



Published in final edited form as:

Neuroimage. 2015 October 15; 120: 285–297. doi:10.1016/j.neuroimage.2015.07.005.

Distinct Effects of Trial-Driven and Task Set-Related Control in Primary Visual Cortex

Joseph C. Griffis^{#1}, Abdurahman S. Elkhetal^{#2}, Ryan J. Vaden², and Kristina M. Visscher²

¹The University of Alabama at Birmingham Department of Psychology

²The University of Alabama at Birmingham Department of Neurobiology

These authors contributed equally to this work.

Abstract

Task sets are task-specific configurations of cognitive processes that facilitate task-appropriate reactions to stimuli. While it is established that the trial-by-trial deployment of visual attention to expected stimuli influences neural responses in primary visual cortex (V1) in a retinotopically specific manner, it is not clear whether the mechanisms that help maintain a task set over many trials also operate with similar retinotopic specificity. Here, we address this question by using BOLD fMRI to characterize how portions of V1 that are specialized for different eccentricities respond during distinct components of an attention-demanding discrimination task: cue-driven preparation for a trial, trial-driven processing, task-initiation at the beginning of a block of trials, and task-maintenance throughout a block of trials. Tasks required either unimodal attention to an auditory or a visual stimulus or selective intermodal attention to the visual or auditory component of simultaneously presented visual and auditory stimuli. We found that while the retinotopic patterns of trial-driven and cue-driven activity depended on the attended stimulus, the retinotopic patterns of task-initiation and task-maintenance activity did not. Further, only the retinotopic patterns of trial-driven activity were found to depend on the presence of intermodal distraction. Participants who performed well on the intermodal selective attention tasks showed strong task-specific modulations of both trial-driven and task-maintenance activity. Importantly, task-related modulations of trial-driven and task-maintenance activity were in opposite directions. Together, these results confirm that there are (at least) two different processes for top-down control of V1: One, working trial-by-trial, differently modulates activity across different eccentricity sectors—portions of V1 corresponding to different visual eccentricities. The second process works across longer epochs of task performance, and does not differ among eccentricity sectors. These results are discussed in the context of previous literature examining top-down control of visual cortical areas.

Corresponding Author: Joseph Griffis, joegriff@uab.edu, Address: 1719 6th Ave S, Birmingham, AL 35233, Telephone: 934-0497.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Visual Attention; Primary Visual Cortex; retinotopy; fMRI; Cognitive Control; Task Set

1. Introduction

Primary visual cortex (V1) is the first cortical stage in the visual information processing hierarchy (Felleman and Van Essen, 1991), and is organized so that posterior regions are most responsive to central vision and anterior regions are most responsive to peripheral vision. Despite its relatively basic role in the processing of visual information, neural activity in V1 is modulated by cognitive processes associated with top-down control, such as the deployment of attention to locations in space (e.g. Huk and Heeger, 2000) or to task-relevant visual features (e.g. Watanabe et al., 1998). The effects of attention on neural activity in V1 and other early visual areas commonly occur as cue-driven or expectation-driven shifts in neural activity (Murray, 2008; Chawla et al., 1999; McMains et al., 2007; Sylvester et al., 2009) that enhance responses in cortex corresponding to attended spatial locations (Kastner et al., 1999; Munneke et al., 2008) and suppress responses in cortex corresponding to ignored spatial locations (Sylvester et al., 2008). Such attention-related shifts in activation persist even during delay periods when visual stimuli are absent but attention is still deployed (Silver et al., 2007), and influence the magnitude of subsequent neural responses evoked by visual stimuli (Sylvester et al., 2009; Cardoso et al., 2012). Further, directing attention to different features of a stimulus (Runeson et al., 2013) alters activation in hierarchically higher visual areas that contain neurons that are specifically tuned to the attended feature. Thus, top-down modulations of activity in visual cortex serve to facilitate behaviorally relevant visual processing.

Attention-related modulations of neural activity in V1 likely reflect top-down signals from frontal and parietal cortex that facilitate the processing of cued or expected visual stimuli (Kastner et al., 1999; Hopfinger et al., 2000; Bressler et al., 2008; Vossel et al., 2012; Liu et al., 2014). The “fronto-parietal” control network (including, e.g. frontal eye fields and intraparietal sulcus) has been implicated as a primary source of attentional control that operates at the level of individual trials and adapts to real-time feedback about performance (Dosenbach et al., 2007). Control mechanisms originating in other regions (e.g. regions associated with the “cingulo-opercular” control network, including dorsal anterior cingulate and frontal operculum) may also contribute to task-specific responses in early visual areas (Dosenbach et al., 2007; Ebisch et al., 2013). This network is thought to operate on a slower timescale than the fronto-parietal network and is likely involved in implementing and maintaining task sets – task-specific configurations of cognitive processes that enable stimuli to be processed and responses to be generated according to the rules and associations inherent to a given task (Sakai, 2008). Thus, trial-by-trial and task-set related control of visual processing may be implemented via distinct processes that operate at different temporal resolutions.

Functional MRI (fMRI) experiments can distinguish effects that occur at different timescales (Visscher et al., 2003; Dubis et al., 2011). For example, experiments employing mixed

blocked/event-related designs to study control-related neural signals that occur during different aspects of task performance indicate that task-dependent changes in transient, trial-driven neural responses and sustained shifts in background neural activity likely correspond to distinct cognitive processes related to processing the presented stimuli and maintaining a task set, respectively (e.g. Braver et al., 2003; Burgund et al., 2006; Velanova et al., 2003, Dosenbach et al., 2006). Accordingly, recent experiments by our lab indicate that task-set-related and trial-related modulations differ even in V1. In a previous study, participants performed discrimination tasks on unimodal stimuli (visual or auditory) or bimodal stimuli (intermodal attention to only the visual or auditory portion of simultaneously presented visual and auditory stimuli). We found that trial-driven activity in V1 depended on presented stimuli, while activity related to the initiation and maintenance of task sets depended instead on the attended modality of the task regardless of presence or absence of intermodal distractors (Elkhetali et al., 2015). This suggests that even at the level of primary visual cortex, task-appropriate visual processing may be accomplished by the joint action of temporally distinct top-down control processes that differ in their stimulus-dependence vs. task-dependence.

Trial-driven vs. task set-related changes in neural activity likely reflect temporally distinct mechanisms for the control of visual processing in V1. It might be expected that they would also differ in their spatial specificity. The topography of cue-driven and trial-driven activity in V1 has been reliably shown to depend on target location (e.g. Sylvester et al., 2008). However, it is unclear whether activity related to task set initiation and maintenance also shows such retinotopic specificity. Characterizing the retinotopic patterns of task set-related changes in activity in V1 will allow us to better understand the mechanisms at work and how they influence visual processing.

In our previous study, we observed that trial-driven and task-set related responses in V1 varied in different ways depending on a participant's task (Elkhetali et al., 2015). Task set-related activity in V1 was greater during visual attention than auditory attention for a task with simultaneous visual and auditory stimuli, whereas this was not true for trial-driven activity. In fact, trial-driven effects were in the opposite direction, and whole brain analyses suggested that the trial-driven effect differed for sectors in V1 corresponding to central vs. peripheral vision (Figure 2, Elkhetali et al., 2015).

We reasoned that the observed trial-driven effects might result from the presence of attentional suppression in task-irrelevant portions of V1 during attention to vision. If such suppression were released during attention to audition, this might result in more widespread, retinotopically non-specific activation patterns during attention to audition. Such an effect would be expected based on previous reports that the suppressive effects of attention on activity in V1 are released during attention to an auditory stimulus (Ciaramitaro et al., 2007). If this were the case, we hypothesized that the observed differences in magnitude between trial-driven and task-maintenance activity might reflect differences in their effects on task-relevant vs. task-irrelevant eccentricity sectors during the bimodal tasks. We further hypothesized that task-maintenance activity might either lack retinotopic specificity, or show a specific retinotopic pattern that is task-invariant. In either case, only the magnitude of task-maintenance activity would change between tasks, with the retinotopic pattern

staying the same. To summarize, our observations suggested that 1) there are (at least) two control processes whose effects can be observed within V1—one with a trial-by-trial temporal profile, and the other which has a slower temporal profile, affecting V1 at the level of the task set (Elkhetali et al., 2015), and 2) these two control processes may act on V1 at different spatial scales.

In this experiment, in order to characterize the retinotopic effects of these two control processes in V1, we divided V1 into separate, smaller regions of interest based on retinotopy. We analyzed previously collected fMRI data that were obtained as participants performed an attention-demanding discrimination task (described fully in Methods). This task involved four conditions in which attention was directed to an auditory stimulus presented alone (auditory unimodal, AU), a centrally presented visual stimulus presented alone (visual unimodal, VU), an auditory stimulus paired with a task-irrelevant centrally presented visual stimulus (auditory bimodal, AB), or a centrally presented visual stimulus paired with a task-irrelevant auditory stimulus (visual bimodal, VB). Importantly, the bimodal conditions involved identical stimuli, and differed only in instructions for which stimulus modality to attend. This design allowed us to control for stimulus features and general attentional deployment while examining the effects of task instructions on the retinotopic patterns of activity in V1 during time periods corresponding to cue-driven preparation, trial-by-trial processing, task set initiation, and task set maintenance.

2. Materials and methods

2.1 Participants

Twenty healthy right-handed participants took part in this study. Participants were 8 males and 12 females with a mean age of 26 years (range 19-32 years) who had normal hearing and normal or corrected-to-normal vision. Participants were recruited through a campus wide advertisement. Recruitment procedures adhered to ethical standards as set and reviewed by the IRB at the University of Alabama at Birmingham. All participants provided a written consent prior to admission to the study. The study consisted of a total of 3 sessions, with a first session for behavioral measurement and 2 subsequent MRI sessions.

