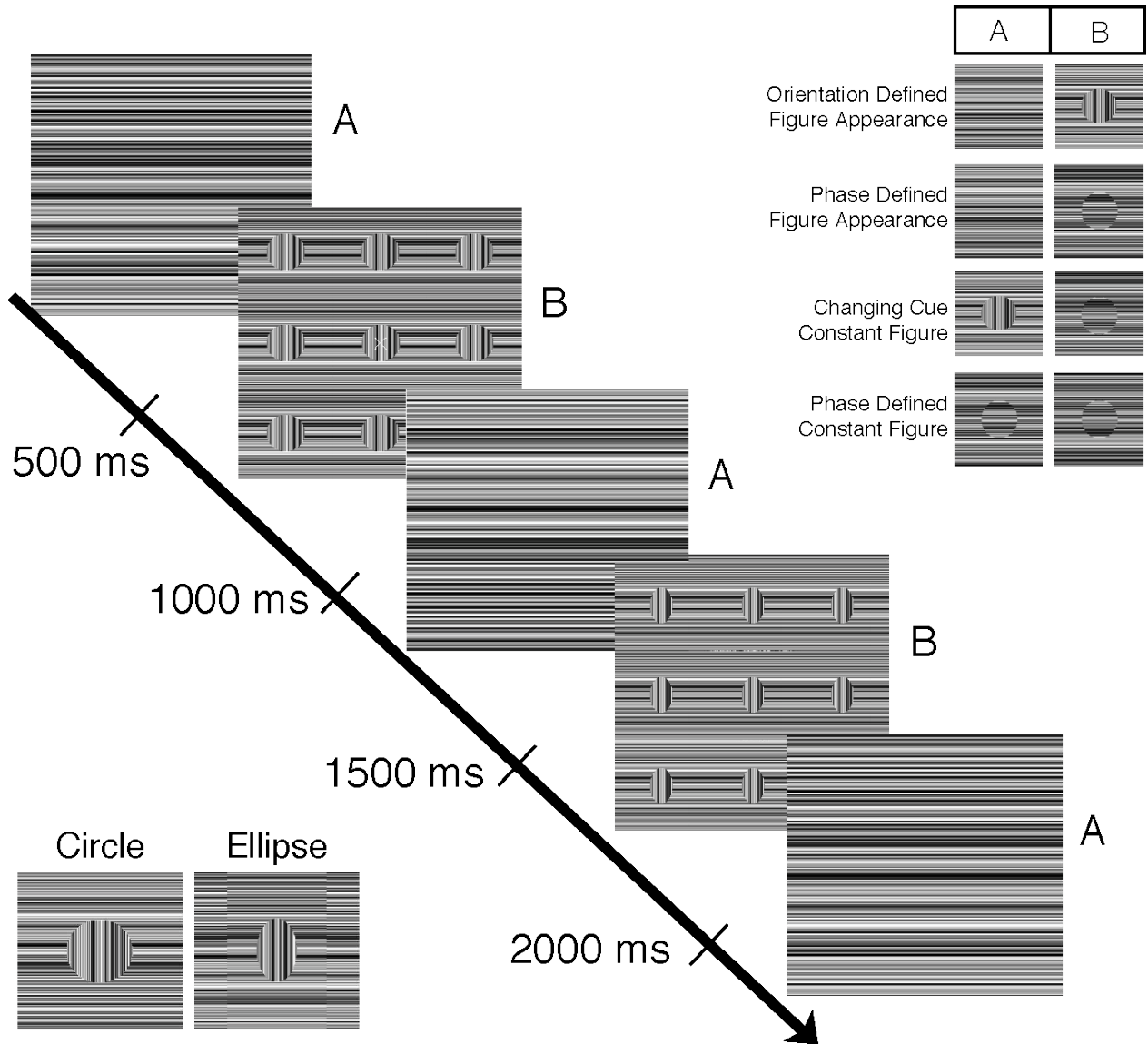


Figure 1
Schematic illustration of stimulus timing and configuration. Texture-defined forms were portrayed using dynamic 1-D spatial noise updated at 36 Hz. The texture within nine disk-shaped regions was modulated at 1 Hz between two states (identified with the A and B labels). A 90° orientation change of the disk regions is illustrated here for the condition in which the disk regions appeared from a uniform background (Orientation defined figure appearance as illustrated in the upper right inset). There were 9 disk regions, each 5° in diameter spaced 7.5° center-to-center. On 20% of the stimulus transitions the aspect ratio of the disk regions was deformed by a threshold amount. A highly suprathreshold ellipse with an aspect ratio of 0.64 is illustrated in the second interval, on average participants were shown a .94 aspect ratio as the target.

separated by 7.5° in a 3x3 array that could appear to be segmented from the background depending on whether the figure and background regions matched exactly. Stimuli were viewed binocularly in a dark and quiet room as whole head, 128-channel EEG was simultaneously recorded. For each condition observers viewed 20 individual trials each lasting 12 seconds. Steady-State analysis was restricted to the central 10 seconds of each trial.

Conditions were randomized within a task block. This results in 200 one-second presentations. A typical session lasted roughly 45 minutes.

Shape Discrimination Task

Before the EEG recording began, we used the method of adjustment to find each subject's threshold for detecting a slight change in the aspect ratio of the circular figure regions. During the EEG recording session, targets comprised of elliptical regions were substituted for circular regions on 20% of the 1 Hz image alternations (40 targets per condition). The observers were asked to detect these deviant targets as quickly and as accurately as possible using a mouse key. The aspect ratio for the targets was set to yield an approximately 80% correct detection rate. Targets had a minimum spacing of 2-seconds so that it was possible to uniquely assign a response to the deviant targets, the 2-second window was also the maximum allowed response time.

Attentional Control Task

In a sub-set of six observers we presented the same series of standard and deviant targets while the subjects performed a difficult letter discrimination task at threshold. The goal here was to test whether responses to the deviant stimuli could be generated automatically, even when the deviants were not task relevant. Randomly oriented T or L letters were presented at 5 positions within the central disk region of the 9-disk display. The array could contain all L's or it could contain one T at a random position. The subject was instructed to press the left key if the array contained a T and the right key if the array comprised all L's. The appearance of the target array was preceded and followed by an array of randomly oriented F's. The time between the preceding and following F letter arrays was placed on a staircase that converged to the 82% correct level. Durations were shortened if the subject made a correct response and were lengthened if they made an error according to the staircase rule. This procedure kept the subjects working at a consistent level of difficulty on the distracter-task while we recorded evoked potentials to the now task-irrelevant figure-ground segmentation stimuli. The segmentation stimuli included embedded deviants that matched those used during the trials where the shape of the figures was task relevant.

EEG Recording

The procedures for this experiment (EEG signal acquisition, head conductivity modeling, source estimation, visual areas definition, region-of-interest quantification, and statistical analysis) are similar to those utilized in our previous studies. In the interest of brevity, we provide an overview of these methods, and reference the reader to our previous work (Appelbaum et al., 2006; Appelbaum et al., 2008; Ales and Norcia, 2009; Appelbaum and Norcia, 2009) for a more a more detailed description.

The electroencephalogram (EEG) was collected with 128-sensor HydroCell Sensor Nets (Electrical Geodesics, Eugene OR) and was band-pass filtered from 0.1 to 200 Hz. Following each experimental session, the 3D locations of all electrodes and three major fiducials (nasion, left and right peri-auricular points) were digitized using a 3Space Fastrack 3-D digitizer (Polhemus, Colchester, VT). For all subjects, the 3D digitized locations were used to co-

register the electrodes to their T1-weighted anatomical MRI scans. Raw data were evaluated off line according to a sample-by-sample thresholding procedure to remove noisy sensors that were replaced by the average of the six nearest spatial neighbors. Once noisy sensors were substituted, the EEG was re-referenced to the common average of all the sensors. Additionally, EEG epochs that contained a large percentage of data samples exceeding threshold (50 microvolts) were excluded on a sensor-by-sensor basis. Typically these epochs are associated with subject motion, eye movements or blinks.

Structural and Functional Magnetic Resonance Imaging (MRI)

Structural and functional MRI scanning was conducted at 3T (Siemens Tim Trio, Erlangen, Germany) using a 12-channel head coil. We acquired a T1-weighted MRI dataset (3-D MP-RAGE sequence, 0.8 x 0.8 x 0.8 mm³ and a 3-D T2-weighted dataset (SE sequence at 1 x 1 x 1 mm³ resolution) for tissue segmentation and registration with the functional scans. For fMRI, we employed a single-shot, gradient-echo EPI sequence (TR/TE = 2000/28 ms, flip angle 80, 126 volumes per run) with a voxel size of 1.7 x 1.7 x 2 mm³ (128 x 128 acquisition matrix, 220 mm FOV, bandwidth 1860 Hz/pixel, echo spacing 0.71 ms). We acquired 30 slices without gaps, positioned in the transverse-to-coronal plane approximately parallel to the corpus callosum and covering the whole cerebrum. Once per session, a 2-D SE T1-weighted volume was acquired with the same slice specifications as the functional series in order to facilitate registration of the fMRI data to the anatomical scan.

