Selection of events in time enhances activity throughout early visual cortex

Khena M. Swallow, Tal Makovski, and Yuhong V. Jiang
Department of Psychology and Center for Cognitive Sciences, University of Minnesota, Minneapolis, Minnesota

Swallow KM, Makovski T, Jiang YV. Selection of events in time enhances activity throughout early visual cortex. *J Neurophysiol* 108: 3239–3252, 2012. First published September 19, 2012; doi:10.1152/jn.00472.2012.—Temporal selection poses unique challenges to the perceptual system. Selection is needed to protect goal-relevant stimuli from interference from new sensory input. In addition, contextual information that occurs at the same time as goal-relevant stimuli may be critical for learning. Using fMRI, we characterized how visual cortical regions respond to the temporal selection of auditory and visual stimuli. Critically, we focused on brain regions that are not involved in processing the target itself. Participants pressed a button when they heard a prespecified target tone and did not respond to other tones. Although more attention was directed to auditory input when the target tone was selected, activity in primary visual cortex increased more after target tones than after distractor tones. In contrast to spatial attention, this effect was larger in V1 than in V2 and V3. It was present in regions not typically involved in representing the target stimulus. Additional experiments demonstrated that these effects were not due to multimodal processing, rare targets, or motor responses to the targets. Thus temporal selection of behaviorally relevant stimuli enhances, rather than reduces, activity in perceptual regions involved in processing other information.

attention; primary visual cortex

ALTHOUGH THE NATURAL ENVIRONMENT is usually stable over time, changes in sensory input occur with the appearance of new objects and navigation through the environment. Some of these changes may be more relevant to a person's goals than others. Adaptive perception requires attentional selection over time (Chun and Potter 1995; Neisser 1976; Pashler 1994). Previous studies have characterized temporal selection as a late process that facilitates encoding into working memory (Bowman and Wyble 2007; Chun and Potter 1995; Olivers and Meeter 2008). However, its impact on early visual cortical activity is poorly understood. In this study, we used functional magnetic resonance imaging (fMRI) to examine how the temporal selection of brief auditory and visual stimuli affects activity in early visual cortical regions that are not involved in coding them.

One way that temporal selection may affect early visual activity is by recruiting spatial selection mechanisms for a brief period of time. Spatial selection prioritizes the processing of selected locations. It ensures that objects in those locations successfully compete for neural representation within a neuron's receptive field (Desimone and Duncan 1995; Reynolds and Chelazzi 2004). The resulting bias manifests as increased activity in regions representing the attended location and decreased activity in regions representing nearby locations (Desimone and Duncan 1995; Reynolds and Heeger 2009). This

modulation is greater in later visual areas that have larger receptive fields (Kastner et al. 1998). Selecting information in time, however, poses a distinct set of computational challenges. Unlike simultaneously presented stimuli, sequentially presented stimuli do not strongly compete within a neuron's receptive field (Kastner et al. 1998; Luck et al. 1997). Rather, competition in time results from the need to accumulate sensory information over time (Gold and Shadlen 2007; Ploran et al. 2007) and the fact that new sensory input tends to override older sensory input (Becker et al. 2000; Breitmeyer and Ganz 1976; Enns and Di Lollo 2003). Temporal selection therefore must ensure that relevant sensory input from one moment in time is sufficiently available for later processing before new input is encountered. The different computational challenges facing temporal and spatial selection make it unlikely that temporal selection is just the brief application of spatial selection.

The perceptual context of behaviorally relevant stimuli may be critical for representing and responding to them (Davenport and Potter 2004; Oliva and Torralba 2007; Shinoda et al. 2001), and for learning when and where to anticipate them (Brockmole et al. 2006; Chun and Jiang 1998). Because sensory input can change rapidly, temporal selection may need to influence perceptual processing in a temporally constrained manner that is not necessarily restricted to the selected input.

This study investigated the impact of temporal selection on visual cortical activity. Participants selected auditory or visual targets from a stream of distractors. Extensive studies have shown that regions involved in processing these stimuli respond more strongly to attended than unattended stimuli (Hon et al. 2009; Jäncke et al. 1999; Reynolds and Chelazzi 2004). Our study is unique in that, rather than examining how temporal selection affects processing of the selected targets, we ask how temporal selection influences activity in regions that are not involved in processing them.

One possibility is that temporal selection of a target interferes with activity in regions representing other perceptual information. Interference is predicted based on the idea that attention is competitive both within and across modalities (Desimone and Duncan 1995; Johnson and Zatorre 2006; Shomstein and Yantis 2004; Spence and Driver 1997). Indeed, attending to, rather than ignoring, auditory stimuli reduces early visual cortical responses to simultaneously presented visual stimuli, and vice versa (Johnson and Zatorre 2005, 2006). Likewise, within the visual modality, directing attention to one location reduces cortical responses to stimuli at other locations (Brefczynski and DeYoe 1999; Luck et al. 1997; Schwartz et al. 2005; Silver et al. 2007). Because selecting targets in time exerts greater attentional demands than rejecting distractors (cf. the attentional blink; Chun and Potter 1995; Raymond et al. 1992), detecting auditory targets could reduce

Address for reprint requests and other correspondence: K. M. Swallow, Dept. of Psychology, Univ. of Minnesota, N218 Elliott Hall, 75 East River Rd., Minneapolis, MN 55455 (e-mail: swall011@umn.edu).

activity in the visual cortex and detecting centrally presented visual targets could reduce activity in the peripheral visual cortex.

The second possibility is that temporal selection could result in increased (rather than decreased) activity in visual cortical areas that are not involved in processing the selected stimuli. The appearance of a target in a temporal stream constitutes a goal-relevant change in the environment. This change may trigger cognitive processes that update representations of the current context in memory. Target detection produces a late positive deflection in the event-related potential (P3) in electrophysiological studies, which may reflect the updating of mental models of the current context (Donchin and Coles 1988). Several theories propose that people update representations of goals and context in active memory at behaviorally relevant moments in time (Bouret and Sara 2005; O'Reilly et al. 1999; Zacks et al. 2007). Consistent with these theories, information that coincides with changes in observed events is better remembered than information presented at other moments (Swallow et al. 2009). In addition, target detection itself can enhance memory for and learning of concurrent stimuli. In the attentional boost effect, visual images presented at the same time as a visual or auditory target are better encoded into memory than those that coincide with distractors (Lin et al. 2010; Swallow and Jiang 2010). In addition, perceptual sensitivity to a subliminally presented motion direction increases after it has been repeatedly paired with centrally presented targets rather than distractors (Seitz and Watanabe 2003).

To examine these divergent predictions, in three experiments participants monitored a series of tones and pressed a button whenever they heard a target tone. We examined how the detection of auditory targets influenced blood oxygen level-dependent (BOLD) activity in the visual cortex. A fourth experiment presented visual targets and distractors at fixation and examined whether detecting visual targets enhances activity in regions of visual cortex representing the periphery. If temporal selection exhibits stimulus and spatial specificity, then activity in early visual cortex should decrease or remain unchanged when an auditory (or visual) target is presented. In contrast, if the effects of temporal selection are not spatially and modality specific, then activity in early visual cortex may increase when an auditory (or visual) target is presented.

Although the main purpose of these experiments was to examine how temporal selection influences early visual cortical activity, we also tested whether its effects interact with the presence or absence of concurrent, task-relevant visual input. Instead of attending to one modality (Johnson and Zatorre 2005, 2006; Shomstein and Yantis 2004), in bimodal conditions participants attended to both visual and auditory stimuli. Our data provide the first clear evidence that temporal selection of a stimulus, even an auditory one, enhances, rather than reduces, visual cortical activity in regions that do not typically represent it. In addition, the pattern of modulation differs qualitatively from spatial selection.

METHODS

Overview of Experiments

We performed five fMRI experiments (Table 1). For most experiments participants monitored a stream of auditory (*experiments 1, 3,* and *4*) or visual (*experiment 2*) stimuli for a prespecified target. They

Table 1. *Summary of task parameters across experiments*

	Task	T:D	No-Image Sessions	Image Sessions	No. Vols./Session
Experiment 1	Auditory	1:4	2	4	211
Experiment 2	Visual	1:4	0	6	211
Experiment 3	Auditory	1:1	2	0	101
Experiment 4	Auditory	1:1	0	2	171
Experiment 5	Button press	–	2	–	101

Auditory and visual tasks were continuous detection tasks. T:D, target-to-distractor ratio.

pressed a button as quickly as possible whenever a target occurred. For example, in the auditory task participants pressed the button whenever they heard a high-pitched tone rather than a low-pitched tone. Tone timing and status as a target or distractor were irregular and unpredictable, preventing hemodynamic and oscillatory effects associated with stimulus entrainment and expectation from influencing the data (Lakatos et al. 2008; Sirotin and Das 2009). These experiments contrasted the response of early visual cortical areas to stimuli that required temporal selection (target) with their response to stimuli that did not require selection (distractor). On some scans, images of faces and scenes were presented during the detection task to evaluate whether its effects interact with visual processing.