2.2 Task

During the fMRI experiment, participants performed an attention demanding discrimination task in which they had to correctly discriminate between two successive auditory or visual stimuli, schematized in Figure 1A. The stimuli and task parameters are described in detail in a previous publication (Elkhetali et al., 2015), and thus will only be briefly described here. During the Auditory Unimodal (AU) and Visual Unimodal (VU) tasks, the auditory or visual stimuli were presented alone. During the bimodal conditions, both the auditory and visual stimuli were presented simultaneously and the participant discriminated between the stimuli of only the cued modality. During these Auditory Bimodal (AB) and Visual Bimodal (VB) tasks, the unattended stimuli followed a random pattern so that the participant could gain no advantage by paying attention to the irrelevant stimulus. While data from the unimodal tasks were analyzed in this experiment to characterize the effects of attention to each stimulus modality on activity in V1, the bimodal tasks were of primary interest for the

purpose of addressing our hypothesis. Because the stimuli were identical for both the AB and VB tasks, changes in activity between the two tasks cannot be attributed to differences in stimuli and are likely to result from task-related control factors.

The participants were instructed to keep central vision fixed on the location of the fixation mark in the middle of the screen throughout the task. In order to monitor compliance with these instructions (and to confirm participants did not adopt a strategy of, e.g., closing eyes during presentation of irrelevant visual stimuli), participant's eye movements were monitored during the experiment using an Eyelink 1000 fMRI eye tracking system (SR Research Ontario, Canada). Eye position was calibrated at the beginning of each run, and monitored throughout.

Trials contained two successive stimuli of the target modality that were either identical or different. The simple auditory or visual stimuli used here were employed in a previous study examining psychophysical and short-term memory characteristics of these stimuli (Visscher et al., 2007). Auditory stimuli varied sinusoidally in time and tone. Stimuli that were "different" were modulated with different temporal frequencies, while the "same" trials contained exactly the same temporal frequencies. Visual stimuli varied sinusoidally in luminance over space. The stimuli were gray-scale horizontal gratings referred to as "Gabor patches" and were presented centrally. Visual stimuli that were "different" varied from each other in the width of the gratings, while the "same" trials contained the exact same grating width. The Gaussian window defining the contrast of the bars in the Gabor patch had a standard deviation of 2.7 degrees visual angle. Moving from the center (highest contrast) to the edges (lowest contrast) of the Gabor, the stimuli become less visible. At 5.4 degrees visual angle, the image presented had exactly zero contrast (that is, the values at those pixels were constrained to be all gray). This edge was not easily detectable by the participants. Thus the visible portion of the Gabor ended between 2.7 to 5.4 degrees eccentricity.

Four different cues were used to indicate the task type and modality of the upcoming task. A small, white, centrally presented fixation cross remained on screen during runs when no other stimuli or cues were present. The cues appeared at the location of the fixation cross and were small and of similar luminance, in order to minimize sensory bottom-up processing in response to the cue. A blue circle indicated the visual unimodal task while a blue circle with a cross within it indicated the visual bimodal task. A yellow cross indicated the auditory unimodal task, while a yellow circle with a cross within it indicated the auditory bimodal task.

The timeline of a trial in the event-related version of the task is schematized in Figure 1. First, the cue was presented, indicating the task type. The cue was followed by two stimuli, each with a duration of 500 ms. The two stimuli were separated by a mask for 500 ms. For the auditory stimuli, the mask was white auditory noise, filtered to include similar temporal frequencies to the range of auditory stimuli. For the visual stimuli, the mask was a white visual noise pattern filtered to include spatial frequencies similar to the range of frequencies of the visual stimuli. A question mark replaced the fixation cross during the two seconds during which the participant could make a response. In order to standardize the difficulty of the task across participants, each participant's just noticeable difference (JND) threshold for

auditory and visual stimuli were measured prior to the scanning sessions. Thresholds were defined using the QUEST algorithm (Watson and Pelli, 1983) as the stimulus difference (in % difference between two stimuli) at which participants could correctly discriminate two sequentially presented 500 ms duration stimuli 70% of the time, and were measured independently for auditory and visual stimuli as previously described (Elkhetali et al., 2015).

2.3 MRI data acquisition

Twenty participants performed the auditory and visual discrimination tasks in a 3T Siemens Allegra fMRI scanner. Whole-brain BOLD-weighted images were obtained with a TR of 2 s, TE of 30 ms, and a voxel size $3.75 \times 3.75 \times 4 \text{ mm}^3$. The visual stimuli were presented using a rear projection screen located outside of the magnet bore. The screen was visible through an angled mirror attached to the head coil that was placed above the participant's eyes. The auditory stimuli were delivered to the first 7 participants through MR safe Etymotics ER 30 earphones, with additional MR compatible ear protectors. However, due to participant discomfort with the in-ear devices, their use was substituted with auditory stimuli fed through the Siemens sound system via specialized headphones for the final 13 participants (there was no significant difference in performance between the two earphone models used).

Scanning time was divided into two sessions of about 2 hours each, performed on different days no more than 2 weeks apart. At the beginning of each session, an anatomical MPRAGE scan was taken of a participants' brain, producing an image with a voxel size of $1.0 \times 1.0 \times 1.1 \text{ mm}^3$. Retinotopic scans were performed during the first session. During the remainder of the first session and all of the second session, participants performed the tasks described above in alternating runs following an event-related design, and a mixed blocked/event-related design.

2.4 Event-related design

We used an event-related design for 8 of the runs (Rosen et al., 1998) to measure activity in response to transient cues independently from activity in response to cues followed by a trial (Figure 1B). Twenty five percent of cues were not followed by the stimuli to be discriminated (cue only trials). Both the cue only and cue+stimulus trials were presented in a randomized order. The trials were presented with varying intervals to allow for jittering of the signal. The length of the intervals varied between 0 and 5 TRs, and was determined in a pseudorandom manner. Between trials, a cross was presented in the center of the screen. Participants were instructed to look at this fixation mark in the center of the screen throughout each run. Each run was 276 TRs (552 s) long. We will refer to these runs as "cued runs." The event-related design allowed us to measure activity associated with switching task set, as task type changed from trial to trial. To measure sustained activity associated with maintaining a task set over a period of time we used a separate design called the mixed block/event-related design.

2.5 Mixed blocked/event-related design

We used a mixed blocked/event-related design for an additional 8 runs to measure trial-driven activity, task-maintenance activity, and task-initiation activity (Braver et al., 2003; Donaldson et al., 2001; Dosenbach et al., 2006; Velanova et al., 2003; Wenger et al., 2004).

This design involves the presentation of identical trials jittered in time during a block, as shown in Figure 1C. The length of the inter-trial intervals between the start of successive 4-second-long trials was between 2 and 7 TR's (4 to 14 s), and was chosen in a pseudorandom manner. This design allowed us to use a general linear model to estimate different components of fMRI activity (Visscher et al., 2003). There were a total of 8 runs, each lasting 251 TRs (502 s). Each run contained 5 blocks lasting 35 TR's each (70 s). There were 8 trials per task block, and between blocks there were intervening rest periods of 12 TR's (24 s).

2.6 Preprocessing steps

The BOLD images were preprocessed using MATLAB scripts using SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK; <http://www.fil.ion.ucl.ac.uk/spm/software/spm8/>). Images were slice time corrected, realigned and re-sliced. The images were resampled to 2 mm isotropic voxel dimensions, normalized to an EPI template using rigid body translation and rotation and then smoothed using a 5 mm kernel. The Matlab toolbox Artrepair (Mazaika et al., 2005) was used to minimize artifacts caused by movement. Artrepair replaced images in which more than .5 mm/TR of movement occurred with an interpolated image made from adjacent images. Runs were excluded from the analysis if more than 16 images total or 6 consecutive images had movement above the threshold of 0.5 mm. Runs were additionally excluded if the displacement ever reached more than 3 mm, that is, if the difference between any two timepoints was greater than 3 mm. Each participant included in the analysis had at least 4 cued and 4 blocked runs that met our strict criteria.

2.7 GLM analysis

BOLD images were analyzed using SPM8 based on the general linear model (Friston et al., 1994). Data for cued runs and blocked runs were analyzed separately. For the blocked runs, as shown in Figure 1C, we modeled trial-driven effects, sustained effects, block transition (task onset/offset) effects. Detailed methods for analysis and validation that these different time-courses can be independently estimated can be found elsewhere (Dosenbach et al., 2006; Visscher et al., 2003). A finite impulse response (FIR) model was used to model cue-driven activity, trial-driven activity, and task-initiation activity. The FIR model does not require the hemodynamic response to fit a particular shape and is considered the most flexible model that can be applied using a GLM (Lindquist et al., 2009), and is the least likely to result in the mis-attribution of trial-driven and task-maintenance effects (Visscher et al., 2003). Trial-driven effects were modeled with 12 regressors in a finite impulse response (FIR) model, representing 24 seconds following stimulus presentation. The peak magnitude timepoint for all effects that were modeled using the FIR model was defined as the timepoint with the greatest mean absolute difference from time-point zero (effect onset) across the four different task conditions as previously described (Elkhetali et al., 2015). For example, the magnitude of trial-driven effects in the mixed blocked event-related design condition was defined as the FIR regressor for the peak time (at time 8s) minus the start (at time 0s), as illustrated in Figure 3.

Task-initiation effects were modeled using 7 FIR regressors representing the first 14 seconds after the start of the block. The magnitude of task-initiation effects was similarly defined as

the peak (8s) minus the start (0s). We also modeled effects time locked to the end of a block (“task-termination”), representing the 12 FIR timepoints after the end of a block, but did not examine these closely, treating them as regressors of no interest. Task-maintenance effects were modeled using a single boxcar-shaped regressor that started 16 seconds from the beginning of the block (immediately following the task-initiation regressors) and ended at the end of a block. The regressor for task-maintenance activity was not convolved with any canonical hemodynamic response function because it represents a stable shift in baseline. The task-initiation and task-termination effects modeled the hemodynamic lag in the rise and fall of the signal.