The FreeSurfer software package (<http://surfer.nmr.mgh.harvard.edu>) was used to perform gray and white matter segmentation. The FreeSurfer package extracts both gray/white and gray/cerebrospinal fluid (CSF) boundaries, but these surfaces can have different surface orientations. In particular, the gray/white boundary has sharp gyri (the curvature changes rapidly) and smooth sulci (slowly changing surface curvature), while the gray/CSF boundary is the inverse, with smooth gyri and sharp sulci. We created a new surface that had a similar curvature for both gyri and sulci, avoiding these curvature discontinuities. The new surface generated by interpolating a position that was midway between the gray/white surface and the gray/CSF surface.

The highest accuracy for source-imaging is obtained when there is an accurate model that connects activity at each location on the surface of cortex with how it will be measured at the scalp. We made separate models for each participant in the study using the Boundary Element Method (BEM) with conductivity models that were derived from the T1 and T2 weighted MRI scans of each observer. The FSL toolbox (<http://www.fmrib.ox.ac.uk/fsl/>) was also used to segment contiguous volume regions for the scalp, outer skull, and inner skull and to convert these MRI volumes into inner skull, outer skull, and scalp surfaces (Smith, 2002; Smith et al., 2004).

The general procedures for the scans used to define the visual areas (head stabilization, visual display system, etc) are standard and have been described in detail elsewhere (Brewer et al., 2005). Retinotopic field mapping defined Regions-of-Interest (ROIs) for visual cortical areas V1, V2v, V2d, V3v, V3d, V3A, and V4 in each hemisphere (Tootell and Hadjikhani, 2001; Wade et al., 2002). ROIs corresponding to hMT+ were identified using low contrast motion stimuli similar to those described Huk et al. (2002).

The Lateral Occipital Complex (LOC) was defined using a block-design fMRI localizer scan. During this scan, the observers viewed blocks of images depicting common objects (18s/block) alternating with blocks containing scrambled versions of the same objects. The stimuli were those used in a previous study (Kourtzi and Kanwisher, 2000). The regions activated by these scans included an area lying between the V1/V2/V3 foveal confluence and hMT+ that we identified as LOC. Only the portion lying on the lateral surface, posterior and adjacent to hMT+ was included in our definition. Activations in ventral areas were more variable and sources in these areas are less visible in the EEG due to their greater depth. Our definition covers almost all regions (e.g. V4d, LOC, LOp) that have previously been identified as lying within object-responsive lateral occipital cortex (Kourtzi and Kanwisher, 2000; Tootell and Hadjikhani, 2001). Cross-subject surface averages of our LOC ROI can be seen Figure 6c. Examples of the ROI definitions used here on individual subjects can be found in Figure 2 of (Appelbaum et al., 2010). In this study, the fMRI data was used purely to define ROIs for the EEG analysis.

Source Imaging Procedure

In order to recover activity on the cortex we used the L2 minimum-norm inverse. Detailed discussion of the source imaging procedure is provided in Cottareau et al. (2011), we provide some essential details below. The sources were constrained to the location and orientation of the cortical surface (Hamalainen and Ilmoniemi, 1994; Dale et al., 2000; Cottareau et al., 2011). We used a mid-gray cortical surface that was down sampled to have 20,484 isotropically spaced vertices. In order to avoid having the difference between signal-to-noise ratios across conditions change the amount of regularization, and thus the magnitude of the recovered sources, we fixed the regularization to be consistent across all conditions for a single participant. The amount of regularization for each participant was determined using generalized cross validation averaging over all conditions (Wahba, 1990). Given this cortically constrained activity we estimated the response magnitude from each fMRI defined ROI by coherently averaging across all source locations within that ROI. We averaged the ROI defined activations to combine data across subjects. The lightly shaded area around waveforms represents ± 1 standard error of the mean.

Subject Surface Averaging

Cross subject, surface-based averaging was used to extend the source-imaging data outside of the functionally-defined ROIs. Averaging was performed with the FreeSurfer image analysis suite (<http://surfer.nmr.mgh.harvard.edu/>) as documented in (Fischl et al., 1999; Yeo et al., 2010a; Yeo et al., 2010b). First, each subject's cortical surface was inflated to a sphere. Using the pattern of gyral and sulcal curvature, each subject's cortical surface was then aligned to a reference subject. This alignment defines a unique unidirectional transform from every single subject to the target subject. The alignment was then used to average the source data across subjects. For cross subject, surface-based averaging only, the source reconstructions were smoothed over 1st and 2nd order neighboring vertices, this resulted in a boxcar smoothing with a 4mm width. EEG time courses were derived from the vertex containing the response maxima within each ROI.

Permutation Tests

Differences between the experimental conditions were identified by a permutation test based on methods devised by Blair and Karniski (1993) and described in detail in Appelbaum et al. (2006). Briefly, the null hypothesis that no differences were present between experimental conditions was tested by making synthetic datasets in which the condition labels for an individual subject's data were randomly permuted. For each permutation we calculated *t*-scores of the difference and found the longest run of consecutive time points with *p*-values less than .05. This procedure provides a nonparametric reference distribution of consecutive significant *p*-values. We then rejected the null hypothesis if the length of any consecutive sequence of significant *t*-scores in the original, nonrandomized data exceeded 95% of the values in the null distribution. Because each permutation sample contributes only its *longest* significant sequence to the reference distribution this procedure implicitly compensates for the problem of multiple comparisons, and is a valid test for the omnibus hypothesis of no difference between the waveforms at any time point. Furthermore, this test not only detects significant departures from the null hypothesis, it also localizes the time periods when such departures occur. However, since the correction procedure is tied to the length of the data and the somewhat arbitrary choice of keeping family-wise error at 5%, we therefore also present the uncorrected significance values (see red/yellow color maps in Figures 4 and 7). By evaluating the data using both statistical approaches, we are better able to identify time periods when the responses depart from the null hypothesis.

Results

Stimulus-locked evoked responses in visual cortex

There were two types of stimulus-locked evoked responses --- those to the frequent standards and those to the infrequent targets. Evoked responses selective for detection of the target were isolated by subtracting responses evoked when the target was present and the subject pressed a button (hits) from those when the standard was presented and the subject did not press a button (correct rejects). As an illustration of the raw sensor data that forms the basis of our source analysis, we first show averaged data at the sensors for both Hit (black) and Correct Reject (red) trials (Figure 2b) for the orientation-defined figure appearance condition. In this condition, the display alternated between a uniform field (first half cycle) and one that was segmented on the basis of an orientation cue (second half cycle; see also Figure 1, orientation defined/figure appearance condition for stimulus depiction). On hits (81.6% of target presentations), an oval figure appeared, but on correct rejects the standard circular figures appeared. Figure 2b shows the averaged time courses beginning 250 ms before the transition to the segmented state. After presentation of the orientation transient that defined the figures, the response for both Hit and Correct Reject trials was the same, but starting at around

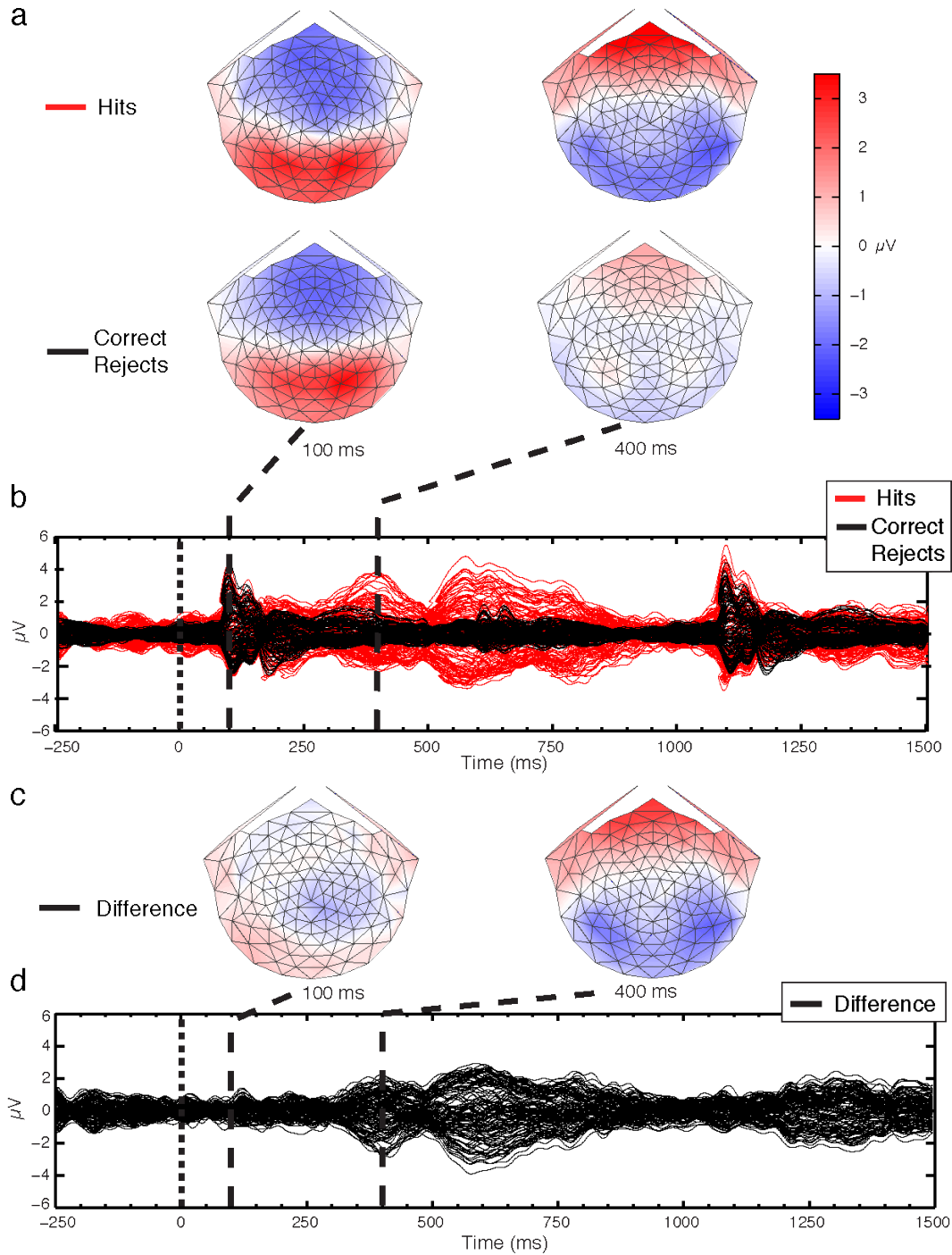


Figure 2.