Experiment 1 established that the temporal selection of auditory targets is associated with increased activity in early visual cortex. Subsequent experiments tested whether these effects can be attributed to multimodal processing (*experiment 2*) and occur when targets are as common as distractors (*experiment 3*). Finally, the potential contributions of eye movements (*experiment 4*) and manual button presses (*experiment 5*) were evaluated.

Participants

Participants were healthy 18- to 36-yr-old volunteers with normal or corrected-to-normal visual acuity and hearing. There were 10 volunteers in *experiment 1*, 9 volunteers in *experiment 2*, 8 volunteers in *experiments 3* and *5*, and 10 volunteers in *experiment 4*. The same participants were tested in *experiments 3* and *5*, and three of these also completed *experiment 1*. All participants provided informed consent and were compensated for their time. The University of Minnesota Institutional Review Board approved all experimental procedures.

MRI Image Acquisition and Preprocessing

Experiments 1–5 were performed in a Siemens 3T MRI Scanner with a standard 12-channel head coil at the University of Minnesota Center for Magnetic Resonance Research. A high-resolution T1-weighted MPRAGE ($1 \times 1 \times 1$ mm) anatomical scan was acquired for each participant. This scan was used for cortical reconstruction in Freesurfer (Fischl et al. 1999). A standard T2*-weighted EPI sequence measured the BOLD signal during the functional scans. BOLD data for the main tasks were collected in 32 contiguous transverse slices (4-mm thick, 3.4-mm isotropic voxels; TR = 2 s, TE = 30 ms, flip angle = 75°; for *experiment 4* there were 34 3.5-mm-thick slices), providing full brain coverage except for the base of the cerebellum. For retinotopic mapping, BOLD data were acquired in 16 contiguous coronal slices oriented perpendicular to the calcarine sulcus (4-mm thick, 3-mm isotropic voxels; TR = 1 s, TE = 30 ms, flip angle = 60°). Functional data were motion corrected, smoothed with a 6-mm full-width half-maximum (FWHM) Gaussian filter, and aligned to the reconstructed surface. For whole-brain analyses, structural data were aligned to the MNI305 atlas.

Experimental Design and Procedure

Experiment 1: auditory detection task. To test the effect of temporal selection on activity in early visual cortex, participants were asked to monitor intermittently presented auditory tones (650 Hz for high-pitched tones; 350 Hz for low-pitched tones; 45-ms duration plus 1,955-ms blank) for a tone of a prespecified pitch (target; Fig. 1). They pressed a button as soon they heard a target tone but made no response to tones of a different pitch (distractor). The pitch of the target tone was counterbalanced across scans. There were 211 2-s-long trials per scan. The first three and last eight trials were fixation periods. The remaining 200 trials included 50 no-tone baseline trials, 30 target tone trials, and 120 distractor tone trials. Tones were presented at the beginning of a volume acquisition and no more than once every 2 s. To optimize estimation efficiency, the trial sequence was determined with Freesurfer's optseq2 algorithm.

The presence of visual images during the detection task was manipulated across scans. In the two no-image (blank) scans the only visual stimulus was a red fixation cross (0.26° × 0.26° viewing angle)
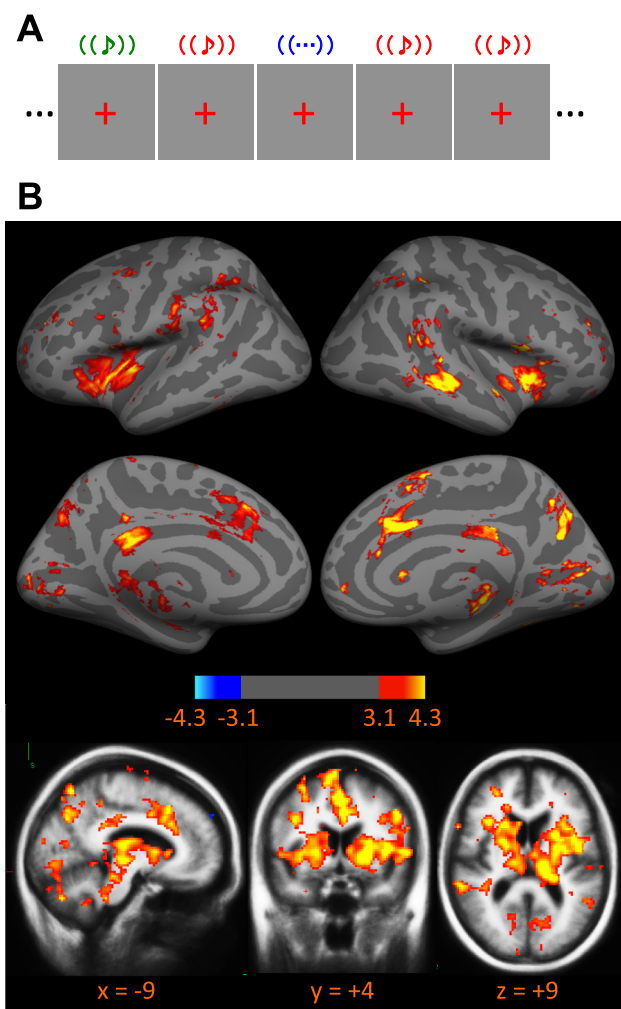
Fig. 1. Design and group-level data from *experiment 1*. *A*: for the auditory detection task, participants monitored a series of tones and pressed a button whenever the tone was a prespecified pitch (a target; green note). They made no response to other tones (distractor; red note). Variable intervals of time in which no tones were presented (blue) separated the tones. In some scans participants viewed a gray screen throughout the task. In other scans visual images were also presented. *B*: regions whose activity was greater after target than distractor tones [$t > 3.1$, $P < 0.001$, false discovery rate (FDR) < 0.05] on the cortical surface and in subcortical regions.

on a gray background. In four image scans[1] visual images (4.5° × 4.5° viewing angle) were presented in the central visual field during the detection task. On each trial a face, scene, or scrambled image onset at the same time as a target or distractor tone. The image was presented for 500 ms and then masked with a scrambled version of itself for 1,500 ms. A red fixation cross appeared in the center of the screen at all times. In addition to responding to the target tones, participants were instructed to remember the faces and scenes for a later memory test. Faces and scenes were acquired through online sources, and scrambled images were generated from the face and scene images. Faces, scenes, and scrambled images were evenly and randomly divided among target and distractor trials for each participant. Scrambled images were presented on the no-tone fixation trials. Each image was presented twice, each time with the same type of tone (e.g., a target or distractor). A demo can be viewed online at **http://jianglab.psych.umn.edu/targetdetection/targetdetection.htm**.

After scanning was complete, participants performed a two-alternative forced choice recognition test on the faces and scenes. One old and one new image were presented on the left and right side of the screen on each trial. Participants selected the image they believed was shown to them during the continuous detection task. Tests of faces and scenes were randomly intermixed.

Experiment 2: visual detection task (with images). Experiment 2 investigated the effect of temporal selection of visual stimuli on activity in nonstimulated visual regions. For the visual detection task participants monitored a stream of intermittently presented black or white squares (2 s/item; 0.34° × 0.34° viewing angle) that appeared for 80 ms at fixation. Participants pressed a key as quickly as possible whenever the square was white (target) and made no response when the square was black (distractor). On each trial the square onset at the same time as the image (500-ms duration), which was then masked for 1,500 ms. Other than replacement of the auditory tones with the squares, *experiment 2* was the same as the image scans in *experiment 1*. We did not include no-image scans.

Experiment 3: equal frequency targets and distractors (no image). Experiment 3 equated the proportion of target and distractor tones. Participants performed the same auditory detection task used in *experiment 1*, but with 30 target trials, 30 distractor trials, and 30 no-tone fixation trials per scan. No visual images were presented. Other than the target-to-distractor ratio and the total number of trials, *experiment 3* was the same as the no-image scans in *experiment 1*.