The event-related design runs were analyzed using a separate GLM that independently modeled cue-only trials and cue+stimulus trials, as shown in Figure 1B. Each trial type was modeled with 12 regressors in a finite impulse response model (Ollinger et al., 2001). The peak magnitude timepoint for all effects that were modeled using the FIR model was defined as the timepoint with the greatest mean absolute difference from time-point zero (effect onset) across the four different task conditions as previously described (Elkhetali et al., 2015). The magnitude of cue-driven and trial-driven effects was defined as the peak (6s) minus the start (0s). This peak occurs slightly earlier than the peak for the trial-driven effects in the other runs simply due to slight differences in trial presentation. Note that for the trial-driven activity measured using the event-related design there was a cue that occurred before presentation of the stimulus on each trial. Neural activity in response to that cue certainly contributes to the response (which is why we treat those trial-driven signals differently from the trial-driven signals measured using the mixed blocked event-related design). Thus the peak response on those trials is a little bit earlier than in the blocked design, where the cue is present throughout the run. Similarly, the cue-only trials were much shorter than the full trials (which lasted 4 seconds), and therefore the cue-driven responses peak a bit earlier than the trial-driven responses from the mixed blocked event-related design.

2.8 Retinotopy

Each participant’s visual cortex was retinotopically mapped using standard methods (Warrnking et al., 2002). Briefly, during three BOLD scans, counter phase flickering checkerboard stimuli were presented to participants as wedges rotating clock wise and counterclockwise, and as contracting circles, all with periods of 24 s, for 10 cycles each. Wedges and circles extended to a maximum of 8.13 degrees eccentricity. We used the data from these scans to create an eccentricity map and a polar angle map for each participant, allowing assessment of the portion of the visual field to which a given voxel was most responsive. The retinotopic maps were generated using the Freesurfer retinotopy processing stream (Dale et al., 1999). Regions of interest were created manually for each individual participant using Freesurfer image analysis suite, which is documented and freely available for download online (<http://surfer.nmr.mgh.harvard.edu/>). The center of the first ROI was manually placed at the most retinotopically central freesurfer ‘vertex’ in the polar part of left hemispheric V1 along the calcarine sulcus, at ~1 degree visual angle. The ROI was ‘dilated’ using freesurfer tools, encompassing the central vertex and all vertices it touched. This dilation was repeated a total of 3 times to create a polar V1 region of interest for each participant like that on the far left in Figure 1D (labeled ROI 1). The subsequent 4 ROIs

were created analogously, with the central vertex placed along the calcarine sulcus at intervals of approximately 1.63 degrees visual angle, as defined by that participant's eccentricity map. Two more ROIs were created manually inside of the calcarine sulcus outside of the retinotopic mapping of V1 (Figure 1D, regions 6 and 7). Although these regions extended beyond the mapped region, we are confident that these regions correspond to peripheral V1 based on work showing that the retinotopic organization of striate cortex is well predicted by surface topology (Benson et al., 2012). Thus, our analyses focused on seven ROIs spanning from central to early peripheral regions of V1, with the first ROI corresponding to ~1 degree visual angle, the second ROI corresponding to ~2.6 degrees visual angle, the third ROI corresponding to ~4.2 degrees visual angle, the fourth ROI corresponding to ~5.8 degrees visual angle, the fifth ROI corresponding to ~7.4 degrees visual angle, and the last two ROIs corresponding to > 7.5 degrees visual angle. Thus, seven ROIs were created for each individual participant according to their retinotopy (an example from a single participant is shown in Figure 1D). These ROIs were then transformed to MNI space in order to conduct the group-level analyses.

3. Results

3.1 Behavioral results

Accuracy was similar for all four behavioral conditions and for both cued runs (Mean Percent Correct +/- SD: AU: 84.20 +/- 11.34 VU: 82.02 +/- 8.87 AB: 83.11 +/- 9.97 VB: 82.76 +/- 9.47) and blocked runs (Mean Percent Correct +/- SD: AU: 80.97 +/- 12.8; VU: 82.54 +/- 9.01; AB: 82.80 +/- 11.11; VB: 81.85 +/- 10.56). A two-way ANOVA on accuracy data during the blocked runs with factors attended modality (Auditory vs. Visual) and number of stimuli (unimodal vs. bimodal) showed no effect of attended modality ($F_{1,19}=0.010$, $p=0.9207$) or number of stimuli ($F_{1,19}=0.101$, $p=0.7542$). The results were similar for the cued runs: no effect of attended modality ($F_{1,19}=0.0136$, $p=0.7159$) or number of stimuli ($F_{1,19}=0.048$, $p=0.8294$). The absence of a difference in accuracy between conditions implies that the auditory and visual tasks were relatively similar in difficulty.

A two-way ANOVA with factors of attended modality (Auditory vs. Visual) and number of stimuli (unimodal vs. bimodal) showed that there was a significant difference in reaction time between auditory and visual tasks for both the cued ($F_{1,19}=14.059$, $p=0.0014$) and blocked ($F_{1,19}=15.54$, $p<0.001$) trials, but there was no significant difference due to number of stimuli (unimodal vs. bimodal) in either the cued ($F_{1,19}=1.883$, $p=0.1859$) or blocked runs ($F_{1,19}=1.883$, $p=0.1893$). Only correct trials were considered for this analysis. Longer reaction times for the auditory tasks may reflect the fact that the auditory stimuli varied along a temporal dimension, which may have caused participants to require more time to perceive the ripple sounds' frequency.

3.2 Mixed blocked/event-related design results

Results for task-initiation activity are shown in Figure 2. For task-initiation activity, a three-way repeated measures ANOVA with factors of attended modality, number of stimuli, and ROI revealed a significant main effect of attended modality ($F_{1,19}=6.89$, $p=0.02$) and a

marginally significant effect of ROI ($F_{6, 114} = 2.22, p = 0.05$), but did not reveal significant interactions of attended modality and number of stimuli ($F_{1,19} = 2.04, p=0.17$), attended modality and ROI ($F_{6,114} = 0.74, p=0.62$), number of stimuli and ROI ($F_{6,114} = 0.63, p = 0.71$), or attended modality, number of stimuli, and ROI ($F_{6,114} = 0.79, p = 0.58$).

Main effect contrasts were performed to further characterize the specific effects detected by the ANOVA. To follow up the main effect of task, a dependent samples t-test was performed comparing the difference in mean levels of task-initiation activity between visual (VU and VB) and auditory (AU and AB) tasks in V1, revealing significantly higher levels of task-initiation activity for visual tasks compared to auditory tasks ($t_{19} = 2.63, p < 0.001$). To follow up the main effect of ROI, Tukey's honestly significant difference (HSD) test was performed with an alpha of 0.05 to compare the overall difference in mean levels of task-initiation activity between all pairs of ROIs, but did not reveal significant differences between any pair of ROIs.

Results for trial-driven activity are shown in Figure 3. For trial-driven activity, a three-way repeated-measures ANOVA with factors of attended modality, number of stimuli, and ROI revealed significant two-way interactions of attended modality by ROI ($F_{6, 114} = 12.51, p < 0.001$), number of stimuli by ROI ($F_{6, 114} = 3.50, p = 0.003$), and attended modality by number of stimuli ($F_{1, 19} = 7.86, p = 0.011$), but no significant three-way interaction ($F_{6, 114} = 1.65, p = 0.14$).

To follow up the significant interaction of attended modality by ROI, one-way repeated measures ANOVAs with a factor of ROI were performed for the visual (VU and VB) and auditory (AU and AB) tasks, revealing a significant main effect of ROI on the magnitude of trial-driven activity for visual tasks ($F_{6,114} = 5.86, p < 0.001$), but not for auditory tasks ($F_{6,114} = 0.69, p = 0.65$). Post-hoc tests using Tukey's HSD with an alpha of 0.05 to compare the mean magnitude of trial-driven activity during visual tasks with the mean magnitude of trial-driven activity during auditory tasks for each ROI revealed significantly greater activity during visual tasks at ROI 1. Post-hoc tests using Tukey's HSD with an alpha of 0.05 to compare the mean magnitudes of trial-driven activity during the visual tasks (VU and VB) between all pairs of ROIs revealed that the mean magnitude of trial-driven activity was greater for ROI 1 than for ROIs 5,6, and 7.

To follow up the significant interaction of number of stimuli by ROI, one-way repeated measures ANOVAs with a factor of number of stimuli were performed for the unimodal (AU and VU) and bimodal (AB and VB) conditions of each task, revealing a significant main effect of ROI for bimodal tasks ($F_{6,114} = 2.60, p = 0.02$) but not for unimodal tasks ($F_{6,114} = 0.84, p = 0.54$). Post-hoc tests using Tukey's HSD with an alpha of 0.05 to compare the mean levels of trial-driven activity during bimodal tasks between all pairs of ROIs revealed that trial-driven activity for the bimodal tasks was significantly higher for ROI 1 than for ROI 7.

To follow up the significant interaction of attended modality by number of stimuli, post-hoc tests were performed using dependent-samples t-tests using Bonferroni correction to correct for multiple comparisons. Bonferroni correction was used here (rather than Tukey's HSD)

because there are only two pairwise comparisons. This revealed that trial-driven activity was higher for AB trials than for AU trials ($t_{19} = 3.67$, $p=0.003$), but did not differ significantly between VB trials and VU trials ($t_{19} = -0.05$, $p = 0.95$).

Results for task-maintenance activity are shown in Figure 4. For task-maintenance effects, a three-way repeated measures ANOVA with factors of attended modality, number of stimuli, and ROI revealed a significant main effect of attended modality ($F_{1, 19} = 8.65$, $p=0.008$), but no main effect of number of stimuli ($F_{1,19} = 0.13$, $p = 0.72$), no main effect of ROI ($F_{6, 114} = 1.67$, $p = 0.17$), no interaction of attended modality and number of stimuli ($F_{1,19} = 0.05$, $p=0.83$), no interaction of attended modality and ROI ($F_{6,114} = 1.09$, $p=0.37$), no interaction of number of stimuli and ROI ($F_{6,114} = 0.69$, $p= 0.65$), and no three-way interaction ($F_{6,114} = 0.53$, $p = 0.78$). A main effect contrast of attended modality was performed using a dependent samples t-test to analyze differences in task-maintenance activity between visual (VU and VB) vs. auditory (AU and AB) tasks, and revealed (as can be seen in Figure 4) that mean levels of task-maintenance activity were significantly higher during the visual tasks than during the auditory tasks ($t_{19} = 2.94$, $p = 0.008$).