Scalp topography and time courses for the orientation-defined figure appearance condition. a-b, Scalp topographies and time courses from Hit and Correct Reject trials. **c-d,** Scalp topographies and time courses for the difference between hits and correct rejects. Between -250 and 0 ms, the display was uniform. At 0 and 1000 ms, 9 orientation-defined disks appeared abruptly from the uniform background. The disks disappeared at 500 ms and the field was uniform between 500 and 1000 ms. Figure onset produced an initial transient response peaking at approximately 100 ms that was similar in amplitude and spatial distribution for hits and correct rejections. Initial target-selective activity (hits – correct rejects, shown in **d**) reached a peak at approximately 400 ms and was relatively more negative over lateral occipital electrodes, bilaterally.

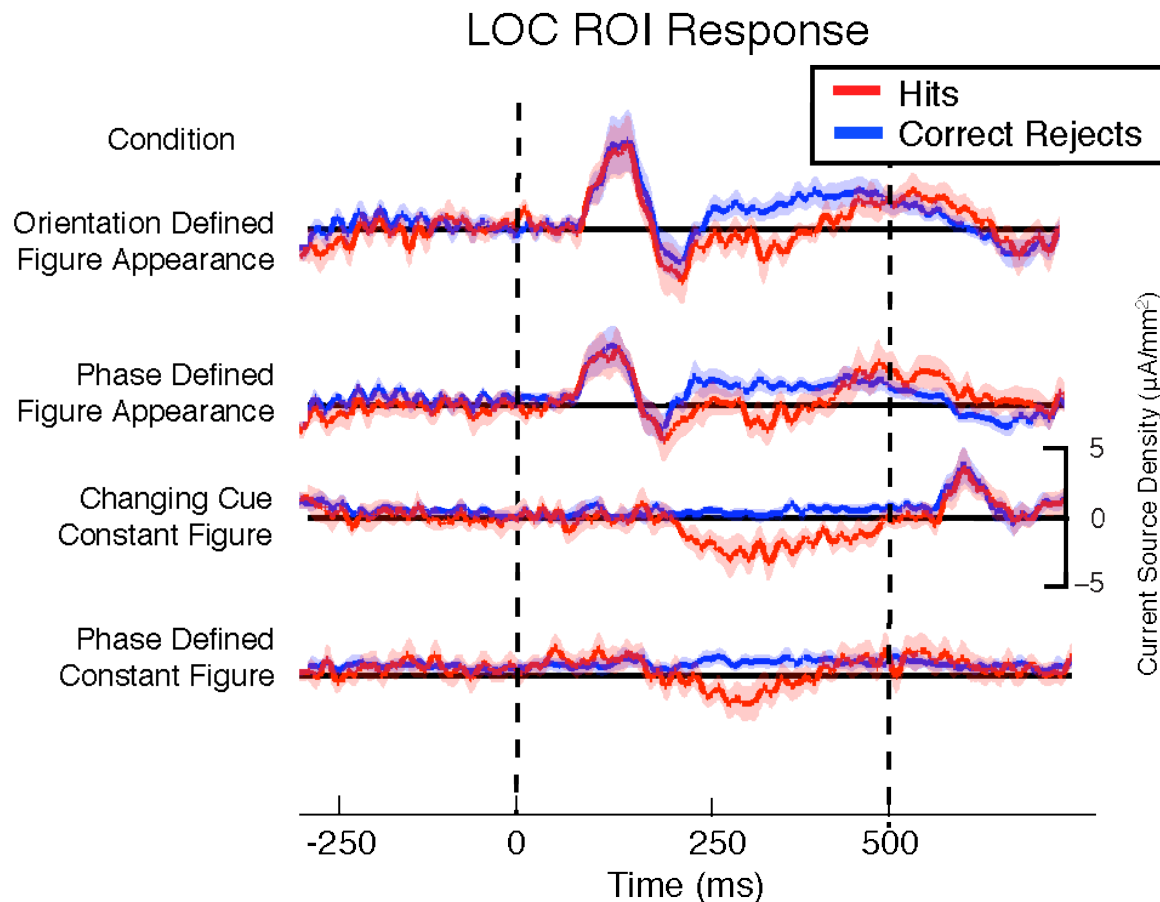


Figure 3.

Hit versus Correct Reject time courses for each stimulus condition in the LOC ROI. The time courses in the LOC show a strong transient response following the onset of a figure (top two rows). Differential activity in each condition comprises a relatively negative response beginning around 200 msec. The differential activity can be easily seen in the responses for the conditions that had a constant figure (bottom two rows). Since the figure was constantly present in these conditions no transient figure appearance response was evoked. The lightly shaded area around waveforms represents ± 1 standard error of the mean.

300 ms after figure onset, the two responses began to diverge (Figure 2d). The target-selective response peaks 400 ms after figure onset with a topography that is bilateral, occipital and more negative for hits than for correct rejects (see topographic map inset in Figure 2c). This relatively negative and bilateral occipital distributed differential activity on correctly detected targets differs from the widely study P300 ERP component which is positive and peaks at parietal electrodes (Polich and Kok, 1995).

In order to recover the brain regions responsible for the effects seen at the electrodes, we estimated the cortical current density distribution for each subject using a distributed inverse method. The current density distributions were averaged across participants with respect to individually defined functional and retinotopic ROIs determined via fMRI (see the "Methods" section). To illustrate subsequent analysis steps, we first focus on the responses to each of the four stimulus conditions as recorded from the LOC ROI, (see Figure 3). Following the onset of the elliptical target (0 ms), the response from the LOC ROI to hits (red trace) is