Experiment 4: eye tracking during auditory detection task (with images). Experiment 4 was similar to the image scans in *experiment 1*, except that target and distractor tones were equally likely to occur and eye gaze position was measured. There were 60 target tone trials, 60 distractor tone trials, and 40 no-tone trials per scan. Target and distractor tone trials were evenly divided across face, scene, and scrambled images. No-tone trials were presented with scrambled images only.

During scanning, eye gaze position was measured with an MRI-compatible ASL LRO-6 eye-tracker (60 Hz sampling rate). The *x* and *y* coordinates of gaze position and pupil diameter of one eye were recorded. Linear interpolation was used to estimate gaze position during periods of signal loss due to blinks or noise. The data were smoothed with a normal filter (bandwidth = 5 samples) and resampled to 12 data points per second. Four participants were excluded because of the poor quality of their eye data (>80% of the eye data samples were acquired during a signal loss; for the other 6 participants <30% of the samples were acquired during a signal loss).

Experiment 5: self-generated button press task. Participants in *experiment 3* also performed a self-generated button press task in two additional scans.[2] In each 202-s long scan, participants were instructed to press a button at any time they wanted. Prior to scanning

---

[1] One participant completed six image scans.

[2] One participant completed four scans in the self-generated button press task.

participants practiced the task to ensure that button presses were not too frequent or infrequent. The mean interval between button presses was 5.67 s (SD = 1.21; mean min and max = 2.43–14.2 s), similar to that between targets in *experiment 3* (mean = 6 s, SD = 0.1; mean min and max = 2–21 s) [$t(7) = -0.78$, $P = 0.46$]. Throughout the scan participants fixated a cross (0.26° × 0.26° viewing angle) in the center of a gray background. Other than cues to start and end the task, no other visual or auditory stimuli were presented.

Functional Data Analysis

Region of interest (ROI) and whole-brain analyses of the functional data were performed in a standard two-step analysis in Freesurfer using the general linear model (GLM; Friston et al. 1995). Linear drift and autocorrelated noise (20-s window) were removed for all analyses.

For the whole-brain analysis the shape of the hemodynamic response was modeled as a gamma function (delta = 2.25, tau = 1.25) at each voxel, resulting in one regressor per voxel per condition. For each voxel, beta weights for the response to distractors were subtracted from those for targets and submitted to a *t*-test. The resulting statistical parametric maps were thresholded at an uncorrected *P* value of 0.001 ($t > 3.1$) for cortical regions and a *P* value of 0.0001 ($t > 3.7$) for subcortical regions. Thresholds all resulted in a false discovery rate (FDR) of <0.05 (Genovese et al. 2002). Correction for multiple comparisons was performed during cluster identification. Clusters were defined as a set of activated voxels whose area was greater than would be expected by chance. Chance was determined in a Monte Carlo simulation in which the size of clusters of activated voxels under the null hypothesis was determined over 10,000 permutations separately for the left and right hemispheres and for subcortical structures. Only clusters with a brainwise *P* value < 0.05 are reported.

ROI analyses estimated the hemodynamic response to the different types of events, using the finite impulse response approach. For *experiments 1–4*, the hemodynamic response was modeled over a 22-s peristimulus window beginning 4 s before the onset of the event. This analysis produced 11 regressors per condition, 1 for each time point in the peristimulus window. For *experiment 5*, a 26-s-long peristimulus window that began 8 s before the button press was used, resulting in 13 regressors. Beta values were used to calculate signal intensity, which was averaged across all voxels within an ROI for each individual, each time point, and each condition.

Random-effects analyses on the ROI data were performed with analysis of variance (ANOVA). To simplify these analyses, the peak response to events of each condition was estimated for *experiments 1–4*. Peak signal change was defined as the difference between the mean prestimulus signal and the maximum signal observed 2–6 s after stimulus presentation (units are % signal change from the prestimulus baseline). For *experiment 1* these values were submitted to an ANOVA with tone status (target/distractor), image presence (blank/image), region eccentricity (central/periphery), and area (V1/V2/V3) as factors. For *experiment 2* square status and eccentricity were included as factors (only image scans were included in that experiment, and ROIs were only available for the pericalcarine cortex; see below). For *experiments 3* and *4*, tone status, eccentricity, and area were included as factors. Analyses of the fusiform face area (FFA) and parahippocampal place area (PPA) for *experiments 1* and *2* included only detection stimulus status (target/distractor) and image type (face/scene/scrambled) as factors. For *experiment 5*, an ANOVA with time point (13 levels), area (V1/V2/V3), and eccentricity (central/periphery) was performed to determine whether early visual cortex responded to self-generated button presses.

Region Localization

FFA and PPA localizer. To localize visual regions selectively involved in processing faces (FFA) and scenes (PPA), participants completed two scans of a standard blocked-design localizer task (Yovel and Kanwisher 2004). Participants monitored a series of images of faces, scenes, objects, and scrambled images for immediate image repetitions. For each participant the FFA was defined as the portion of cortex in and around the mid-fusiform gyrus whose activity was greater when faces were presented than when objects were presented ($t > 2.7$). The PPA was defined as the portion of cortex in and around parahippocampal gyrus that was more active when scenes were presented than when scrambled images were presented ($t > 2.7$).

Retinotopic mapping. Early visual cortical areas were identified with a standard traveling wave retinotopic mapping procedure that included two polar angle and two eccentricity mapping scans (Engel et al. 1997; Schira et al. 2009; Sereno et al. 1995). We identified the boundaries between V1, V2, and V3 based on shared horizontal and vertical meridian maps. These areas were then separated into regions representing the central and peripheral visual fields with data from the eccentricity and localizer scans. Central regions included all voxels activated by images in the localizer scans (6.1° wide), exceeding the region activated by images in the continuous detection task (4.5° wide). Peripheral regions were approximately the same length as the central regions. Because clear boundaries between V1, V2, and V3 could not be discerned from the retinotopic data in *experiment 2*, V1 was anatomically defined in Freesurfer as pericalcarine cortex (Desikan et al. 2006). To avoid overlap, voxels were included in the retinotopically defined ROIs only if at least 50% of their volume was contained within the boundaries for that ROI.

Primary auditory cortex. To examine its response to auditory and visual stimuli, primary auditory cortex (A1), corresponding to the transverse temporal gyrus (Howard et al. 2000), was defined for each participant in *experiments 1* and *2* with Freesurfer's cortical parcellation (Desikan et al. 2006).

RESULTS

Behavioral Data from Experiments 1–4

Participants accurately followed the detection task instructions (Table 2). They responded quickly to the targets and made few responses to distractors. Two participants (1 each in *experiments 1* and *2*) for whom equipment problems prevented recording behavioral data were excluded from these analyses. The experimenter verified correct performance of the task for these two participants during scanning.

Recognition memory for the images was also examined. Data from *experiments 1*, *2*, and *4* were analyzed in a single ANOVA with detection stimulus status (target/distractor), image type (face/scene), and *experiment* as factors (Table 3). Although the effect was small (2.8%) relative to previous reports (cf. Swallow and Jiang 2010), images that were presented with a target were better recognized than those presented with a distractor, resulting in a main effect of detection stimulus status [$F(1,20) = 4.42$, $P = 0.048$, $\eta_p^2 = 0.181$]. In addition, faces were better recognized than scenes [main effect

Table 2. *Mean hit rates, response times, and false alarm rates in continuous detection tasks for each experiment*

	Hit Rate	Response Time, ms	False Alarm Rate
Experiment 1	0.969 (0.016)	507 (45)	0.022 (0.028)
Experiment 2	0.926 (0.088)	464 (44)	0.009 (0.006)
Experiment 3	0.988 (0.035)	494 (69)	0.01 (0.012)
Experiment 4	0.962 (0.08)	436 (84)	0.019 (0.014)
Overall	0.961 (0.062)	479 (63)	0.015 (0.018)

Data are means (SD).

Table 3. *Proportion of correctly recognized faces and scenes presented with auditory (or visual) targets and distractors in experiments 1, 2, and 4*

	Faces		Scenes	
	Targets	Distractors	Targets	Distractors
Experiment 1	0.892 (0.092)	0.834 (0.124)	0.628 (0.123)	0.578 (0.045)
Experiment 2	0.844 (0.134)	0.848 (0.124)	0.622 (0.183)	0.624 (0.134)
Experiment 4	0.8 (0.145)	0.746 (0.086)	0.581 (0.068)	0.566 (0.057)
Overall	0.851 (0.122)	0.818 (0.118)	0.614 (0.133)	0.591 (0.09)

SD values are in parentheses. One participant each in *experiments 1* and *2* did not complete the recognition test.

of image type, $F(1,20) = 72.3$, $P < 0.001$, $\eta_p^2 = 0.783$]. No other effects or interactions, including those involving experiment, were significant (all $F < 1.58$, all $P > 0.23$).