3.3 Event-related design results

Results for cue-driven activity are shown in Figure 5. For cue-driven activity, a three-way repeated measures ANOVA revealed a significant two-way interaction of attended modality by ROI ($F_{6, 114} = 5.01$, $p<0.001$) and a significant two-way interaction of number of stimuli by ROI ($F_{6, 114} = 2.51$, $p=0.026$), but no significant interaction of attended modality by number of stimuli ($F_{1,19} = 0.11$, $p = 0.74$), and no three-way interaction ($F_{6, 114} = 1.08$, $p = 0.38$).

To follow up the significant interaction of attended modality by ROI, one-way repeated measures ANOVAs with a factor of ROI were performed for the visual (VU and VB) and auditory (AU and AB) tasks. These revealed a significant main effect of ROI on the magnitude of cue-driven activity for visual tasks ($F_{6,114} = 3.12$, $p=0.007$), but not for auditory tasks ($F_{6,114} = 1.04$, $p =0.40$). Post-hoc tests using Tukey's HSD with an alpha of 0.05 to compare mean levels of cue-driven activity between the visual and auditory tasks at each ROI revealed greater cue-driven activity during the visual tasks (VU and VB) than during the auditory tasks (AU and AB) for ROIs 1, 2, 3, and 4. Post-hoc tests using Tukey's HSD with an alpha of 0.05 to compare mean levels of activity during the visual tasks (VU and VB) between all pairs of ROIs revealed greater cue-driven activity for the visual tasks for ROI 1 than for ROIs 5, 6, and 7.

To follow up the significant interaction of number of stimuli by ROI, one-way repeated measures ANOVAs with a factor of ROI were performed for the unimodal (VU and AU) and bimodal (VB and AB) tasks, but did not reveal significant effects of ROI for the auditory ($F_{6,114} = 0.69$ $p = 0.66$) or visual ($F_{6,114} = 1.30$, $p = 0.26$) tasks.

Results for trial-driven activity in the event-related design are shown in Figure 5. Recall that these trials differ slightly from the trials in the mixed blocked event-related design, because a cue always precedes the beginning of the trial by 2 seconds (see Figure 1), and thus these effects also include responses to that cue. For trial-driven activity, a three-way repeated

measures ANOVA with factors of attended modality, number of stimuli, and ROI revealed a significant three-way interaction of attentional modality, number of stimuli, and ROI ($F_{6,114}=9.26$, $p<0.001$).

To follow up this significant 3-way interaction, a two-way repeated measures ANOVA with factors of attended modality and ROI was performed for the unimodal condition (VU vs AU) of both tasks. This test revealed a significant interaction of ROI by attended modality ($F_{6,114} = 25.149$, $p < 0.001$). Post-hoc tests using Tukey's HSD were performed to investigate differences in trial-driven activity between all pairs of ROIs for each condition with an alpha level of 0.05. For the AU task, ROI 7 showed greater trial-driven activity than ROIs 1, 2, and 3; ROI 6 showed greater trial-driven activity than ROIs 1, 2, 3, and 4; ROI 5 showed greater trial-driven activity than ROIs 1, 2, and 3; ROI 4 showed greater trial-driven activity than ROI 2. For the VU task, ROI 1 showed greater trial-driven activity than ROIs 4, 5, 6, and 7; ROI 2 showed greater trial-driven activity than ROIs 6 and 7; ROI 3 showed greater trial-driven activity than ROIs 6 and 7; ROI 4 showed greater trial-driven activity than ROI 7.

In order to investigate differences in the mean levels of trial-driven activity between the VU and AU task at each ROI, we performed post-hoc tests using Tukey's HSD with an alpha of 0.05. These revealed that trial-driven activity was higher for the VU task at ROIs 1, 2, and 3.

A follow-up two-way repeated measures ANOVA for the bimodal conditions of each task revealed a significant interaction of attended modality and ROI ($F_{6,114}= 7.27$, $p < 0.001$). Identical post-hoc tests were performed comparing the VB and AB tasks, and revealed higher levels of activity during the AB task for ROIs 4,5, and 6. Identical tests investigating differences between the VB and VU conditions at each ROI did not reveal any significant differences. Identical tests investigating differences between the AB and AU conditions revealed greater trial-driven activity during the AB task for ROIs 1, 2, 3, 4, 5, and 6. For the AB task, no ROIs significantly differed in trial-driven activity magnitudes. For the VB task, ROI 1 showed greater trial-driven activity than ROIs 4, 5, 6, and 7; ROI 3 showed greater trial-driven activity than ROIs 6 and 7.

3.4. Differences in activity between good and poor performers

In order to investigate whether task set or eccentricity-related differences in V1 activity were related to performance on the tasks, we divided the participants into two groups based on their performance on the bimodal tasks and performed further analyses using these groups. "Good" performers were defined as those 9 participants who performed above 75 percent correct on both the AB and VB tasks. The other 11 participants (all of whom performed above chance) were defined as "poor" performers.

Three-way mixed measures ANOVAs with within-subjects factors of attentional modality and ROI, and a between-subjects factor of group were performed for each time-course of activity for the bimodal tasks only. The bimodal condition of the tasks was chosen because they used identical stimuli, differing only in the attended modality. This is important because task-driven modulations of activity that occur during the bimodal tasks cannot be attributed to differences in stimuli, and therefore are likely attributable to purely task-related

processes. Because the purpose of these analyses was to identify differences in activity that relate to performance, only between-group main effects and interactions are reported.

A significant attended modality by group interaction ($F_{1, 18} = 7.79, p = 0.01$) was found for trial-driven activations from the mixed design (Figure 7A), but we did not find significant interactions of ROI by group ($F_{6,108} = 0.95, p=0.43$) or interactions of ROI, by group, by attended modality ($F_{6,108} = 1.59, p=0.22$). The attended modality by group interaction was followed up using Tukey's HSD with an alpha of 0.05, which revealed that good performers showed higher mean levels of trial-driven activity during AB than VB trials. Poor performers did not show significant differences in trial-driven activity between the two conditions.

A significant attended modality by group interaction ($F_{1, 18} = 8.55, p = 0.009$) was also found for task-maintenance activity (Figure 7B), but we did not find significant interactions of ROI and group ($F_{6,108} = 0.44, p=0.72$) or interactions of ROI, group, and attended modality ($F_{6,108} = 0.96, p=0.42$). The attended modality by group interaction was followed up using Tukey's HSD with an alpha of 0.05, which revealed that good performers had higher levels of task-maintenance activity during VB blocks than during AB blocks. Poor performers did not show significant differences in task-maintenance activity between the two conditions.

Analyses of task-initiation activity did not reveal significant interactions of attended modality by group ($F_{1, 18} = 0.27, p = 0.77$), ROI by group ($F_{6,108} = 0.0001, p=0.992$) or ROI by group by attended modality ($F_{6,108} = 0.83, p=0.54$). Analyses of cue-driven activity did not reveal significant interactions of attended modality by group ($F_{1, 18} = 0.37, p = 0.54$), ROI by group ($F_{6,108} = 0.95, p=0.42$) or ROI by group by attended modality ($F_{6,108} = 0.89, p=0.36$). Analyses of trial-driven activity from the event-related runs did not reveal significant interactions of attended modality by group ($F_{1, 18} = 0.31, p = 0.86$), ROI by group ($F_{6,108} = 0.34, p=0.80$) or ROI by group by attended modality ($F_{6,108} = 1.12, p=0.30$)

4. Discussion

Early visual cortical areas are influenced by top-down control (e.g., Somers et al., 1999; Kastner et al., 1999; Szczepanski et al., 2010; McMains et al., 2007; McMains and Kastner, 2011; Hopfinger et al., 2000; Zanto et al., 2010; Pooremaeili et al., 2014). Our data add to this literature, showing that (at least) two separate processes are involved in the control of vision, influencing neural activity at the level of V1. The first operates on a trial-by-trial temporal scale and modulates neural activity in V1 in a retinotopically specific manner. The second operates on a task-length temporal scale and does not appear to show retinotopic specificity.

4.1. Different retinotopic effects of trial-related and task set-related control in V1

The use of a mixed blocked/event-related design allows the measurement of changes in BOLD activity that correspond to distinct aspects of task performance, such as the initiation and maintenance of a task set or the response to individual trials (Visscher et al., 2003; Petersen and Dubis, 2011). Results from previous studies employing mixed designs have

shown that trial-driven and task-maintenance time-courses of BOLD activity can be differently modulated by task demands, and may even show opposite effects within the same brain region (Burgund et al., 2005). These observations have been interpreted as evidence that task-driven changes in task-maintenance and trial-driven signals likely reflect distinct components underlying complex cognitive processes such as recognition memory (Donaldson et al., 2001) and cognitive control (Braver et al., 2003).

Recent work by our lab has found that even in early visual cortex, task-driven changes in task-maintenance and trial-driven signals likely reflect distinct processes: one related to maintaining a state of attentional readiness and another related to processing stimuli according to task demands (Elkhetali et al., 2015). Specifically, we previously found that trial-driven activity in V1 depended on presented stimuli, while activity related to the initiation and maintenance of task sets depended instead on the modality of the task regardless of presence or absence of intermodal distractors. This finding led us to the hypothesis that the two control processes may act on V1 at different spatial scales. The current analyses characterize this spatial scale, showing how these trial-driven and task-maintenance signals depend on what portion of the visual field a given sector of V1 represents.