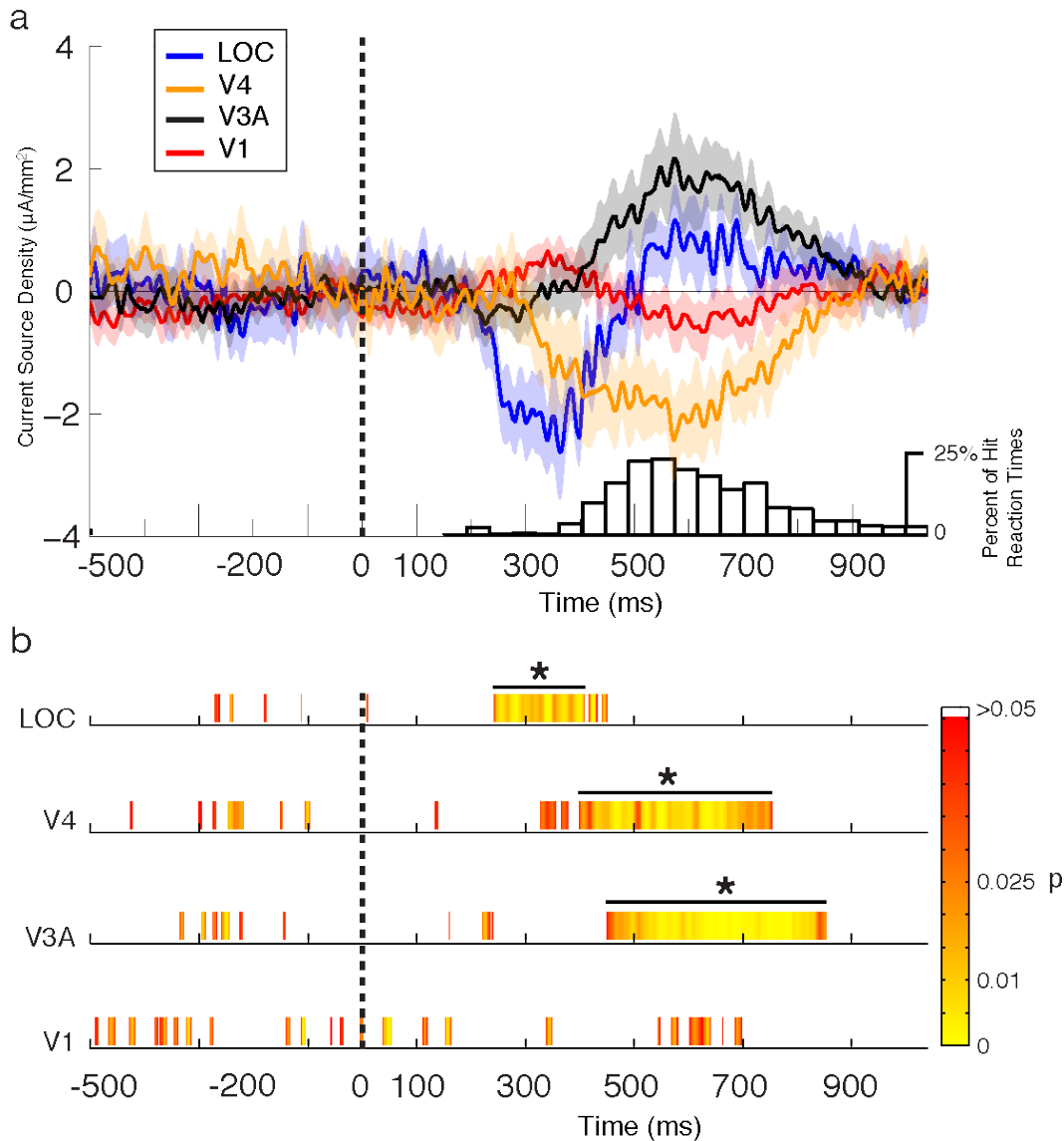


Figure 4

Target-selective time courses in visual ROIs. **a**, Stimulus-locked time courses for target-selective activity for each of four visual ROIs. Here target-selective activity was derived as the difference between Hit and Correct Rejection trial types collapsed over all 4 conditions. A histogram of the reaction times for all hits is plotted underneath the ROI waveforms. **b**, Times of significant differences between Hit and Correct Rejection trials. The colormap shows the uncorrected significance values, and black bars indicate periods that pass run-length correction for multiple comparisons. Target-selective activity is first significant in the LOC ROI, followed by the V4 and V3A ROI's in that order. No measurable target-selective activity was detected in the V1 ROI. The lightly shaded area around waveforms represents ± 1 standard error of the mean.

initially identical to the response to the correct reject of a standard (blue trace). The details of the early phase (roughly $< 250\text{ms}$) of the response vary greatly across the four stimulus conditions, ranging from a brisk multi-phasic response when an orientation transient is present (top two panels) to nothing when only the figural cue changes (bottom panel, Figure 3), but no

newly segmented figure appears. Because this early phase of activity is common to hits and correct rejects, it is not target-selective, and we do not consider it further.

Beginning at around 250 ms after a correctly discriminated target (hit) onset, the LOC response becomes relatively more “negative” (i.e. the current flow is directed from the surface inwards) than the response to the correct reject. Target-selective activity is similar across the four conditions and we therefore averaged across them to measure the common target-selective response and to increase SNR. In order to focus on activity that is specific for target presentations the analyses that follow will use the difference between hit and correct reject presentations. The difference between these presentations subtracts out the common, non-selective evoked responses.

The target-selective response varies in strength and time of emergence in the LOC, V4, V3A, and V1 ROIs, as shown in Figure 4. A histogram of the behavioral response latencies from all trials in all subjects is shown below the waveforms. The results of a permutation analysis to test for significant differences between hits and correct rejects are shown at the bottom (see Methods). Target-selective activity is first significant in the LOC ROI, followed by the V4 and V3A ROI's in that order. No significant target-selective activity was detected in the V1 ROI. A target-selective response is first seen in the LOC ROI, beginning around 250 ms after target presentation. The first statistically significant LOC activity started before 98.8% of button presses and ended before 96.4 % of button presses. Significant activity in LOC overlapped with 2.4% of button presses. The response in the LOC is followed by another relative negativity in the V4 ROI beginning at 310ms. Initial significant activity in V4 started before 97.1% of button presses, and ended before 29.3 % of button presses. Significant activity in V4 overlapped with 67.8% of button presses. The differential response in the V3A ROI is a relative positivity and begins at around 400 ms post-target. Initial significant activity in V3A started before 91.4% of button presses, and ended before 16.3 % of button presses. Significant activity in V3A overlapped with 75.1% of button presses. The V1 ROI exhibits only weak differential activity that is selective for the target presentation.

The target-selective activity recorded in the LOC and V4 ROIs may be related to the previously described “Selection Negativity” recorded at the scalp (Hillyard and Munte, 1984). Recordings in macaque V4 during a target detection tasks have shown a similar relative negativity in feedback recipient laminae (Mehta et al., 2000). In our recordings, activity in the LOC ROI precedes that recorded in the V4 ROI, consistent with the macaque results. We see no target-related activity in the V1 ROI, which may reflect the higher-order segmentation needed to perform the task. Given the resolution of our inverse procedure we can expect some crosstalk between nearby cortical regions. The ROIs chosen here are widely spaced on cortex and have been shown to have a low amount of crosstalk between them (see Lauritzen et al. (2010), supplement). The V1 ROI likely reflects some contribution from V2 and V3.

In a control experiment in 6 of the observers, we presented the same shape changes during trials in which attention was focused on a difficult letter discrimination task presented inside the central disk of the figure-array. In this manner we could determine if the responses described above depend on the allocation of volitional attention, or if they reflect a form of pre-attentive processing. Figure 5 shows the data from this experiment in the same format used for Figure 3. Differential target-related activity is not apparent when attention to the shape changes is not present.

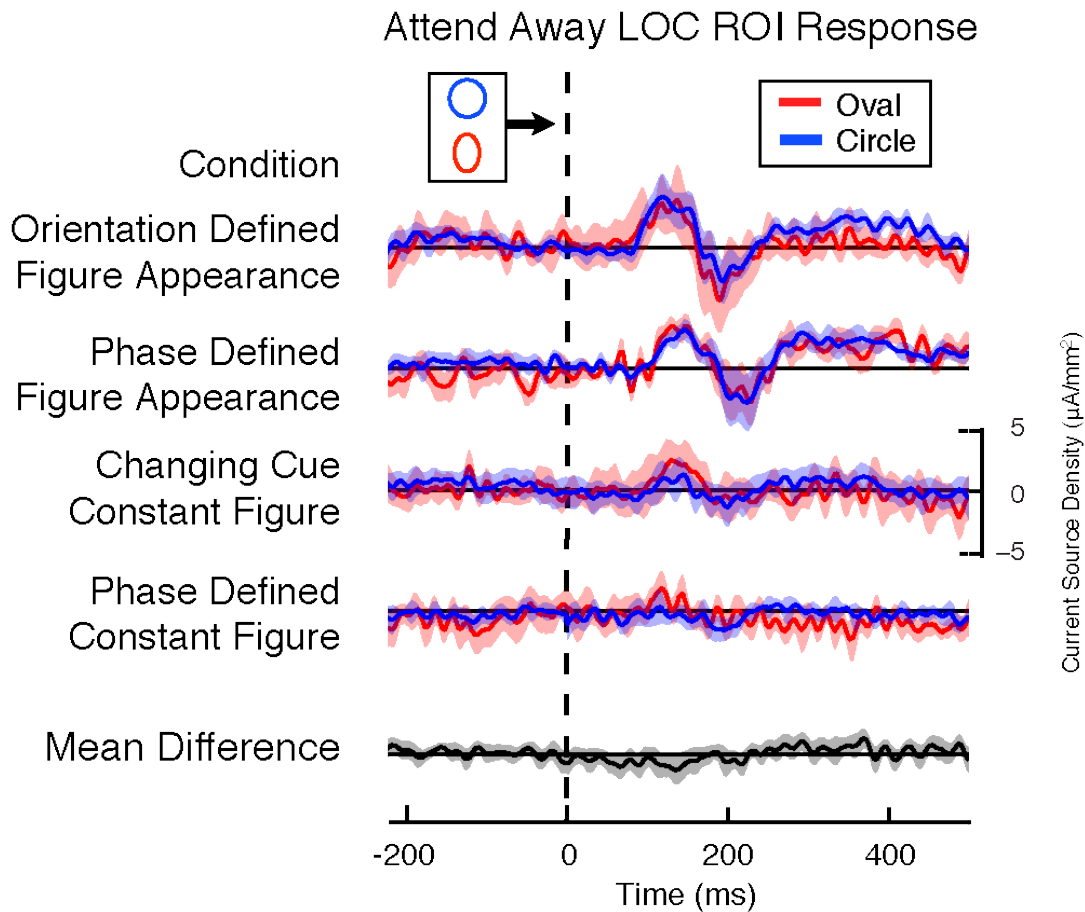


Figure 5

Attentional control task results: Response time courses when texture stimuli were task-irrelevant. The response in the LOC to circles (frequent) and ovals (infrequent) were recorded when attention was directed towards a letter discrimination task at the fovea. This figure is plotted using the conventions of Fig. 3, with the 4 sets of traces corresponding to the 4 separate stimulus conditions. The within-subject difference was collapsed over the four conditions and calculated identically as that done in Fig. 4 is shown at the bottom (Difference). This confirms that there is no difference in response between these stimuli when they are not task-relevant. The lightly shaded area around waveforms represents ± 1 standard error of the mean.