Experiment 1: Whole-Brain Analysis of Temporal Selection of Auditory Targets

To confirm that the tones activated auditory cortex, whole-brain and ROI analyses examined activity following tones relative to fixation periods (see METHODS). A cluster of reliably activated voxels ($t > 2.3$, $p < 0.01$, false positives controlled for by cluster size, see METHODS) was identified in the right superior temporal sulcus and middle temporal gyrus (peak: $[61, -35, -5]$). In addition, the estimated response of anatomically defined A1 to tones was submitted to an ANOVA with time, tone status, and image presence as factors. A main effect of time indicated that it was activated by tones [$F(10,90) =$

20.7, $P < 0.001$, $\eta_p^2 = 0.696$]. It also responded more strongly to target than distractor tones, as indicated by a reliable time × tone status interaction [$F(10,90) = 8.3$, $P < 0.001$, $\eta_p^2 = 0.48$]. Thus auditory cortex was reliably activated by the tones, and this response was modulated by temporal selection.

Voxels whose response to target and distractor tones reliably differed were also identified (see METHODS). Regions that responded more strongly to target than distractor tones included those typically activated in attentional selection tasks (Fig. 1; Table 4): the anterior insula, the anterior cingulate, the intraparietal sulcus, and the supramarginal gyrus (Bledowski 2004; Corbetta 2008; Duncan 2010; Hon et al. 2009).[3] In addition, the pericalcarine cortex, right middle temporal gyrus, precuneus, basal ganglia, thalamus, cerebellum, and posterior brain stem in the vicinity of the locus coeruleus (LC) were more active after target than after distractor tones.

Effect of Temporal Selection on Ventral Visual Areas

If the effects of temporal selection on brain activity are not specific to processing the relevant stimulus itself, then it should affect activity in visual cortical areas. To test this, we first contrasted the response of early visual cortex to target and distractor tones in retinotopically defined regions of V1, V2, and V3 representing the central and peripheral visual fields (Fig. 2). Peak signal changes to events in each condition (see

[3] Voxels in bilateral intraparietal cortex were reliably activated by target tones relative to distractor tones ($t > 3.1$, $P < 0.001$, FDR < 0.05) prior to correction for multiple comparisons based on cluster size. The cluster that included activated voxels in the right supramarginal gyrus extended into the middle temporal lobe.

Table 4. *Peak coordinates and size of regions that were more active after target tones than distractor tones in experiment 1, after correction for multiple comparisons*

Region	Hemi	Talairach			Size	P
		X	Y	Z		
Middle frontal gyrus	L	−31.4	2.1	45	193	0.036
Pars opercularis	R	47	9.7	2.6	267	0.027
Superior frontal gyrus	L	−9	22.6	37.4	750	0.005
	R	9.7	13	49	346	0.023
Anterior cingulate	R	5.3	12.1	31.3	623	0.003
Posterior cingulate	L	−6.4	−34.8	24.9	307	0.018
	R	5.5	−29.7	29	292	0.025
Postcentral sulcus	L	−54.2	−22.1	28.3	532	0.008
Precuneus	L	−12.5	−65.2	37.3	355	0.015
	R	17.5	−58	29.7	588	0.006
Supramarginal gyrus	L	−49.8	−39.7	28.9	1,045	0.005
Insula	L	−43.6	0.1	12.3	1,969	0.001
	R	34.8	16.8	−1.4	717	0.003
Middle temporal gyrus	R	45.1	−27.4	−5.9	1,353	0.001
Cuneus	R	26.3	−57.8	9.4	189	0.037
Pericalcarine cortex	L	−8.9	−84	3.1	393	0.011
	R	15.6	−75.8	12.7	545	0.006
Caudate	R	21.8	16.8	5.6	1,152	0.001
Pallidum	R	17.8	−3	−2.4	2,272	0.001
Putamen	L	−23.8	6.9	2.4	6,688	0.001
Thalamus	L	−15.8	−27.7	9.7	4,488	0.001
	R	15.8	−14.6	−0.1	1,680	0.001
Cerebellum	L	−39.6	−62.2	−19.6	4,672	0.001
		−13.9	−76.2	−29	464	0.022
	R	21.8	−47.6	−38.8	21,176	0.001
Brain stem		−5.9	−25.1	−16.4	368	0.037

Sizes are in mm^2 for cortical regions and mm^3 for subcortical regions. Only those regions whose size was unlikely to be observed by chance ($P < 0.05$) are reported.
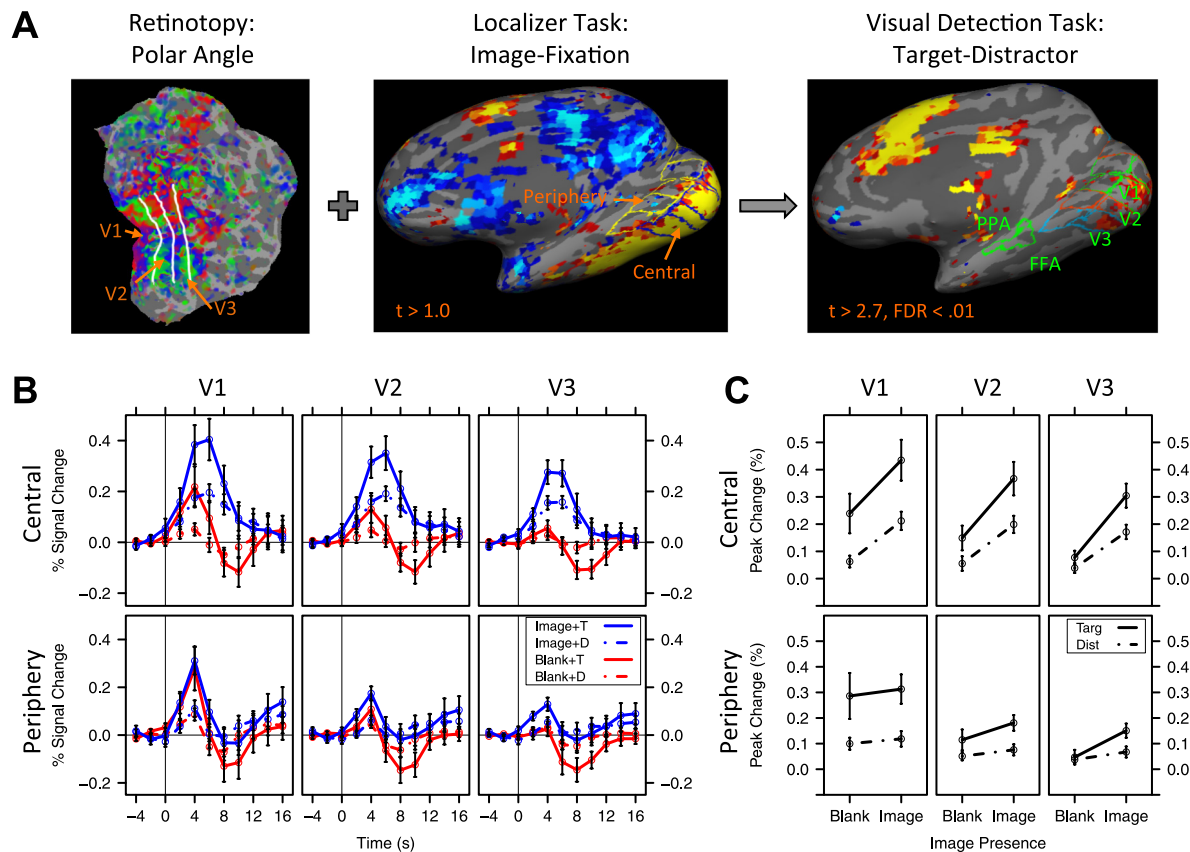
Fig. 2. Definition of regions of interest (ROIs) and their response to auditory target and distractor tones in *experiment 1*. *A*: for each individual, central and periphery V1, V2, and V3 ROIs were defined with polar angle mapping data and localizer data. Data are shown for 1 individual (on the flattened occipital lobe for retinotopy and on the cortical surface for the remaining contrasts). The fusiform face area (FFA) and parahippocampal place area (PPA) were also identified from the localizer data. All voxels in an ROI were used to estimate the hemodynamic response to the different types of trials. *B*: mean time course of the response of early visual cortical areas to target (solid lines) and distractor tones (dashed lines) presented with images (blue lines) and without images (red lines). *C*: peak signal change following target tones (solid lines) and distractor tones (dashed lines) presented with and without visual images in early visual ROIs. Error bars represent ±SE around the mean in *B* and *C*.