Even when stimuli are kept identical (in the AB and VB conditions), both trial-driven activity and task-maintenance activity in V1 depend on task. However, there are important differences in this task dependence. For example, Figure 4B and associated statistics show that task maintenance activity is stronger in the VB than the AB condition, and this effect does not differ significantly across V1. Conversely, the task-dependence of trial-driven activity depends strongly on which portion of V1 is examined. Figure 3 and associated statistics show that trial-driven activity for the AB condition is strongest in peripheral V1, but not in central V1. These observations are in accordance with our hypothesis that previously reported differences in trial-driven and task-maintenance signals might stem from differences in retinotopic specificity. This difference in direction of the modulations by task give further evidence that they represent different control processes acting on V1. In addition, they show that the processes acting with a trial-by-trial temporal profile can influence V1 in a retinotopically-specific manner (likely by suppressing activity in peripheral V1 in this case), while the processes acting on a longer timescale appear to influence V1 as a whole. Task-maintenance signals also show nearly identical magnitudes for both unimodal and bimodal conditions of each task, (Figure 4). However, that was not true for trial-driven signals (Figure 3/ Figure 7A). This suggests that task maintenance signals are not strongly influenced by the presence or absence of task-irrelevant visual stimuli during attention to other modalities.

Our data show that trial-driven responses in V1 are strongly different across eccentricity when vision is attended, but not when vision is not attended: during the visual bimodal condition, there was a strong decrease in trial-driven responses from central to peripheral eccentricity sectors. Conversely, no such effect was present during the auditory bimodal condition. This implies that the differences in trial-driven activity as a function of eccentricity sector are enhanced during attention. It is worth noting that the lack of retinotopic specificity observed for task-maintenance activity suggests that it likely reflects

different processes than those responsible for sustained activations in V1 that occur during trial-specific delay periods during cued attention (Silver et al., 2007). These previously reported delay period effects likely reflect prolonged cue-driven changes in activation that are trial-specific, similar to the cue-driven signals observed here, which, as shown in Figure 5, do depend on retinotopy.

Similarly, while both cue-driven preparatory activity and task-initiation activity were stronger for visual than for auditory tasks (Figure 2/5), cue-driven activity showed some retinotopic specificity during visual tasks compared to auditory tasks. Cue-driven responses were significantly greater for the most central V1 sector (Figure 5 ROI 1) than for the most peripheral V1 sector (Figure 5 ROI 7) during visual tasks, and were stronger during visual tasks than during auditory tasks in central and paracentral sectors (Figure 5 ROIs 1/2/3/4). In contrast, task-initiation activity did not show significant differences in retinotopic specificity between visual and auditory tasks, and overall levels of task-initiation activity were not found to differ between any eccentricity sectors (Figure 2).

To interpret cue-driven and task-initiation activity, it is worth noting that while both signals are time-locked to the presentation of a cue, this cue appears following a long rest period for task-initiation activity, and during task performance for cue-driven activity. While, compared to baseline, cue-driven and task-initiation activity during auditory tasks appear to show opposite effects, task-initiation activity likely reflects shifts in baseline that occur at the boundary of rest periods and need to be sustained over the course of task performance (Elkhetali et al., 2015). Cue-driven effects, in contrast, likely reflect transient preparatory activations that occur during periods of prolonged alertness (e.g. Kastner et al., 1999; Sylvester et al., 2008; Elkhetali et al., 2015). Thus, the observed cue-driven effects may reflect transient preparatory processes that are only maintained until the end of the trial, whereas task-initiation effects may reflect the configuration of a task set that must be maintained for a prolonged period of time.

Importantly, although participants who performed well on the tasks showed strong task-dependent modulations of both task-maintenance and trial-driven activity, we found that these modulations were in opposite directions (Figure 7). This is consistent with the proposal that sustained and transient responses reflect distinct processes that interact to enable task-appropriate stimulus processing (Burgund et al., 2005). This further suggests that task-maintenance and trial-driven signals in V1 may reflect the actions of distinct control processes that operate on different temporal and spatial resolutions.

4.2 Interpretation of these data in the context of top-down control

Where might these control processes originate? One possibility is that task-related changes in trial-driven and task-maintenance signals reflect the actions of distinct but complementary brain networks that operate within different temporal domains and influence stimulus processing and response generation via separable processes (Dosenbach et al., 2008, 2007). One of these networks, the fronto-parietal network, consists of areas classically associated with the top-down control of visual attention and includes the intraparietal sulcus (IPS), inferior parietal lobule (IPL), and frontal eye fields (FEF) (Szczepanski et al., 2010; Liu et al., 2014; Hopfinger et al., 2000; Bressler et al., 2008; Giesbrecht et al., 2003). This network

is primarily thought to be responsible for the goal-directed adjustment of sensory processing during task performance, and is involved in implementing trial-by-trial modulations of neural activity following task-relevant cues, expected stimuli, and errors (Dosenbach et al., 2006). This interpretation is consistent with reports that activity in the fronto-parietal network predicts the level of activity in early visual areas prior to the presentation of expected stimulus (Bressler et al., 2008; Vossel et al., 2012), as well as with evidence that the suppression of activity in parts of early visual cortex that correspond to unattended spatial locations likely results from top-down modulations by the FEF and IPS (Sylvester et al., 2008). The observations that trial-driven effects in V1 depend on both the task and the stimulus in a way that is specific to retinotopy are consistent with the interpretation that trial-to-trial control processes we observe in visual cortex may be produced through the fronto-parietal network, though future work is needed to show this directly. Regardless of the origin of the trial-to-trial control processes we observe, our data imply that this control process does not treat V1 as a unit, but rather exerts top-down influence on precise sectors within V1.

Another network, the cingulo-opercular network, consists of areas that show sustained increases in neural activity across a diverse range of task paradigms and includes the anterior insula and dorsal anterior cingulate cortex (dACC) (Dosenbach et al., 2006). Thus, this network is considered to be a likely source of signals related to the initiation and maintenance of task sets – configurations of cognitive processes that allow for stimuli to be processed and behavioral responses to be generated according to task-specific rules and associations (Sakai, 2008b). As such, the cingulo-opercular network likely operates on longer temporal scales than the fronto-parietal network (Dosenbach et al., 2006). This is consistent with recent reports suggesting that the anterior insula selectively interacts with early visual areas during visuo-spatial tasks (Ebisch et al., 2013). The observations that task-maintenance and task-initiation effects in V1 depend on the task rather than the stimulus and show more generalized retinotopic effects are consistent with the interpretation that task-maintenance signals in visual cortex may be produced through the cingulo-opercular network, though future work is needed to show this directly. Regardless of the origin of the slower control processes we observe, our data imply that this control process treats V1 as a unit, rather than exerting top-down influence on segments of the area.

Thus, one possible interpretation of our findings is that task-related changes in the trial-driven and cue-driven responses of early visual areas reflect top-down modifications of visual processing by regions that instantiate control with both high temporal (e.g. trial-specific) and spatial (e.g. retinotopically specific) resolution, most probably fronto-parietal regions. In contrast, changes in the neural activity of early visual areas during periods of task initiation and task maintenance may reflect a lower temporal (e.g. block duration) and spatial resolution (e.g. retinotopically non-specific) form of control that, presumably, may reflect the action of the cingulo-opercular network. Thus, future experiments should investigate how the functional connectivity of different eccentricity sectors in V1 is modulated according to task demands. Based on these findings, it would be expected that changes in connectivity with fronto-parietal control regions would be retinotopically specific where changes in connectivity with cingulo-opercular control regions would not show such specificity. Nonetheless, while future work is needed to clarify precise mechanisms, the

results from this experiment provide strong evidence that task-appropriate visual processing is enabled by distinct but complimentary control processes that influence neural activity even at the level of V1.

4.3. Trial-driven activity in peripheral V1 during auditory attention

Auditory attention has been found to modulate activity in human V1 (Ciaramitaro et al., 2007), and recent experimental evidence has shown that auditory attention may selectively enhance activity in regions of early visual cortex that correspond to peripheral visual space (Cate et al., 2009). While not a primary goal of our study, our data are consistent with trial-by-trial enhancement of peripheral V1 with attention to audition. For example, in Figure 6B, when attention was oriented toward the auditory stimulus and the visual stimulus was ignored, peripheral V1 was more active. A similar effect can be seen for unimodal trials: when only auditory stimuli were presented, activity was highest in peripheral than central V1 sectors (Figure 6A). While this effect was only found to be significant for trials of the event-related design (which included cues on each trial), a similar trend was observed for trials of the mixed design (Figure 3). The role of predominantly peripheral trial-driven activations during auditory attention is not clear. One possibility is that peripherally responsive regions of V1 are involved in orienting attention to the location of auditory stimuli (Cate et al., 2009) – an explanation that is supported by studies of cortical anatomy in non-human primates that have found projections from primary auditory cortex that terminate in peripheral sectors of V1 (Clavagnier et al., 2004; Falchier et al., 2002; Hall and Lomber, 2008).

Interestingly, functional connectivity between peripheral V1 and auditory cortex has been reported in humans (Eckert et al., 2008), and recent evidence indicates that task-irrelevant sounds activate extrastriate visual areas and enhance performance on visual discrimination tasks (Feng et al., 2014). However, our behavioral results do not indicate that the addition of task-irrelevant sounds had any substantial effect on accuracy or response times during the VB task. While our data replicates the findings of previous studies in regard to the effects of auditory attention on peripheral visual areas (Cate et al., 2009), further research is needed to fully characterize the mechanisms underlying auditory activations in peripheral visual cortex and to clarify their functional relevance.

5. Conclusion

A previous study by our lab found that in V1, task-maintenance activity is much more influenced by a participant's attention to vision than is trial-driven activity (Elkhetali et al., 2015). In the current study, we hypothesized that the observed effects might stem from differences in the retinotopic specificity of attentional modulations related to trial-by-trial processing vs. the initiation/maintenance of task sets. Using mixed block/event-related and event-related fMRI, we comprehensively examined BOLD activity in different eccentricity sectors of V1 during distinct temporal epochs of task performance for unimodal and intermodal visual and auditory attention tasks. We found that the retinotopic patterns of cue-driven and trial-driven activity varied according to the task being performed, and the retinotopic patterns of trial-driven activity were further influenced by the presence of

intermodal distractors. In contrast, the retinotopic patterns of task-maintenance activity showed task-dependent changes in overall magnitude, which didn't depend on retinotopy. Further, good performers showed strong task-driven changes in both trial-driven and task-maintenance activity between AB and VB tasks that were in opposite directions, while poor performers did not show strong task-driven modulations of either trial-driven or task-maintenance activity. The observed differences in the retinotopic specificity and direction of task-driven changes in trial-driven and task-maintenance activity provide strong evidence that top-down control of visual processing is accomplished via distinct processes that operate with different spatial and temporal resolutions, and may reflect the actions of distinct brain networks that are involved in the control of different aspects of visual processing.