Whole-brain analysis of target-selective activity locked to stimulus onset

To localize regions outside of the functionally defined visual areas that were also active in our task, we made surface-based averages of the current density distribution. This approach aligns each individual participant's brain to each other using gyral and sulcal landmarks (Fischl et al., 1999). The surface-based alignment method is useful for regions of cortex for which we do not have functionally defined areas to compare across observers. Surface-aligned and averaged current source density distributions are shown in Figure 6a along with the separate waveforms emanating from the location with peak activation in each of 5 anatomically defined regions of interest. The surface averages indicate that the earliest stimulus-locked, target-selective responses occur in bilateral cortex in the region of the LOC (~254 ms). This finding from the surface average is consistent with the location of the functionally defined LOC across subjects, which is shown in Figure 6c as a cross-subject, surface-based average of the fMRI-

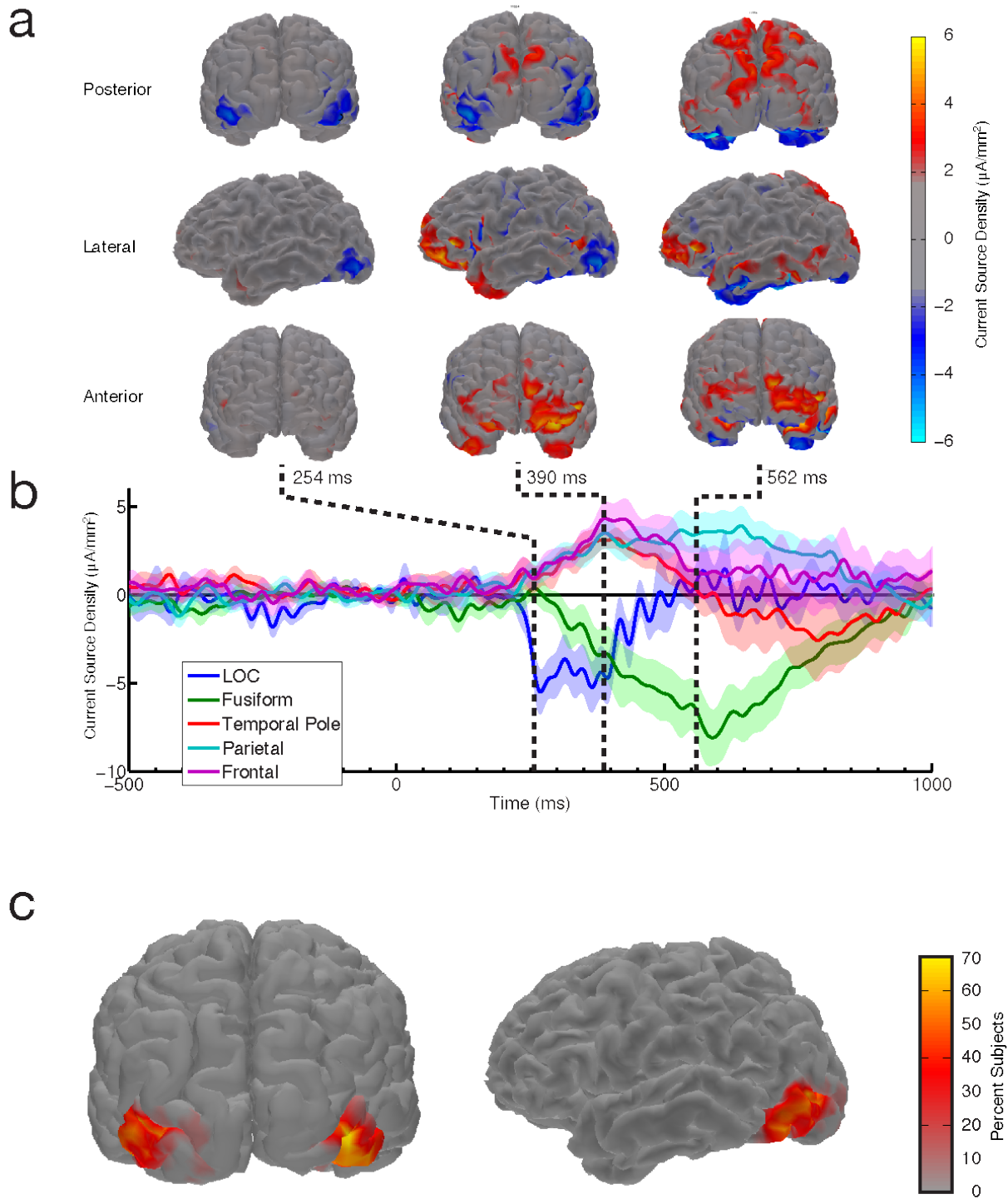


Figure 6

Whole-brain analysis of target-selective activity time-locked to the stimulus. a, Surface-based average of current density ($n=12$). **b**, Waveforms from LOC, fusiform, temporal pole, parietal and frontal locations. The earliest target-selective activity is found in the LOC, beginning before 254 ms. By 390 ms, target selective activity is widely distributed over all of the identified locations. By 562 ms, target selective activity has declined in the LOC but is sustained in the remaining locations. **c**, Cross-subject average of the fMRI definition of the LOC. Each location on the cortical surface shows the percentage of participants having the LOC ROI defined at that location. Note the close overlap between the fMRI-defined LOC ROI and the earliest target-selective evoked activity. The lightly shaded area around waveforms represents ± 1 standard error of the mean.

defined LOC ROI. By focusing on the earliest activity from a small change in a second-order stimulus, we isolate a focal region of lateral occipital cortex. The combination of sensitivity, selectivity and early latency exhibited by the LOC region we identify suggests that this activity represents the initial first step in fine-grained shape discrimination. There is also a strong sustained activation in the fusiform gyrus, and as time progresses, activity becomes evident in the IPS, the frontal lobe, the anterior part of the temporal lobe.

Response-locked evoked responses in visual cortex

The stimulus-locked analyses demonstrate that target-selective responses are tightly coupled to the time of target presentation. There was, however, considerable variability in the response time for target identification (see histogram below waveforms in Fig 4). In order to determine which cortical areas are more tightly associated with behavioral response generation, we calculated response-locked potentials by averaging the data from all Hit trials, aligning them to the time of the button press. This analysis preferentially emphasizes activity that bears a fixed timing relationship to the behavioral decision over activity that bears a fixed time relationship to the stimulus (DiCarlo and Maunsell, 2005).

Before describing the results of the response-locked analysis, we first consider a potential confound in this sort of analysis and then describe an analysis procedure that controls for it. Response-locked averages may also contain a residual neural activity from sustained stimulus-locked evoked potentials. If these responses are sufficiently long lasting, they can survive the blur introduced by the temporal jitter of response locking. We can conservatively estimate what portion of the data is not just blurred stimulus-related

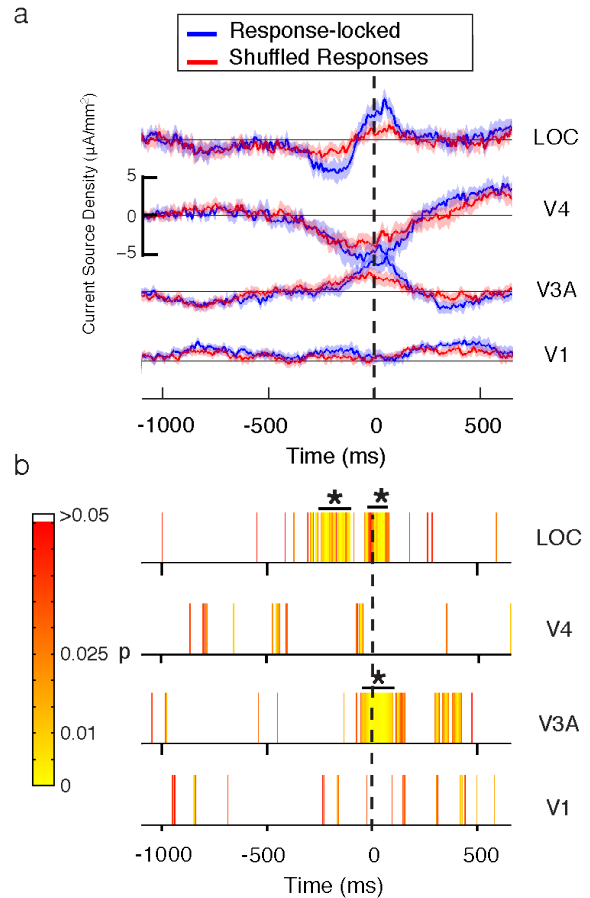


Figure 7.

Response-locked time courses in visual ROI's. **a**, Comparison of response-locked and shuffle corrected time courses. Blue curves plot the measured response-locked average; red curves plot the shuffled-time comparison time course. **b**, Latencies of significant differences as determined by run-length corrected (black bars) and uncorrected (yellow/red tic marks) permutation tests. The lightly shaded area around waveforms represents ± 1 standard error of the mean. Significant differences are present in the LOC ROI beginning approximately 300 ms before the button press, indicating that this region is associated with response generation. Significant differences are also present in the V3A ROI at the time of the button press.

activity, but actually response-locked by comparing the response-locked data to a surrogate dataset that contains the same amount of temporal blurring but without the true response-time alignment. The surrogate dataset was created by applying the same analysis as used for the true response-locked data, but instead of using the correct response times, we use random response times drawn from the subjects' actual response time distributions. That is, we permuted the trial labels for each response, destroying the actual response-time correspondence, but preserving the statistical temporal structure of the underlying evoked responses.