METHODS) were analyzed with an ANOVA that included tone status, image presence, eccentricity, and area as factors. The results of this analysis are presented in two parts.

Despite greater attentional demands when target tones were detected, early visual cortex responded more strongly to target tones than to distractor tones, resulting in a reliable main effect of tone status on peak percent signal change [$F(1,9) = 23.4$, $P < 0.001$, $\eta_p^2 = 0.722$]. In addition, temporal selection enhanced activity throughout early visual cortex, although its effects decreased from early to late visual areas. The effect of tone status was similar in central and peripheral ROIs, as there were no reliable interactions between tone status and region eccentricity [largest $F(2,18) = 1.17$, all $P > 0.333$]. However, tone status more strongly modulated activity in V1 than in V2 and V3, leading to a reliable interaction between tone status and area [$F(2,18) = 15.4$, $P < 0.001$, $\eta_p^2 = 0.631$]. The overall effect of tones on early visual cortical activity decreased from V1 to V3, particularly in the peripheral eccentricities, as indicated by an area × eccentricity interaction [$F(2,18) = 3.75$, $P = 0.043$, $\eta_p^2 = 0.294$], a main effect of area [$F(2,18) = 13.9$, $P < 0.001$, $\eta_p^2 = 0.607$], and a marginal main effect of eccentricity [$F(1,9) = 4.48$, $P = 0.063$, $\eta_p^2 = 0.332$]. The decrease in the magnitude of the effect of target tones through the visual processing stream is readily apparent in Fig. 2*C*,

which plots peak signal change for target and distractor tones presented with and without images in each region.

Early visual cortical regions responded more strongly to target tones, which required selection, than to distractor tones, which did not. Thus temporal selection of auditory stimuli appears to elicit increased activity in visual cortical areas. Surprisingly, the effects of temporal selection were not spatially or modality specific and appeared to decrease along the ventral visual processing stream. These data are in stark contrast to those of spatial selection. In addition to increasing activity in perceptual regions involved in processing the selected stimulus (Luck et al. 1997; Silver et al. 2007; Tootell et al. 1998), spatial selection follows a reverse hierarchy, more strongly modulating activity in later than in early visual areas (Buffalo et al. 2010; Hochstein and Ahissar 2002; Kastner et al. 1998).

Interaction of Temporal Selection and Early Visual Stimulus Processing

A second goal of *experiment 1* was to examine the interaction of the temporal selection of a behaviorally relevant stimulus (the target tone) and the processing of separate, concurrent stimuli. Responses to target and distractor tones were therefore evaluated when visual stimuli were or were not presented.

Surprisingly, the effect of temporal selection of auditory tones on early visual cortical activity was not affected by a concurrent visual task [there were no interactions involving tone status and image presence; largest $F(1,9) = 1.23$, $P = 0.295$]. Furthermore, the progression of the effect of target tones from V1 to V2 to V3 did not change when an image was presented [there were no reliable interactions involving image presence and area, including the 3- and 4-way interactions with eccentricity; largest $F(2,18) = 2.65$, $P = 0.098$ for the image presence $\times$ area $\times$ eccentricity interaction]. Image presence increased activity in the central, but not peripheral, visual fields, resulting in a reliable interaction between image presence and eccentricity [$F(1,9) = 16.4$, $P = 0.003$, $\eta_p^2 = 0.645$]. This finding confirms that these regions distinguished between stimulated and nonstimulated regions of the visual field. However, a concurrent image encoding task does not appear to influence the persistence or distribution of the effect of auditory targets on early visual cortical activity.

Additional analyses were performed to determine whether face- and scene-selective visual areas are differentially modulated by temporal selection when their preferred stimuli are presented (Fig. 3). The FFA and PPA were identified in seven participants with anatomical criteria and functional data from a separate localizer task. For both regions, peak signal change estimates were submitted to an ANOVA with tone status and image type (face/scene/scrambled) as factors. Main effects of image type indicated that the FFA responded most strongly to faces [$F(1,6) = 52.8$, $P < 0.001$, $\eta_p^2 = 0.898$] and the PPA responded most strongly to scenes [$F(1,6) = 59.6$, $P < 0.001$,

$\eta_p^2 = 0.908$]. However, there was little evidence of an effect of auditory targets on activity in either region, particularly for their preferred stimuli [FFA: no main effect of tone status, $F(1,6) = 2.95$, $P = 0.136$, no tone status $\times$ image type interaction, $F(2,12) = 1.68$, $P = 0.228$; PPA: no main effect of tone status, $F(1,6) = 3.17$, $P = 0.125$, no tone status $\times$ image type interaction, $F(2,12) = 0.58$, $P = 0.575$]. Thus the effect of auditory targets was absent in the FFA and the PPA.

The data from *experiment 1* demonstrated a clear and robust effect of auditory target stimuli on activity in early visual cortex. This response was present in both central and peripheral regions and was stronger in V1 than in V2 and V3 and absent in the FFA and PPA. Moreover, it appeared to interact minimally with the presence of attended and easily perceived visual stimuli.

Its lack of specificity, its decrease through ventral visual cortex, and its insensitivity to the presence of competing stimuli clearly distinguish the effect of temporal selection from those of visuo-spatial attention, visual imagery, arousal, and alerting. The modulatory effects of visuo-spatial attention and imagery on visual cortex are spatially constrained and larger in late than in early visual areas (Buffalo et al. 2010; Cichy et al. 2012; Kastner et al. 1998; Reynolds and Heeger 2009; Slotnick et al. 2005). In addition, enhanced activity in the fusiform gyrus, but not early visual cortex, is often observed in response to arousing stimuli and alerting signals (Anderson et al. 2003; Fan et al. 2005; Jiang and He 2006; Thiel et al. 2004).

The data from *experiment 1* also stand in contrast to previous reports on the effects of directing attention to a single modality. Typically, selective attention to a single modality results in decreased activity in regions processing the nonselected modality (Johnson and Zatorre 2005, 2006; Shomstein and Yantis 2004). In those studies, visual and auditory stimuli were presented to participants who were instructed to attend to either the visual or the auditory modality at different times. When sustained attention was directed to the auditory modality, activity in visual cortex decreased. The data from *experiment 1* suggest that transient attention to auditory stimuli has a markedly different effect on activity in visual perceptual areas, both when attention is also directed to visual stimuli (as in the image scans) and when it is not (as in the no-image scans).

Despite the unusual distribution of the effect of temporal selection on early visual cortical activity, these data are not without precedent. One other study has reported nonperceptual enhancements of visual cortical activity in response to task-relevant events that marked transitions in the task (Jack et al. 2006). In that study, activity in early visual cortex, particularly in peripheral regions of V1, increased in response to a variety of task-relevant events. The nonperceptual modulation of activity in early visual cortex was dissociated from spatial selection both in terms of its cortical distribution and by its occurrence in regions that did not contain visual stimuli.

The data from *experiment 1* provide substantial support in favor of the idea that nonperceptual factors can modulate activity in early visual cortex. However, they begin to provide greater insight into when these modulations are likely to occur by linking them to temporal selection. They also begin to investigate how these modulations may interact with visual stimulus processing. Because target detection appears to increase the amplitude of the response of early visual cortex to
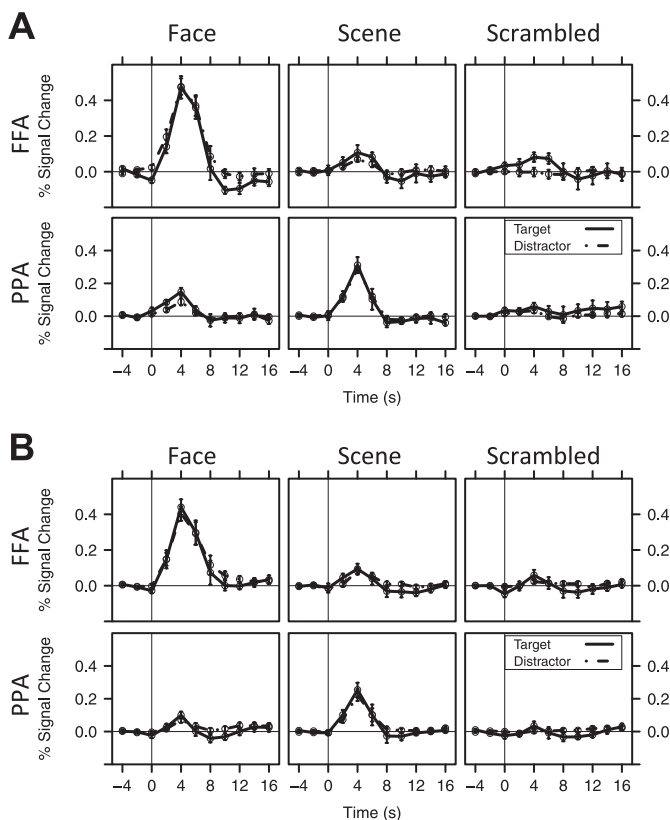

Fig. 3. Time course of the response of FFA and PPA to auditory targets and distractors in *experiment 1* (*A*) and fixation targets and distractors in *experiment 2* (*B*). Error bars represent $\pm$SE.

auditory tones, we refer to the effect of targets on early visual cortical activity as the target-mediated boost.