Acknowledgements

We would like to thank the following people for excellent comments on this manuscript and help with data acquisition: Wesley Burge, Rodolphe Nenert and Sean Pool. We would also like to thank three anonymous reviewers for their excellent suggestions.

Funding

This work was supported by the UAB Center for Clinical And Translational Science (grant number UL1 TR000165); Vision Science Research Center P30 (grant number EY003039); the Civitan International Research Center; the McKnight Brain Research Foundation; and the Edward R. Roybal Center for Translational Research on Aging and Mobility (grant number NIA 2 P30 AG022838).

References

- Benson NC, Butt OH, Datta R, Radoeva PD, Brainard DH, Aguirre GK. The retinotopic organization of striate cortex is well predicted by surface topology. *Curr. Biol.* 2012; 22:2081–5. doi:10.1016/j.cub.2012.09.014. [PubMed: 23041195]
- Braver TS, Reynolds JR, Donaldson DI. Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron.* 2003; 39:713–26. [PubMed: 12925284]
- Bressler SL, Tang W, Sylvester CM, Shulman GL, Corbetta M. Top-Down Control of Human Visual Cortex by Frontal and Parietal Cortex in Anticipatory Visual Spatial Attention. *J. Neurosci.* 2008; 28:10056–10061. doi:10.1523/JNEUROSCI.1776-08.2008. [PubMed: 18829963]
- Burgund ED, Lugar HM, Miezin FM, Schlaggar BL, Petersen SE. The development of sustained and transient neural activity. *Neuroimage.* 2006; 29:812–821. doi:10.1016/j.neuroimage.2005.08.056. [PubMed: 16236529]
- Burgund E, Lugar HM, Schlaggar BL, Petersen SE. Task demands modulate sustained and transient neural activity during visual-matching tasks. *Neuroimage.* 2005; 25:511–9. doi:10.1016/j.neuroimage.2004.12.039. [PubMed: 15784430]
- Cardoso MMB, Sirotin YB, Lima B, Glushenkova E, Das A. The neuroimaging signal is a linear sum of neurally distinct stimulus- and task-related components. *Nat Neurosci.* 2012; 15:1298–1306. doi:10.1038/nn.3170. [PubMed: 22842146]
- Cate AD, Herron TJ, Yund EW, Stecker GC, Rinne T, Kang X, Petkov CI, Disbrow EA, Woods DL. Auditory Attention Activates Peripheral Visual Cortex. *PLoS One.* 2009;4. doi:10.1371/journal.pone.0004645.
- Chawla D, Rees G, Friston KJ. The physiological basis of attentional modulation in extrastriate visual areas. *Nat. Neurosci.* 1999; 2:671–676. doi:10.1038/10230. [PubMed: 10404202]
- Ciaramitaro VM, Buracas GT, Boynton GM. Spatial and cross-modal attention alter responses to unattended sensory information in early visual and auditory human cortex. *J. Neurophysiol.* 2007; 98:2399–413. doi:10.1152/jn.00580.2007. [PubMed: 17715196]

- Clavagnier S, Falchier A, Kennedy H. Long-distance feedback projections to area V1: implications for multisensory integration, spatial awareness, and visual consciousness. *Cogn. Affect. Behav. Neurosci.* 2004; 4:117–26. [PubMed: 15460918]
- Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage.* 1999; 9:179–94. doi:10.1006/nimg.1998.0395. [PubMed: 9931268]
- Donaldson DI, Petersen SE, Ollinger JM, Buckner RL. Dissociating state and item components of recognition memory using fMRI. *Neuroimage.* 2001; 13:129–142. doi:10.1006/nimg.2000.0664. [PubMed: 11133316]
- Dosenbach NUF, Fair DA, Cohen AL, Schlaggar BL, Petersen SE. A dual-networks architecture of top-down control. *Trends Cogn. Sci.* 2008; 12:99–105. doi:10.1016/j.tics.2008.01.001. [PubMed: 18262825]
- Dosenbach NUF, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RAT, Fox MD, Snyder AZ, Vincent JL, Raichle ME, et al. Distinct brain networks for adaptive and stable task control in humans. *Proc. Natl. Acad. Sci. USA.* 2007; 104:11073–11078. doi:10.1073/pnas.0704320104. [PubMed: 17576922]
- Dosenbach NUF, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HC, Burgund ED, Grimes AL, Schlaggar BL, Petersen SE. A core system for the implementation of task sets. *Neuron.* 2006; 50:799–812. doi:10.1016/j.neuron.2006.04.031. [PubMed: 16731517]
- Ebisch SJH, Mantini D, Romanelli R, Tommasi M, Perrucci MG, Romani GL, Colom R, Saggino A. Long-range functional interactions of anterior insula and medial frontal cortex are differently modulated by visuospatial and inductive reasoning tasks. *Neuroimage.* 2013; 78:426–38. doi: 10.1016/j.neuroimage.2013.04.058. [PubMed: 23624492]
- Eckert MA, Kamdar NV, Chang CE, Beckmann CF, Greicius MD, Menon V. A cross-modal system linking primary auditory and visual cortices: Evidence from intrinsic fMRI connectivity analysis. *Hum. Brain Mapp.* 2008; 29:848–857. doi:10.1002/hbm.20560. [PubMed: 18412133]
- Elkhetali AS, Vaden RJ, Pool SM, Visscher KM. Early visual cortex reflects initiation and maintenance of task set. *Neuroimage.* 2015 doi:10.1016/j.neuroimage.2014.11.061.
- Falchier A, Clavagnier S, Barone P, Kennedy H. Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci.* 2002; 22:5749–59. doi:20026562. [PubMed: 12097528]
- Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. cortex.* 1991; 1:1–47. [PubMed: 1822724]
- Feng W, Stormer VS, Martinez a. McDonald JJ, Hillyard S. a. Sounds Activate Visual Cortex and Improve Visual Discrimination. *J. Neurosci.* 2014; 34:9817–9824. doi:10.1523/JNEUROSCI.4869-13.2014. [PubMed: 25031419]
- Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Mapp.* 1994; 2:189–210. doi: 10.1002/hbm.460020402.
- Giesbrecht B, Woldorff M, Song a. Mangun G. Neural mechanisms of top-down control during spatial and feature attention. *Neuroimage.* 2003; 19:496–512. doi:10.1016/S1053-8119(03)00162-9. [PubMed: 12880783]
- Hall AJ, Lomber SG. Auditory cortex projections target the peripheral field representation of primary visual cortex. *Exp. Brain Res.* 2008; 190:413–30. doi:10.1007/s00221-008-1485-7. [PubMed: 18641978]
- Hopfinger JB, Buonocore MH, Mangun GR. The neural mechanisms of top-down attentional control. *Nat. Neurosci.* 2000; 3:284–91. doi:10.1038/72999. [PubMed: 10700262]
- Huk AC, Heeger DJ. Task-Related Modulation of Visual Cortex. *J Neurophysiol.* 2000; 83:3525–3536. [PubMed: 10848568]
- Kastner S, Pinsk M. a. De Weerd P, Desimone R, Ungerleider LG. Increased Activity in Human Visual Cortex during Directed Attention in the Absence of Visual Stimulation. *Neuron.* 1999; 22:751–761. doi:10.1016/S0896-6273(00)80734-5. [PubMed: 10230795]
- Lindquist, M. a.; Meng Loh, J.; Atlas, LY.; Wager, TD. Modeling the hemodynamic response function in fMRI: efficiency, bias and mis-modeling. *Neuroimage.* 2009; 45:S187–S198. doi:10.1016/j.neuroimage.2008.10.065. [PubMed: 19084070]