The true response-locked (blue) and shuffled response-locked (red) data are plotted in Figure 7 for the V1, V4, V3A, hMT+ and LOC ROIs. As indicated by the significant permutation differences depicted in the bottom panel, only the LOC and V3A ROIs contain activity that is significantly different between true response-locked and shuffled responses. Whereas the V4 ROI contains a period of relatively long sustained activity, this does not appear to be specific for the actual response latency as it is still present in the shuffled reconstruction. Importantly, only the LOC ROI contains response-time correlated activity that occurs before the subject's response. The response in this ROI begins around 300 ms before the subject's response. The V3A ROI response, while response-locked, occurs concurrently with the button press.

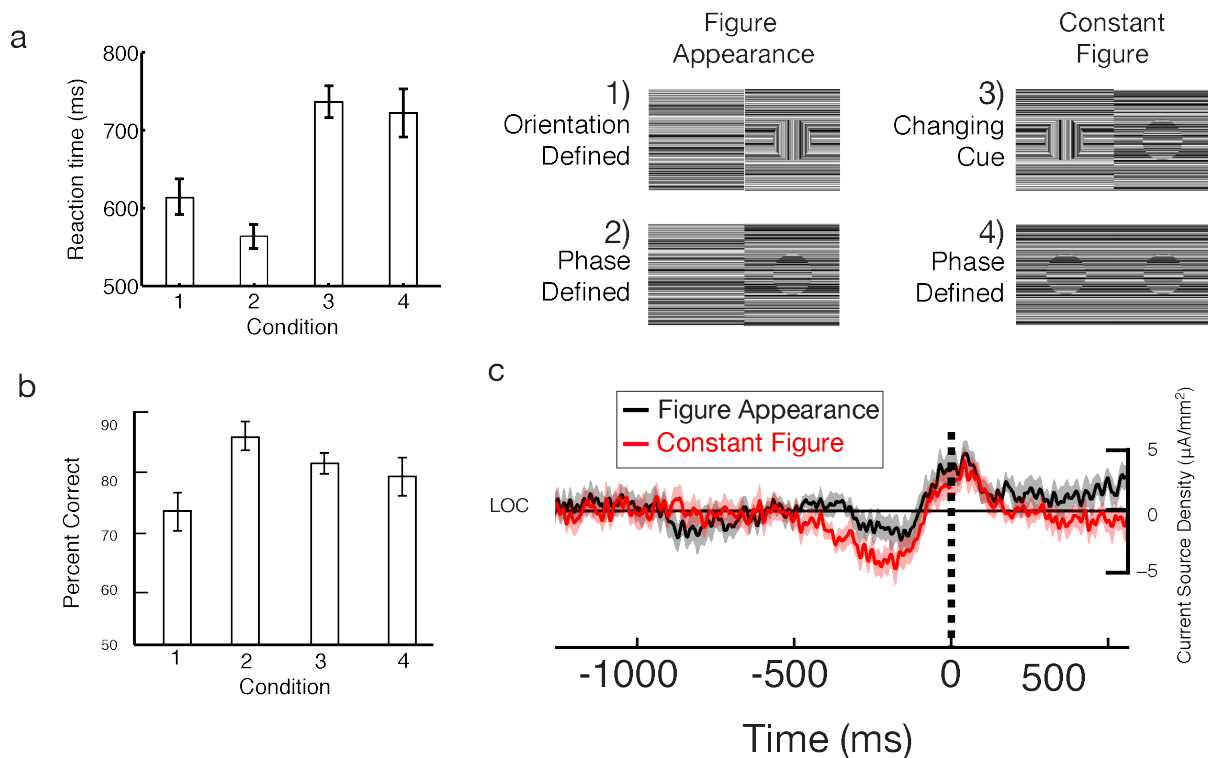


Figure 8

Reaction-time dependent activation in the LOC. **a**, Reaction times measured from the conditions illustrated in the upper left legend. Figure appearance stimuli have faster reaction times than the constant figure stimuli. **b**, Accuracy data shows that this is not caused by a simple speed/accuracy trade-off. **c**, Response-locked waveforms from the LOC for the figure appearance (black) and constant figure (red) stimuli the shaded region around the waveforms represents ± 1 standard error. The conditions with longer reaction times (constant figure) have a longer period of activation in the LOC, while the conditions with quicker reaction times (figure appearance) have shorter activation periods. This is consistent with the activation in the LOC tracking the processing bottleneck.

Using response time differences to localize the decision bottleneck

The stimuli used in this experiment can be divided into two classes: those in which the figure region appears and disappears (Figure Appearance: conditions 1 and 2), and those in which the figure is always present (Constant Figure: conditions 3 and 4). These two stimulus classes lead to different reaction times, and are therefore useful in discriminating responses that occur with a fixed latency before the motor response (motor preparation/execution) from those that are controlled by sensory processing leading to the decision. The stimuli that contain transient appearances from a uniform background have faster reaction times than those where the figure is always present (Figure 8a), even though accuracy was similar across conditions (Figure 8b). This may be because the appearance and disappearance of the figure region provides a timing cue to reduce the temporal uncertainty of the target, making the task less difficult. Alternatively, the appearance of the figure region can serve as an exogenous cue that directs attention to the figure. The important point of this difference, regardless of its source, is that it allows us to determine if the response-locked activations seen in the LOC also correlates with reaction time differences. As can be seen in Figure 8c, the response-locked activity in the LOC splits with reaction time. For the stimuli with longer reaction times, the LOC has a longer period of activation before the subject's response, while for the stimuli with quicker reaction times the LOC shows activation starting closer to the button press.

Whole-brain analysis of target-selective activity locked to the response

As we did for the stimulus-locked analysis of Figure 6, we also made response-locked averages using cross-subject, surface-based averaging. These results, shown in Figure 9, indicate that, within our range of sensitivity the LOC is the first region in the occipital lobe to show response-time predictive activity. This activity begins at about 300 ms before the button press with a time course that is consistent with the LOC response in Figure 7. At approximately the same latency, measurable responses are also present in the anterior temporal lobe, as well in the prefrontal cortex. As time progresses towards the eventual button push, activity in the temporal pole and frontal cortex increases until about 100 ms before the button push, but parietal lobe activation does not occur until less than 50 ms before the button push.

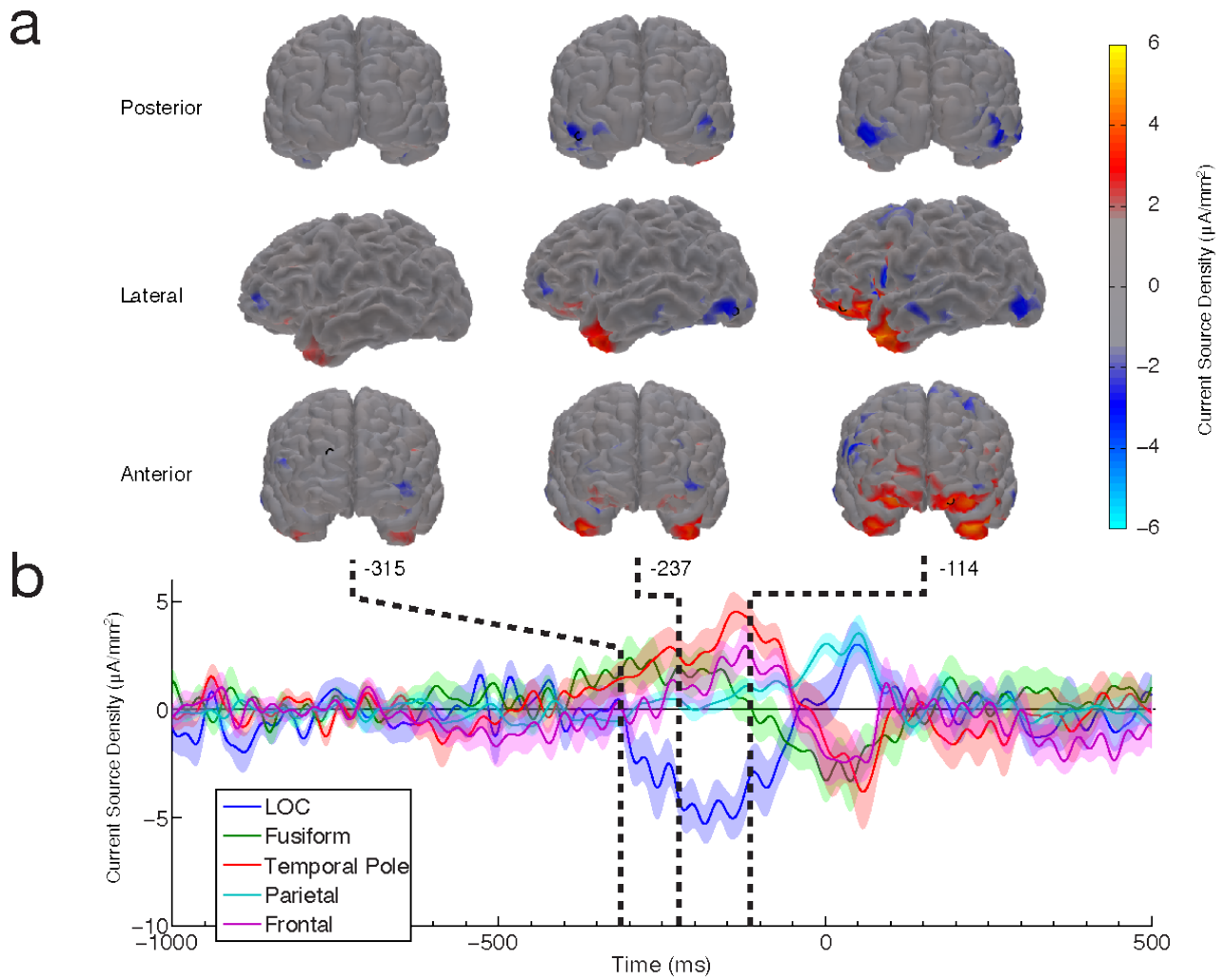


Figure 9

Whole-brain analysis of activity time-locked to the participants' motor response. **a**, Surface-based average of current density ($n=12$). **b**, Waveforms derived from the response maximum in the LOC, fusiform, temporal pole, parietal and frontal anatomically defined ROIs. Shaded region represents ± 1 standard error. The earliest response is found in the temporal pole, beginning at 315 ms before the response. By 237 ms before the response activity is present in the LOC as well. As time progresses to the eventual response activity increases in frontal, temporal pole, and LOC. Activity in the parietal region peaks around the time of response.