Although target tones required temporal selection, there were other potentially relevant differences between target and distractor tones in *experiment 1* that could have produced the target-mediated boost. These were addressed in the next set of experiments, which examined whether the target-mediated boost occurs for visual targets, frequent targets, and self-paced button presses. An additional experiment examined the role of eye movements.

Role of Multimodal Processing in the Target-Mediated Boost: Visual Targets

Efferent projections to early visual cortex, particularly peripheral V1, originate in part from auditory cortex, including the superior temporal sulcus (STS) (Doty 1983; Falchier et al. 2002; Rockland and Ojima 2003). These projections raise the possibility that the target-mediated boost observed in *experiment 1* reflects audiovisual integration. Indeed, the literature on multimodal processing questions the degree to which early sensory areas are unisensory in nature (Brosch et al. 2005; Driver and Noesselt 2008; Ghazanfar and Schroeder 2006). Auditory stimuli appear to facilitate the processing of low-threshold visual stimuli (Noesselt et al. 2010) and enhance early visual cortical responses to visual stimuli (Molholm et al. 2002; Naue et al. 2011), particularly when auditory and visual stimuli are predictably associated (Baier et al. 2006).

Experiment 2 was performed for two reasons. The first was to determine whether the target-mediated boost was specific to the temporal selection of auditory stimuli. The second was to

more strongly produce competitive interactions between the selected stimulus and concurrent visual input. A second group of participants performed a visual, rather than auditory, detection task (Fig. 4). They pressed a button whenever a small, centrally presented fixation square was white instead of black and encoded an unrelated stream of images into memory. Task-relevant faces and scenes were presented in all scans. Attention to the fixation targets should enhance activity in the central visual field. It may also decrease activity in regions representing other spatial locations (Reynolds and Heeger 2009; Silver et al. 2007; Tootell et al. 1998). Of critical interest is how the appearance of a centrally presented target modulates activity in visual regions representing the peripheral visual field. If the target-mediated boost reflects temporal selection of a behaviorally relevant stimulus, regardless of its modality, then it should occur for visual as well as auditory targets, even in regions that are not stimulated. If it instead reflects audio-visual integration, then the target-mediated boost should not occur in *experiment 2*.

Figure 4 illustrates the ROIs and their response to centrally presented target and distractor squares during the visual detection task. V1 was anatomically defined as pericalcarine cortex and divided into central and periphery regions with the localizer data. Peak signal change in the resulting ROIs was then submitted to an ANOVA with square status and eccentricity as factors. Responses were stronger in central V1 than in periphery V1, resulting in a main effect of eccentricity [$F(1,8) = 7.62$, $P = 0.025$, $\eta_p^2 = 0.488$]. More importantly, activity in both the central and periphery regions of V1 was greater when a fixation target was presented than when a distractor was presented
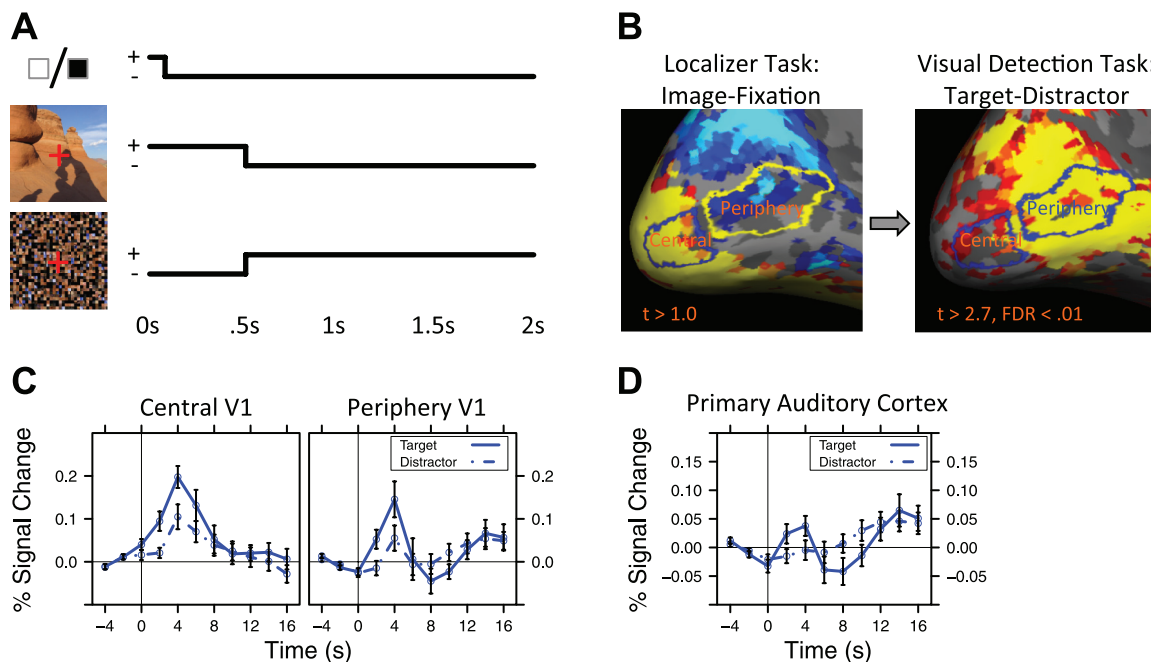

Fig. 4. Design, ROIs, and data from *experiment 2*. *A*: the detection task was similar to the image scans in *experiment 1*, except that the detection stimuli were visual. Participants monitored centrally presented squares (80-ms duration) that appeared in front of the background images. They pressed a button when the square was white rather than black and encoded the background images (500-ms duration) for a later memory test. Images were followed by a scrambled image masked for 1,500 ms. A central fixation cross (not drawn to scale) was presented after the square was removed from the screen. *B*: V1 was anatomically localized to the pericalcarine cortex for each participant (colored map shows data from 1 participant). Localizer data were used to define the boundaries between central and periphery regions as in *experiment 1*. All voxels in an ROI were used to estimate the hemodynamic response to the target and distractor squares. Thresholds and color maps are as in Fig. 2. *C*: time course of the response of central and periphery V1 to centrally presented visual target and distractor squares. *D*: time course of the response of primary auditory cortex to centrally presented target and distractor squares. Error bars in *C* and *D* represent ±SE.

[main effect of square status, $F(1,8) = 15.4$, $P = 0.004$, $\eta_p^2 = 0.658$, and no interaction between square status and eccentricity, $F(1,8) = 0.05$, $P = 0.829$]. The target-mediated boost was present throughout V1, even in regions that were not stimulated and that did not contain the selected target.

The progression of the target-mediated boost through visual cortex and its interaction with cortical responses to concurrent images were also examined for the FFA and the PPA. Peak signal change from these regions was submitted to an ANOVA with square status and image type as factors. Reliable main effects of image type in both the FFA [$F(2,16) = 62.6$, $P < 0.001$, $\eta_p^2 = 0.887$] and the PPA [$F(2,14) = 22.1$, $P < 0.001$, $\eta_p^2 = 0.76$] confirmed that these regions were selectively activated by faces and scenes, respectively. Although selecting a visual target might increase activity in these regions, an interaction of the target-mediated boost with image processing should produce effects that are specific to the type of stimulus preferred by these regions. However, although the FFA increased more in activity for target squares than distractor squares, leading to a main effect of square status [$F(1,8) = 6.18$, $P = 0.038$, $\eta_p^2 = 0.436$], this response was not reliably greater for faces than for scenes or scrambled images, as indicated by a nonsignificant interaction between square status and image type [$F(2,16) = 0.25$, $P = 0.783$]. Similarly, the marginal main effect of square status in the PPA [$F(1,7) = 3.55$, $P = 0.101$] did not depend on whether the concurrent image was a scene or another type of image [nonsignificant interaction between square status and image type, $F(2,14) = 0.17$, $P = 0.844$].