- Liu Y, Bengson J, Huang H, Mangun GR, Ding M. Top-down Modulation of Neural Activity in Anticipatory Visual Attention: Control Mechanisms Revealed by Simultaneous EEG-fMRI. *Cereb. Cortex*. 2014 doi:10.1093/cercor/bhu204.
- Loftus GR, Masson MEJ. Using confidence intervals in within-subject designs. *Psychon. Bull. Rev.* 1994; 1:476–490. doi:10.3758/BF03210951. [PubMed: 24203555]
- Mazaika PK, Whitfield S, Cooper JC. Detection and repair of transient artifacts in fMRI data. *Neuroimage*. 2005; 26:S36.
- McMains, S. a; Fehd, HM.; Emmanouil, T-A.; Kastner, S. Mechanisms of feature- and space-based attention: response modulation and baseline increases. *J. Neurophysiol.* 2007; 98:2110–21. doi: 10.1152/jn.00538.2007. [PubMed: 17671104]
- McMains S, Kastner S. Interactions of top-down and bottom-up mechanisms in human visual cortex. *J. Neurosci.* 2011; 31:587–597. doi:10.1523/JNEUROSCI.3766-10.2011. [PubMed: 21228167]
- Munneke J, Heslenfeld DJ, Theeuwes J. Directing attention to a location in space results in retinotopic activation in primary visual cortex. *Brain Res.* 2008; 1222:184–91. doi:10.1016/j.brainres.2008.05.039. [PubMed: 18589405]
- Murray SO. The effects of spatial attention in early human visual cortex are stimulus independent. *J. Vis.* 2008; 8:1–11. doi:10.1167/8.10.2.
- Ollinger JM, Shulman GL, Corbetta M. Separating processes within a trial in event-related functional MRI. *Neuroimage*. 2001; 13:210–7. doi:10.1006/nimg.2000.0710. [PubMed: 11133323]
- Petersen SE, Dubis JW. The mixed block/event-related design. *Neuroimage*. 2011; 62:1177–1184. doi: 10.1016/j.neuroimage.2011.09.084. [PubMed: 22008373]
- Pooresmaeili A, Poort J, Roelfsema PR. Simultaneous selection by object-based attention in visual and frontal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 2014; 111:6467–72. doi:10.1073/pnas.1316181111. [PubMed: 24711379]
- Rosen BR, Buckner RL, Dale AM. Event-related functional MRI: past, present, and future. *Proc. Natl. Acad. Sci. USA.* 1998; 95:773–780. [PubMed: 9448240]
- Runeson E, Boynton GM, Murray SO. Effects of task and attentional selection on responses in human visual cortex. *J. Neurophysiol.* 2013; 109:2606–17. doi:10.1152/jn.00318.2012. [PubMed: 23427301]
- Sakai K. Task set and prefrontal cortex. *Annu. Rev. Neurosci.* 2008; 31:219–45. doi:10.1146/annurev.neuro.31.060407.125642. [PubMed: 18558854]
- Silver, M. a; Ress, D.; Heeger, DJ. Neural correlates of sustained spatial attention in human early visual cortex. *J. Neurophysiol.* 2007; 97:229–37. doi:10.1152/jn.00677.2006. [PubMed: 16971677]
- Somers DC, Dale A M. Seiffert A E. Tootell RB. Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 1999; 96:1663–8. [PubMed: 9990081]
- Sylvester CM, Jack AI, Corbetta M, Shulman GL. Anticipatory suppression of nonattended locations in visual cortex marks target location and predicts perception. *J. Neurosci.* 2008; 28:6549–6556. doi:10.1523/JNEUROSCI.0275-08.2008. [PubMed: 18579728]
- Sylvester CM, Shulman GL, Jack AI, Corbetta M. Anticipatory and stimulus-evoked blood oxygenation level-dependent modulations related to spatial attention reflect a common additive signal. *J. Neurosci.* 2009; 29:10671–10682. doi:10.1523/JNEUROSCI.1141-09.2009. [PubMed: 19710319]
- Szczepanski SM, Konen CS, Kastner S. Mechanisms of spatial attention control in frontal and parietal cortex. *J. Neurosci.* 2010; 30:148–60. doi:10.1523/JNEUROSCI.3862-09.2010. [PubMed: 20053897]
- Velanova K, Jacoby LL, Wheeler ME, McAvoy MP, Petersen SE, Buckner RL. Functional-anatomic correlates of sustained and transient processing components engaged during controlled retrieval. *J. Neurosci.* 2003; 23:8460–70. [PubMed: 13679414]
- Visscher KKM, Miezin FFM, Kelly JJE, Buckner RL, Donaldson DI, McAvoy MP, Bhalodia VM, Petersen SE. Mixed blocked/event-related designs separate transient and sustained activity in fMRI. *Neuroimage*. 2003; 19:1694–708. [PubMed: 12948724]

- Visscher KM, Kaplan E, Kahana MJ, Sekuler R. Auditory short-term memory behaves like visual short-term memory. *PLoS Biol.* 2007; 5:0662–0672. doi:10.1371/journal.pbio.0050056.
- Vossel S, Weidner R, Driver J, Friston KJ, Fink GR. Deconstructing the architecture of dorsal and ventral attention systems with dynamic causal modeling. *J. Neurosci.* 2012; 32:10637–48. doi: 10.1523/JNEUROSCI.0414-12.2012. [PubMed: 22855813]
- Warnking J, Dojat M, Guérin-Dugué A, Delon-Martin C, Olympeff S, Richard N, Chéhikian A, Segebarth C. fMRI Retinotopic Mapping—Step by Step. *Neuroimage.* 2002; 17:1665–1683. doi: 10.1006/nimg.2002.1304. [PubMed: 12498741]
- Watanabe T, Harner a M, Miyauchi S, Sasaki Y, Nielsen M, Palomo D, Mukai I. Task-dependent influences of attention on the activation of human primary visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 1998; 95:11489–92. [PubMed: 9736764]
- Watson AB, Pelli DG. QUEST: a Bayesian adaptive psychometric method. *Percept. Psychophys.* 1983; 33:113–20. [PubMed: 6844102]
- Wenger KK, Visscher KM, Miezin FM, Petersen SE, Schlaggar BL. Comparison of sustained and transient activity in children and adults using a mixed blocked/event-related fMRI design. *Neuroimage.* 2004; 22:975–985. doi:10.1016/j.neuroimage.2004.02.028. [PubMed: 15193629]
- Zanto TP, Rubens MT, Bollinger J, Gazzaley A. Top-down modulation of visual feature processing: the role of the inferior frontal junction. *Neuroimage.* 2010; 53:736–45. doi:10.1016/j.neuroimage.2010.06.012. [PubMed: 20600999]

Highlights

1. There are at least two control processes observable in primary visual cortex (V1)
2. One, with a trial-by-trial temporal profile also depends on retinal eccentricity
3. A second type of control process is slower and influences V1 as a whole
4. These results can be informative to understanding the cortical mechanisms of control of V1

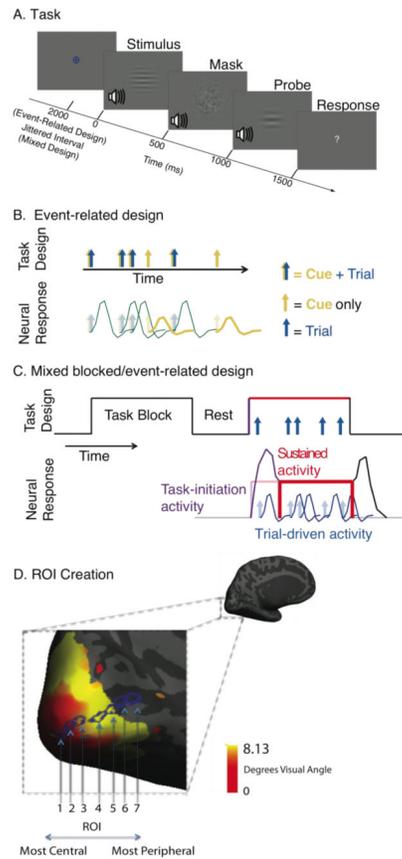


Figure 1. Task Design

A. The diagram shows one trial of the “visual bimodal” task. The fixation symbol indicated which task the participant should perform. For the event-related design, stimuli were presented 2000ms after the fixation symbol. For the mixed design, stimuli were presented at jittered intervals after the fixation symbol. The stimulus, mask, and probe were presented for 500 ms each. Participants were given 2000 ms to respond, indicating whether the relevant stimulus and probe were the same or different. Here, because the relevant stimulus modality was visual, the correct response would be ‘different.’ **B.** A diagram of the event-related design is shown. Blue and gold arrow pairs indicate cue and stimulus presentation, respectively; some cues were not followed by a stimulus. These were used to estimate cue-driven activity. In this experiment, ‘cued runs’ were presented using an event-related design. **C.** A diagram of the mixed blocked/event-related design is shown. Arrows represent individual trials presented during task blocks. The design is used to estimate trial-driven activity associated with individual trials (blue), sustained changes in baseline activity associated with task maintenance (red) and preparatory activity associated with the start of the task block (purple). In this experiment, ‘blocked runs’ were presented with a mixed blocked/event-related design. **D.** V1 regions of interest are shown for a single participant. Each participant’s visual cortex was retinotopically mapped and resulting eccentricity maps were used to create 7 regions of interest (ROIs) for each participant at roughly equal intervals along the calcarine sulcus. The first (most central) ROI corresponded to ~1 degree visual angle, the second ROI corresponds to ~2.6 degrees visual angle, the third

ROI corresponds to ~4.2 degrees visual angle, the fourth ROI corresponds to ~5.8 degrees visual angle, the fifth ROI corresponds to ~7.4 degrees visual angle, and the last two ROIs corresponding to > 7.5 degrees visual angle.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

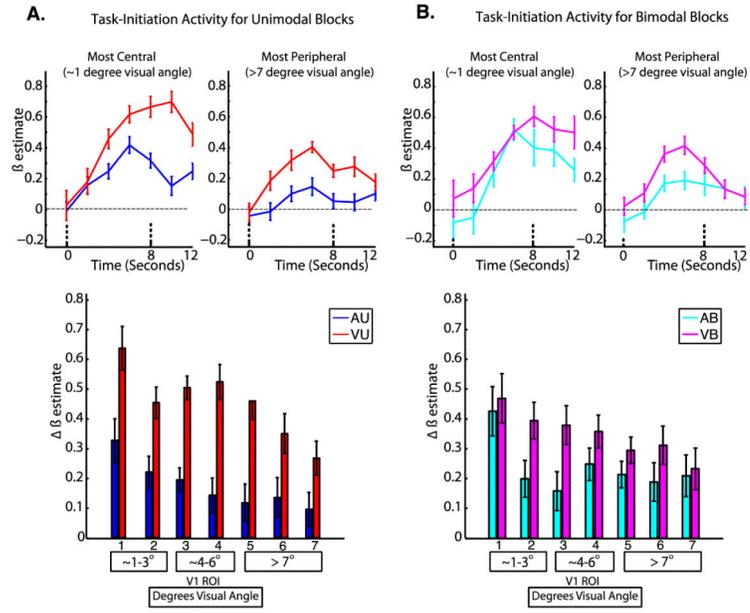


Figure 2. Task-initiation Activity

Group averaged hemodynamic responses are shown for the most central and most peripheral ROIs for unimodal (A, top) and bimodal (B, top) tasks; the dashed lines indicate the timepoints used to compute the magnitude of task initiation effects. The mean magnitude of peak task initiation activation is shown at each ROI for unimodal (A, bottom) and bimodal (B, bottom) tasks; error bars represent within-subjects standard error of the mean (Loftus and Masson, 1994). For the bottom panels, the numbers below each bar indicate the V1 ROI, and the boxes below contain the corresponding retinal eccentricities in degrees visual angle.