Discussion

In this study we used EEG source-imaging of decision-related responses evoked in a threshold-level shape discrimination task and found the following sequence of events: The first sign of target-related activity occurred at 250 ms after target onset in the LOC ROI. Target-specific activity was then found in V4, and parietal, anterior temporal and frontal cortex. This target-specific activity required attention to the task, as it was not present when the subject's attention was diverted. Behavioral responses to the target occurred at an average latency of 550 msec. If we assume a fixed motor delay of 80 ms (Roitman and Shadlen, 2002; Kiani et al., 2008) that leaves a transitional time of approximately 220 ms during which the decision to respond is made. By looking backwards in time from the moment of the behavioral response, we found that activity in LOC reliably reflected the decision time. Activity in LOC started sooner before the button press in conditions with longer average reaction times than in conditions with shorter reaction times. At the time of the earliest response-locked activity, frontal and anterior temporal regions were also active. We thus conclude that neurons in the human LOC contribute to the earliest stages of the decision processes underlying fine shape discrimination.

Tuning for shape

Much has been learned about the flow of visual information along the ventral stream, but it is still largely a mystery how these representations are utilized for shape discrimination. Micro-electrode studies in macaque have used simple luminance-defined, geometrical constructs and found neurons in IT that fired preferentially for related shapes (Gross et al., 1972; Desimone et al., 1984; Tanaka et al., 1991; Gallant et al., 1996) in a cue-invariant fashion (Sary et al., 1993). Neurons that are selective for object shape have also been found in V4 (Gallant et al., 1996; Pasupathy and Connor, 2002). These properties are consistent with our observations of shape-selectivity in LOC, the human homologue of IT (Denys et al., 2004).

Multiple lines of evidence implicate human LOC in shape processing. In fMRI studies the LOC is activated more by intact than by scrambled images of objects (Malach et al., 1995; Grill-Spector et al., 1998; Kourtzi and Kanwisher, 2000). Moreover, the LOC contains a representation of the perceived shape of objects (Kourtzi and Kanwisher, 2001; Vinberg and Grill-Spector, 2008). The shape representation in the LOC is also tuned for small shape changes in the outline of objects (Panis et al., 2008; Gillebert et al., 2009). Our previous source-imaging studies using frequency tagging have suggested that the LOC is tuned for the configuration of a small figure on a larger background (Appelbaum et al., 2010) and that LOC responses to figures are largely cue-invariant (Appelbaum et al., 2006; Appelbaum et al., 2012). Cue-invariance in LOC has also been found with fMRI (Grill-Spector et al., 1998)

From representation to decision

Our dynamic stimulus strongly activates many visual areas, from V1 to the LOC and beyond. Nevertheless, when contrasting responses between targets and distracters, we find that only a few visual areas have activity related to the shape discrimination. These areas include V4 and the LOC. Our previous results from the frequency-tagging paradigm indicated that task-relevant target-related processing selectively arises from the figure-region of texture-defined displays and not the background (Appelbaum and Norcia, 2009). As noted above, prior

research indicates that neurons in both V4 and the LOC are tuned for shape attributes. Because V4 is earlier in the anatomical hierarchy it might be expected that differential activity would arise first in V4, and then propagate to the LOC. Our results indicate, however, that the first visual region to exhibit a differential response for shape is the LOC. This activation is then followed by activity in V4, suggesting that LOC provides the initial substrate for shape discrimination. The stimulus representation in the LOC is thus both generic enough for representing a wide range of visual shapes and specific enough to distinguish between very subtle differences between shapes.

Shape discrimination involves a sequence of increasingly sophisticated processing operations that first signal local contrast, then local orientation, then local orientation discontinuity, then figure-ground assignment and finally the integration of discontinuity information along the entire boundary of the figure into a coherent representation of shape. Prior work using transient VEPs to study the segmentation of textures into regions has indicated that a differentiation between uniform and segmented stimuli occurs as early as 90-115 ms (Bach and Meigen, 1992; Lamme et al., 1992; Caputo et al., 1999; Scholte et al., 2006; Scholte et al., 2008). Texture-segmentation-related activity localizes to V1 (Lamme et al., 1992; Lamme et al., 1998; Appelbaum et al., 2012). Texture-segmentation-related activity is preceded by earlier activity driven by local contrast changes that is independent of segmentation status (Lamme et al., 1998; Marcus and Van Essen, 2002; Appelbaum et al., 2012). For the stimuli used here, texture segmentation is a necessary, but not sufficient, precursor for being able to perform the shape discrimination. As sensory processing advances the local feature representation becomes sufficiently rich as to support decision variables that code figure shapes. These in turn are communicated to motor regions responsible for executing the response. The target-related activity we observe in the LOC places an upper-bound on the timing of this activity at 250 ms.

By looking at how closely a neural response tracks the timing of the participants' reaction, one can selectively emphasize decision-related activity and work backward from the response to find the earliest activity that determines the behavioral response time. This procedure can also be used to determine where along the sensory-motor continuum a given cortical region lies (Di Carlo & Maunsell (2005) and we find that the LOC is as early as any other area on this criterion. The LOC thus appears to be critically involved in the transition from sensory representation of shape to the initial stages of decision. The best evidence for this in our data is comes from the observation *the LOC was the only sensory region we found that tracked between-condition differences in reaction time (cf Fig 8c). This result suggests that the LOC is the bottleneck for processing these stimuli.* Such a role is consistent with prior work in anterior inferotemporal (AIT) cortex of macaque. In a shape decision task, Di Carlo & Maunsell (2005) found that while a majority of neurons in AIT were locked to the sensory presentation, a smaller fraction exhibited response latencies that were significantly correlated with behavioral response times. In addition, the conditions we contrasted contain both a correct sensory discrimination as well as a differential motor response. We therefore cannot explicitly rule out contribution from a motor response-selection process. Given the above described shape selective properties of the LOC and the absence of published data showing motor response-selective units in the occipital visual areas we believe the conservative interpretation is that the LOC does not contribute to response selection. The transition between

representation and decision may occur in different neurons within LOC or within different subdivisions of the complex (Larsson and Heeger, 2006). Techniques with better spatial resolution than the present one will be needed to answer this question.

Relationship to previous Event-Related Potential (ERP) studies

Many previous ERP studies have used paradigms in which participants were tasked with detecting deviant targets among a stream of standards. These studies have frequently identified a target-selective parietal positive component, the P300 (Polich and Kok, 1995; Polich, 2007). Although the latency of our target-selective potential is similar to that of the typical P300 (250-500 ms) the scalp polarity is negative and the potential peaks at rather lateral sites. While the current task does evoke a later, parietal peaking activation, it is somewhat later (~ 560 ms) than the typical P300 range.

The visual mismatch negativity (MMN) is another ERP component sometimes observed in experiments with deviants presented among standards. The MMN, however, exists even when attention is directed away from the deviant stimulus (Heslenfeld, 2003; Pazo-Alvarez et al., 2003). From this definition it follows that the observed response in the present experiment is not a MMN because when the task is changed to direct attention away from the shape the observed negativity to deviant targets disappears (Fig. 5).

The properties of the observed target-selective evoked potential most closely resemble those of the Selection Negativity (SN), a member of the broad class of N2 potentials (Folstein and Van Petten, 2008). The Selection Negativity comprises a more negative response at the scalp when task-relevant features such as color, or spatial frequency are presented (Anllo-Vento and Hillyard, 1996; Hillyard and Anllo-Vento, 1998). Depth electrode recordings in macaque V4 have suggested that this structure contributes to the surface SN (Mehta et al., 2000). Mehta et al. (2000) found that the attentional modulation giving rise to the SN arose in feedback recipient layers of V4. In the present study we find that both V4 and the LOC contribute to the measured target selective activity that is locked to the stimulus. Because the LOC was the first cortical area to differentially respond to the presentation of the target stimulus we suggest that our V4 activity arises due to feedback from LOC.

The current study shows that among visually responsive occipital areas, the LOC shows the earliest latency differential activity for shape discriminations, followed closely by V4. The LOC is also the first occipital region to show response-locked activity predictive of the time of the behavioral response. These results together demonstrate the LOC is involved not only in encoding shape information, but also in making visual shape discriminations, and in feeding this information along to generate a visual perceptual decision.