Thus the FFA and PPA showed weaker effects of targets than did V1 (0.02 for the FFA, 0.03 for the PPA, and 0.1 and 0.09 for central and periphery V1), and these effects were not specific to their preferred stimuli. Just as with auditory target tones, the boost elicited by visual target squares diminishes in a feedforward manner through the ventral visual processing stream and does not depend on the type of stimulus presented.

Finally, because *experiment 2* utilized visual targets and distractors, it was possible to examine activity in primary auditory cortex (A1). A *t*-test indicated that this region showed a reliably larger peak response to visual targets than to distractors [Fig. 4D; $t(8) = 2.42$, $P = 0.042$, $d = 0.753$], suggesting that the target-mediated boost may not be confined to visual perceptual areas.

The data from *experiment 2* demonstrate that the target-mediated boost in visual cortex is not specific to the selection of auditory stimuli. It is therefore unlikely that the target-mediated boost reflects either multimodal processing or feedback from auditory perceptual regions such as STS to V1. This conclusion is consistent with the finding that the target-mediated boost was not stronger in periphery V1, which receives more projections from auditory cortex than does central V1 (Falchier et al. 2002). Rather, the boost appears to reflect processes that are triggered by the temporal selection of behaviorally relevant stimuli.

Early Visual Cortical Responses to Common Auditory Targets

In the previous two experiments the detection stimuli (tones and centrally presented squares) were three times more likely to be distractors than targets. Targets therefore may have triggered processes associated with rare, or unexpected, stimuli, including novelty processing, expectancy violations, and

the orienting response (Donchin and Coles 1988; Polich 2007; Shulman et al. 2009; Sokolov et al. 2002). To determine whether the target-mediated boost reflects processes associated with rare stimuli, a third experiment was run with auditory tones that were equally likely to be targets and distractors. Distractors were as novel and unexpected as targets. If rare or novel stimuli are necessary for the target-mediated boost, then it should be absent in *experiment 3*. In contrast, if the target-mediated boost reflects temporal selection, it should occur when targets are frequent as well as when they are rare.

As can be seen in Fig. 5, early visual cortical responses to tones were greater when they were targets than when they were distractors, even when they were equally frequent. Peak signal changes for the retinotopically defined ROIs were submitted to an ANOVA with tone status, area, and eccentricity as factors. Across all regions the main effect of tone status was marginal [$F(1,7) = 5.11$, $P = 0.058$, $\eta_p^2 = 0.422$], with the magnitude of the effect decreasing from V1 to V3, producing a reliable interaction between tone status and area [$F(2,14) = 11.3$, $P = 0.001$, $\eta_p^2 = 0.618$]. A follow-up ANOVA only on V1 confirmed that it was more active after a target tone than a distractor tone [main effect of tone status $F(1,7) = 7.52$, $P = 0.029$, $\eta_p^2 = 0.518$]. Overall, the main effect of area indicated that responses to tones were larger in V1 than in V2 or V3 [$F(2,14) = 11.6$, $P = 0.001$, $\eta_p^2 = 0.623$] (see Fig. 5). There were no reliable effects of eccentricity [all $P > 0.474$]. These data replicate the target-mediated boost in a task in which images were never presented and were never task relevant. More importantly, the target-mediated boost was present when target tones were as frequent as distractor tones. Hence, rare target stimuli are not necessary for the target-mediated boost.

Testing Contributions of Eye Movements and Button Presses to the Target-Mediated Boost

One consideration was the potential role of the motor response in the target-mediated boost. Although participants were instructed to fixate on the center of the screen, they may have moved their eyes or blinked more following a target than a distractor. In addition, targets, but not distractors, required a manual response. Although a manual response is not necessary for the nonperceptual activity produced by task transitions
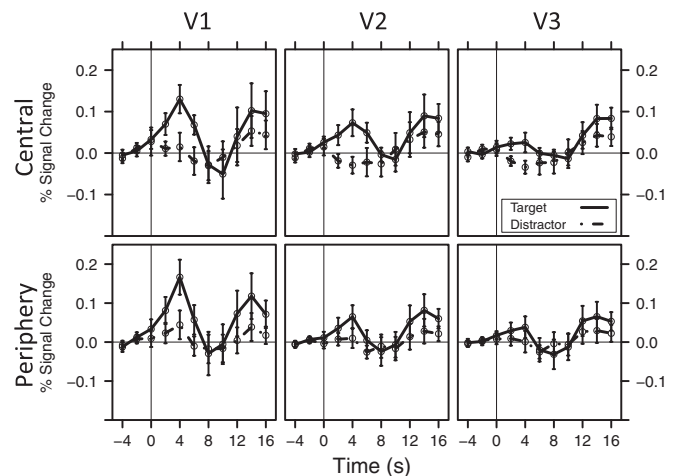

Fig. 5. Time course of the hemodynamic response of early visual regions to auditory target and distractor tones that occurred with equal frequency (*experiment 3*). No visual images were presented in these scans. Error bars represent ±SE.
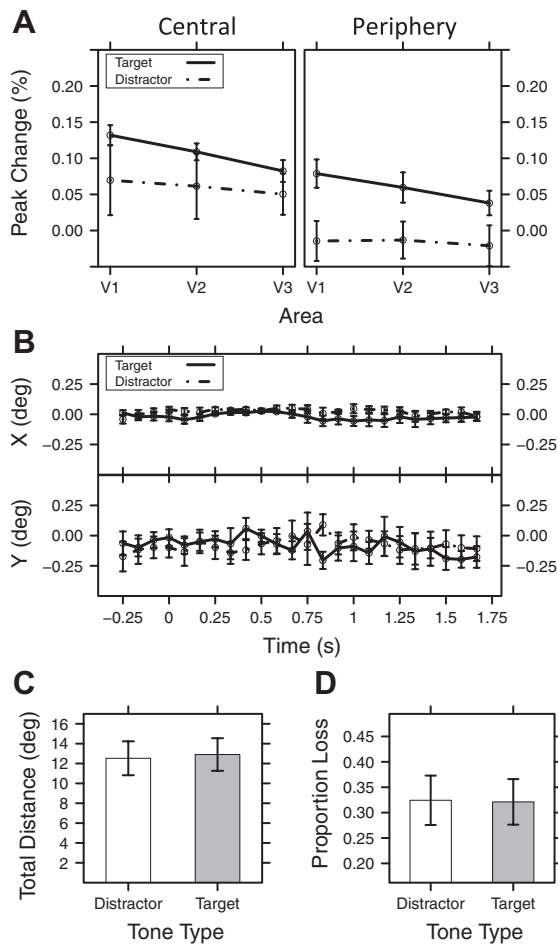
Fig. 6. Target-mediated flash and eye movement data in *experiment 4*. *A*: peak % signal change in retinotopically defined early visual areas V1, V2, and V3 after target and distractor tones. Responses were separately estimated for regions representing the central and peripheral visual fields. *B*: distance of gaze position from fixation coordinates during the 2-s period following target and distractor tones in *x* and *y* coordinates. *C*: total distance that the eyes moved during the 2-s period that followed target and distractor tones. Note that a shift in gaze position of 1°, and then back, would result in a total distance of 2°. *D*: proportion of samples that were flagged as a signal loss in the 2-s period following target and distractor tones. Error bars represent ±SE.

(Jack et al. 2006), it remains possible that the act of pressing a button could increase its magnitude.

To investigate the relationship between eye movements and the target-mediated boost, eye gaze position, blinks (defined as eye data signal losses), and BOLD data were measured simultaneously in *experiment 4*. A new group of participants performed the auditory detection task with equally frequent targets and distractors as they encoded background images into memory.

Analyses of the eye data indicated that there were no reliable differences in blinks or eye movements following targets and distractors [$t(5) = -0.5$, $P = 0.64$ for blinks and $t(5) = 1.12$, $P = 0.315$ for distance]. Importantly, a target-mediated boost was observed (Fig. 6). Peak changes in BOLD signal were submitted to an ANOVA with area, eccentricity, and tone status as factors. A reliable interaction between area and tone status indicated that peak activity in early visual cortex was greater after target tones than after distractor tones but that this effect decreased from V1 to V3 [$F(2,10) = 4.54$, $P = 0.04$,

$\eta_p^2 = 0.476$]. Main effects of area and eccentricity indicated that responses to tones decreased from V1 to V3 [$F(2,10) = 7.51$, $P = 0.01$, $\eta_p^2 = 0.6$] and were larger in central than in peripheral eccentricities [$F(1,5) = 37.6$, $P = 0.002$, $\eta_p^2 = 0.882$]. No other effects or interactions were significant (all $F < 3.69$, all $P > 0.11$).