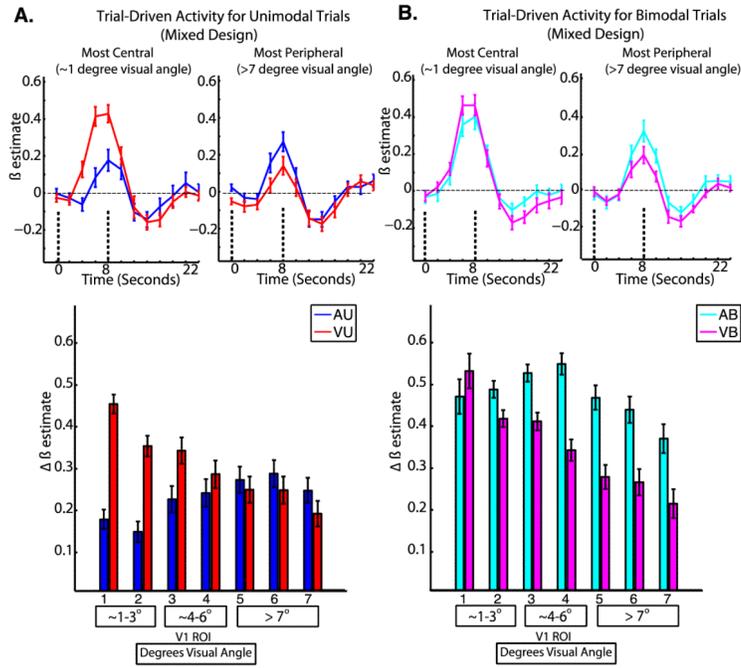


Figure 3. Trial-Driven Activity (Mixed Blocked/Event-Related Design)

Group averaged hemodynamic responses are shown for the most central and most peripheral ROIs for unimodal (A, top) and bimodal (B, top) tasks; the dashed lines indicate the timepoints used to compute the magnitude of trial-driven effects. The mean magnitude of trial-driven activity is shown at each ROI for unimodal (A, bottom) and bimodal (B, bottom) tasks; error bars represent within-subjects standard error of the mean (Loftus and Mason, 1994). For the bottom panels, the numbers below each bar indicate the V1 ROI, and the boxes below contain the corresponding retinal eccentricities in degrees visual angle.

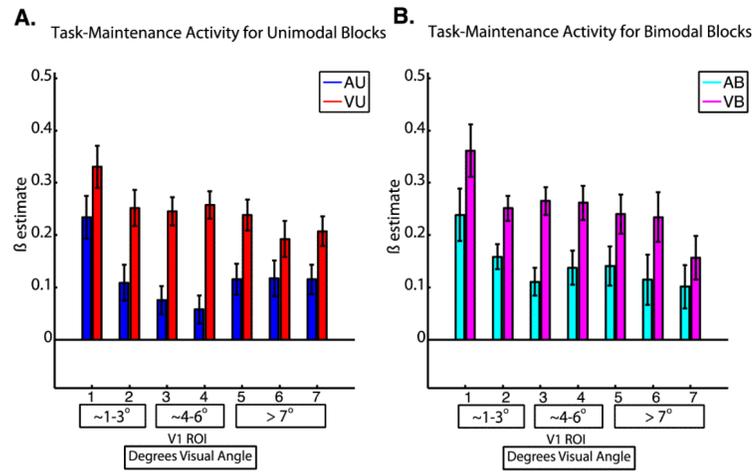


Figure 4. Task-maintenance Activity

The mean magnitude of task initiation activity is shown at each ROI for unimodal (A) and bimodal (B) tasks. Error bars represent within-subjects standard error of the mean (Loftus and Mason, 1994). The numbers below each bar indicate the V1 ROI, and the boxes below contain the corresponding retinal eccentricities in degrees visual angle.

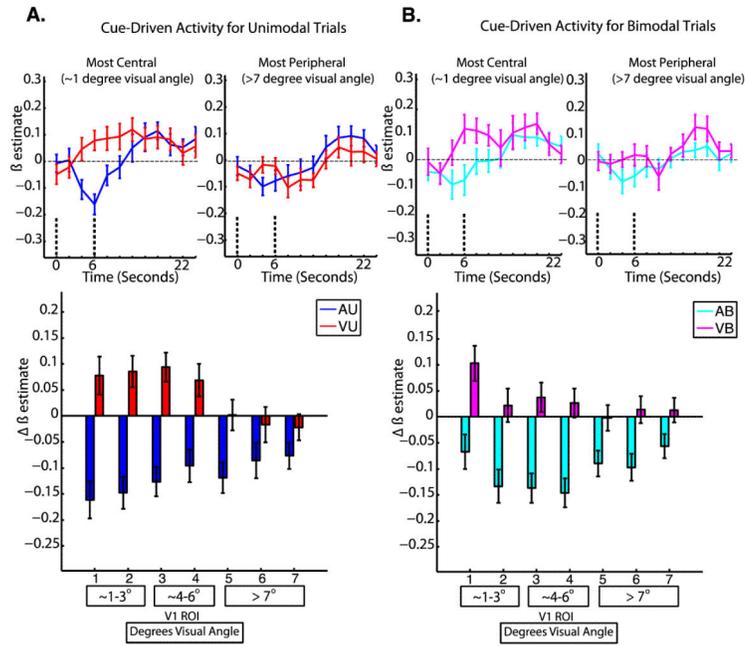


Figure 5. Cue-driven Activity (Event-Related Design)

Group averaged hemodynamic responses are shown for the most central and most peripheral ROIs for unimodal (A, top) and bimodal (B, top) tasks; the dashed lines indicate the timepoints used to compute the magnitude of cue-driven effects. The mean magnitude of cue-driven activity is shown at each ROI for unimodal (A, bottom) and bimodal (B, bottom) tasks. Error bars represent within-subjects standard error of the mean (Loftus and Mason, 1994). The numbers below each bar indicate the V1 ROI, and the boxes below contain the corresponding retinal eccentricities in degrees visual angle.

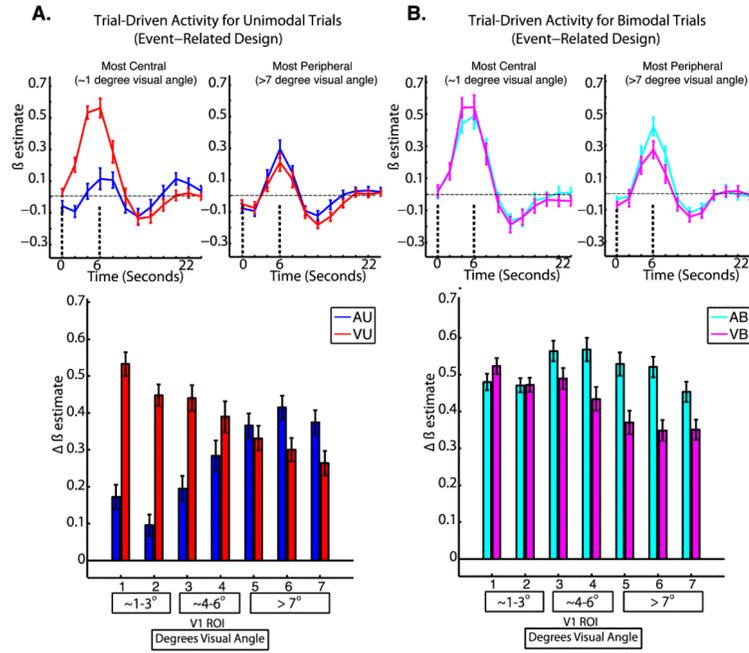


Figure 6. Trial-Driven Activity (Event-Related Design)

Group averaged hemodynamic responses are shown for the most central and most peripheral ROIs for unimodal (A, top) and bimodal (B, top) tasks; the dashed lines indicate the timepoints used to compute the magnitude of trial-driven effects. The mean magnitude of trial-driven activity is shown at each ROI for unimodal (A, bottom) and bimodal (B, bottom) tasks. Error bars represent within-subjects standard error of the mean. The numbers below each bar indicate the V1 ROI, and the boxes below contain the corresponding retinal eccentricities in degrees visual angle.

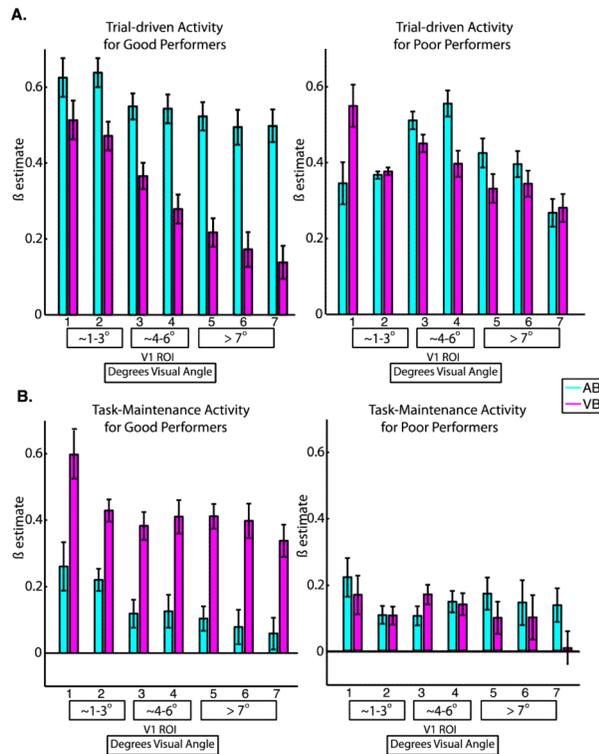


Figure 7. Performance-Related Group Differences in Trial-Driven and Task-Maintenance Activity

Significant group-by-task interactions were observed for trial-driven activity from the mixed blocked/event-related design and for task-maintenance activity. Mean magnitudes of trial-driven activity (top) and task-maintenance activity (bottom) are shown for good performers (left) and poor performers (right). **A.** The mean magnitude of trial-driven activity was significantly greater for auditory than for visual tasks for individuals who performed well on the task (left). Note that while task-related modulations of trial-driven activity show a strong retinotopic pattern for good performers (left), no such pattern is observed for poor performers (right). **B.** The mean magnitude of task-maintenance activity was significantly greater during the visual task than during the auditory task for individuals who performed well on the task (left). Individuals who performed poorly on the task did not show task-specific modulations of task-maintenance activity (right). Error bars represent within-subjects standard error of the mean. The numbers below each bar indicate the V1 ROI, and the boxes below contain the corresponding retinal eccentricities in degrees visual angle.