References

- Ales JM, Norcia AM (2009) Assessing direction-specific adaptation using the steady-state visual evoked potential: results from EEG source imaging. *J Vis* 9:8.
- Anllo-Vento L, Hillyard SA (1996) Selective attention to the color and direction of moving stimuli: electrophysiological correlates of hierarchical feature selection. *Percept Psychophys* 58:191-206.
- Appelbaum LG, Norcia AM (2009) Attentive and pre-attentive aspects of figural processing. *Journal of vision [electronic resource]* 9:18 11-12.
- Appelbaum LG, Ales JM, Norcia AM (2012) The Time Course of Segmentation and Cue-Selectivity in the Human Visual Cortex. *PLoS One*.
- Appelbaum LG, Ales JM, Cottureau B, Norcia AM (2010) Configural specificity of the lateral occipital cortex. *Neuropsychologia* 48:3323-3328.
- Appelbaum LG, Wade AR, Vildavski VY, Pettet MW, Norcia AM (2006) Cue-invariant networks for figure and background processing in human visual cortex. *J Neurosci* 26:11695-11708.
- Appelbaum LG, Wade AR, Pettet MW, Vildavski VY, Norcia AM (2008) Figure-ground interaction in the human visual cortex. *J Vis* 8:8 1-19.
- Bach M, Meigen T (1992) Electrophysiological correlates of texture segregation in the human visual evoked potential. *Vision research* 32:417-424.
- Blair RC, Karniski W (1993) An alternative method for significance testing of waveform difference potentials. *Psychophysiology* 30:518-524.
- Brewer AA, Liu J, Wade AR, Wandell BA (2005) Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nat Neurosci* 8:1102-1109.
- Caputo G, Romani A, Callieco R, Gaspari D, Cosi V (1999) Amodal completion in texture visual evoked potentials. *Vision research* 39:31-38.
- Cichy RM, Chen Y, Haynes JD (2011) Encoding the identity and location of objects in human LOC. *Neuroimage* 54:2297-2307.
- Cottureau BR, Ales JM, Norcia AM (2011) Increasing the accuracy of electromagnetic inverses using functional area source correlation constraints. *Hum Brain Mapp*.
- Dale AM, Liu AK, Fischl BR, Buckner RL, Belliveau JW, Lewine JD, Halgren E (2000) Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* 26:55-67.
- Denys K, Vanduffel W, Fize D, Nelissen K, Peuskens H, Van Essen D, Orban GA (2004) The processing of visual shape in the cerebral cortex of human and nonhuman primates: a functional magnetic resonance imaging study. *J Neurosci* 24:2551-2565.
- Desimone R, Albright TD, Gross CG, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci* 4:2051-2062.
- DiCarlo JJ, Maunsell JH (2005) Using neuronal latency to determine sensory-motor processing pathways in reaction time tasks. *J Neurophysiol* 93:2974-2986.
- Fischl B, Sereno MI, Tootell RB, Dale AM (1999) High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp* 8:272-284.
- Folstein JR, Van Petten C (2008) Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology* 45:152-170.

- Gallant JL, Connor CE, Rakshit S, Lewis JW, Van Essen DC (1996) Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. *J Neurophysiol* 76:2718-2739.
- Gillebert CR, Op de Beeck HP, Panis S, Wagemans J (2009) Subordinate categorization enhances the neural selectivity in human object-selective cortex for fine shape differences. *J Cogn Neurosci* 21:1054-1064.
- Grill-Spector K, Kushnir T, Edelman S, Itzhak Y, Malach R (1998) Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron* 21:191-202.
- Gross CG, Rocha-Miranda CE, Bender DB (1972) Visual properties of neurons in inferotemporal cortex of the Macaque. *J Neurophysiol* 35:96-111.
- Hamalainen MS, Ilmoniemi RJ (1994) Interpreting magnetic fields of the brain: minimum norm estimates. *Med Biol Eng Comput* 32:35-42.
- Haushofer J, Baker CI, Livingstone MS, Kanwisher N (2008) Privileged coding of convex shapes in human object-selective cortex. *J Neurophysiol* 100:753-762.
- Heslenfeld DJ (2003) Visual Mismatch Negativity. In: *Detection of Change: Event-Related Potential and fMRI Findings* (Polich J, ed), pp 41-59. Norwell: Kluwer.
- Hillyard SA, Munte TF (1984) Selective attention to color and location: an analysis with event-related brain potentials. *Percept Psychophys* 36:185-198.
- Hillyard SA, Anllo-Vento L (1998) Event-related brain potentials in the study of visual selective attention. *Proc Natl Acad Sci U S A* 95:781-787.
- Huk AC, Dougherty RF, Heeger DJ (2002) Retinotopy and functional subdivision of human areas MT and MST. *J Neurosci* 22:7195-7205.
- Kendall DG (1984) Shape Manifolds, Procrustean Metrics, and Complex Projective Spaces. *Bulletin of the London Mathematical Society* 16:81-121.
- Kiani R, Hanks TD, Shadlen MN (2008) Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J Neurosci* 28:3017-3029.
- Konen CS, Kastner S (2008) Two hierarchically organized neural systems for object information in human visual cortex. *Nat Neurosci* 11:224-231.
- Kourtzi Z, Kanwisher N (2000) Cortical regions involved in perceiving object shape. *J Neurosci* 20:3310-3318.
- Kourtzi Z, Kanwisher N (2001) Representation of perceived object shape by the human lateral occipital complex. *Science (New York, NY)* 293:1506-1509.
- Lamme VA, Van Dijk BW, Spekreijse H (1992) Texture segregation is processed by primary visual cortex in man and monkey. Evidence from VEP experiments. *Vision research* 32:797-807.
- Lamme VA, Zipser K, Spekreijse H (1998) Figure-ground activity in primary visual cortex is suppressed by anesthesia. *Proceedings of the National Academy of Sciences of the United States of America* 95:3263-3268.
- Larsson J, Heeger DJ (2006) Two retinotopic visual areas in human lateral occipital cortex. *J Neurosci* 26:13128-13142.
- Lauritzen TZ, Ales JM, Wade AR (2010) The effects of visuospatial attention measured across visual cortex using source-imaged, steady-state EEG. *J Vis* 10.

- Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB (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* 92:8135-8139.
- Marcus DS, Van Essen DC (2002) Scene segmentation and attention in primate cortical areas V1 and V2. *J Neurophysiol* 88:2648-2658.
- Mehta AD, Ulbert I, Schroeder CE (2000) Intermodal selective attention in monkeys. II: physiological mechanisms of modulation. *Cereb Cortex* 10:359-370.
- Panis S, Vangeneugden J, Op de Beeck HP, Wagemans J (2008) The representation of subordinate shape similarity in human occipitotemporal cortex. *J Vis* 8:9 1-15.
- Pasupathy A (2006) Neural basis of shape representation in the primate brain. *Prog Brain Res* 154:293-313.
- Pasupathy A, Connor CE (2002) Population coding of shape in area V4. *Nat Neurosci* 5:1332-1338.
- Pazo-Alvarez P, Cadaveira F, Amenedo E (2003) MMN in the visual modality: a review. *Biol Psychol* 63:199-236.
- Polich J (2007) Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol* 118:2128-2148.
- Polich J, Kok A (1995) Cognitive and biological determinants of P300: an integrative review. *Biol Psychol* 41:103-146.
- Roitman JD, Shadlen MN (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J Neurosci* 22:9475-9489.
- Sary G, Vogels R, Orban GA (1993) Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science* 260:995-997.
- Scholte HS, Witteveen SC, Spekreijse H, Lamme VA (2006) The influence of inattention on the neural correlates of scene segmentation. *Brain research* 1076:106-115.
- Scholte HS, Jolij J, Fahrenfort JJ, Lamme VA (2008) Feedforward and recurrent processing in scene segmentation: electroencephalography and functional magnetic resonance imaging. *Journal of cognitive neuroscience* 20:2097-2109.
- Smith SM (2002) Fast robust automated brain extraction. *Hum Brain Mapp* 17:143-155.
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23 Suppl 1:S208-219.
- Summerfield C, Koechlin E (2008) A neural representation of prior information during perceptual inference. *Neuron* 59:336-347.
- Tanaka K, Saito H, Fukada Y, Moriya M (1991) Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J Neurophysiol* 66:170-189.
- Tomba T, Sary G (2010) A review on the inferior temporal cortex of the macaque. *Brain Res Rev* 62:165-182.
- Tootell RB, Hadjikhani N (2001) Where is 'dorsal V4' in human visual cortex? Retinotopic, topographic and functional evidence. *Cereb Cortex* 11:298-311.

- Vinberg J, Grill-Spector K (2008) Representation of shapes, edges, and surfaces across multiple cues in the human visual cortex. *J Neurophysiol* 99:1380-1393.
- Wade AR, Brewer AA, Rieger JW, Wandell BA (2002) Functional measurements of human ventral occipital cortex: retinotopy and colour. *Philos Trans R Soc Lond B Biol Sci* 357:963-973.
- Wahba G (1990) *Spline Models for Observational Data*: SIAM.
- Yeo BT, Sabuncu MR, Vercauteren T, Ayache N, Fischl B, Golland P (2010a) Spherical demons: fast diffeomorphic landmark-free surface registration. *IEEE Trans Med Imaging* 29:650-668.
- Yeo BT, Sabuncu MR, Vercauteren T, Holt DJ, Amunts K, Zilles K, Golland P, Fischl B (2010b) Learning task-optimal registration cost functions for localizing cytoarchitecture and function in the cerebral cortex. *IEEE Trans Med Imaging* 29:1424-1441.