An additional eye tracking experiment replicated *experiment 1* outside the scanner. This experiment had a larger sample size ($N = 9$) and included more trials than *experiment 4*. Its findings were consistent with the conclusion that participants move their eyes a similar amount after target and distractor tones. There were only small deviations in eye position from fixation and no differences in the amount the eyes moved or blinked across target and distractor trials, regardless of whether an image was presented [no reliable effects of tone status or images: eye movements, largest $F(1,8) = 1.27$, $P = 0.292$; blinks, largest $F(1,8) = 1.61$, $P = 0.24$]. Thus the eye movement data indicated that the target-mediated boost occurs even when there are no apparent differences in eye movements or blinks across target and distractor trials.

A final experiment examined the relationship between the target-mediated boost and button presses. For *experiment 5*, participants who completed the auditory detection task in *experiment 3* also completed a self-generated button press task. If the target-mediated boost in *experiment 3* was due to the button press response to targets, then activity in central and periphery V1 in these same participants should increase after a self-generated button press.

Rather than leading to a widespread and immediate enhancement of activity in V1, however, self-generated button presses produced an initial decrease in activity followed ~12 s later by an increase in activity (Fig. 7). These effects were confined to central V1. An ANOVA with time, area, and eccentricity as factors indicated that central V1 showed a stronger response around button presses than the other regions, resulting in reliable interactions between time, area, and eccentricity [$F(24,168) = 2.24$, $P = 0.002$, $\eta_p^2 = 0.243$] and time and area [$F(24,168) = 2.26$, $P = 0.001$, $\eta\{p\varphi2\} = 0.244$] and a trend for an interaction between area and eccentricity [$F(2,14) = 2.23$, $P = 0.093$]. No other effects or interactions were reliable (all $F < 1.5$, all $P > 0.256$). In contrast, the response to targets
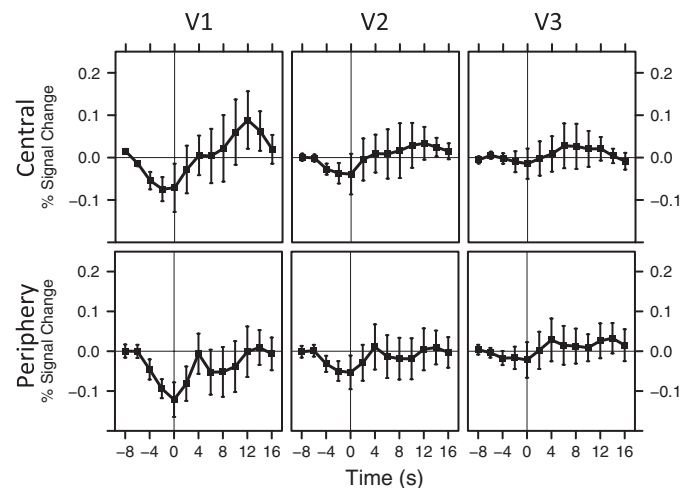

Fig. 7. Time course of the response of retinotopically defined early visual regions to self-generated button presses in *experiment 5*. Note that the pre-event time is longer than in previous experiments. Error bars represent ±SE.

was observed in both central and periphery regions and followed a more or less standard hemodynamic response function, peaking ~4 s after the onset of the tone (Fig. 2). Thus the same group of participants who showed a target-mediated boost in *experiment 3* showed a different response to self-generated button presses in *experiment 5*.

DISCUSSION

Attentional selection is typically considered to be a process that enhances neural responses to the selected stimuli. However, the computational demands of a mechanism that selects stimuli in time suggest that its effects may need to be brief and spatially unconstrained. This study investigated whether temporal selection influences activity in perceptual regions that are not typically involved in processing the selected stimulus. Previous data suggest that temporal selection could either increase or decrease activity in these regions. Whereas some behavioral studies show better encoding of stimuli that coincide with goal-relevant events (Lin et al. 2010; Seitz and Watanabe 2003; Swallow and Jiang 2010), other neuroimaging studies suggest that increasing attention to one stimulus should reduce activity in regions not involved in processing them (Brefczynski and DeYoe 1999; Johnson and Zatorre 2005, 2006; Luck et al. 1997; Schwartz et al. 2005; Shomstein and Yantis 2004; Silver et al. 2007). The data reported here clearly showed that temporal selection of goal-relevant stimuli is associated with a nonspecific increase in activity in early visual cortical regions.

Most neuroscience research on attentional selection has focused on selection in space. Spatial selection results in the modulation of neural activity in visual areas of the brain (Reynolds and Heeger 2009; Tootell et al. 1998). Modulatory or biasing signals are generated in dorsal and ventral attentional networks that include inferior parietal sulcus, angular gyrus, the frontal eye fields, and right middle frontal gyrus (Corbetta et al. 2008; Culham et al. 2001). These networks bias activity in visual regions toward the representation of salient or behaviorally relevant spatial locations or visual features (Desimone and Duncan 1995). Although the exact nature of these modulations is unclear (cf. Reynolds and Heeger 2009), spatial selection proceeds in the opposite direction in the visual processing stream than perceptual processing (Hochstein and Ahissar 2002). Spatial selection tends to produce stronger and earlier modulatory effects in late visual regions such as V4 than in early visual regions such as V1 (Buffalo et al. 2010; Kastner et al. 1998). Moreover, spatial selection enhances neural processing in regions representing the attended region of space (Luck et al. 1997) and can reduce activity in regions representing other spatial locations (Brefczynski and DeYoe 1999; Silver et al. 2007). Thus spatial selection involves the interaction of neural systems that orient attention to goal-relevant or salient regions in space with regions involved in processing sensory information at those and other locations.

In contrast to visuo-spatial attention, in the present study temporal selection was associated with spatially diffuse increases in BOLD activity that were stronger in early than in late visual cortex. Several experiments demonstrated that this target-mediated boost in early visual cortex was due to temporal selection rather than to audio-visual integration, differences in the novelty or expectancy of target and distractor

stimuli, or hand or eye movements in response to targets. Rather, the data suggest a strong relationship between the temporal selection of behaviorally relevant stimuli and spatially nonselective increases in activity in early perceptual cortical regions.

These effects diverge from earlier studies showing that sustained attention to an auditory or visual stimulus reduces activity in regions that are not involved in its representation. Other work has shown that directing attention to the auditory rather than visual modality reduces activity in visual cortex (Johnson and Zatorre 2005, 2006; Shomstein and Yantis 2004). In addition, sustained attention to one spatial location reduces activity in regions representing nonattended spatial locations (Brefczynski and DeYoe 1999; Silver et al. 2007). These and similar data support the suggestion of a push-pull relationship in selective attention: Increasing attention to one modality or spatial location reduces attention to other modalities and locations (Pinsk et al. 2003; Shomstein and Yantis 2004). The observation that transient attention to a stimulus presented in one modality (auditory or visual) or spatial location enhances activity in perceptual regions that are not involved in its processing is a striking contrast to these previous data. However, the critical manipulation in *experiments 1–4* was whether a briefly presented stimulus was a target, rather than which modality or spatial location should be attended. The outcome of these experiments underscores the distinctive computational challenges that face a temporal selection mechanism, suggesting that temporal selection is more than a temporally constrained application of spatial selection.

Although the pattern of activity in early visual cortex reported in this study is unusual in studies of attentional selection, a similar pattern has been reported for task transitions (Jack et al. 2006). In that study, participants performed a simple discrimination task on visual or auditory stimuli. Activity in early visual cortex, particularly in peripheral regions of V1, increased in response to auditory events that signaled the beginning of a trial and that signaled that a response should be made or canceled.

The experiments reported here represent a substantial extension of these findings to a markedly different paradigm—one that required participants to be nearly continuously engaged in a task with no clear trial structure or task transitions. More importantly, they offer new insight into which factors may be important for generating these modulations. In the previous study, all events in a trial were associated with increased activity in peripheral V1 (Jack et al. 2006). *Experiments 1–5* constrain accounts of the V1 and target-mediated boost. They demonstrate that the early visual cortical boost does not depend on stimulus novelty or expectation, that it is weaker for auditory and visual stimuli that do not require a response, and that it does not occur for self-generated button presses. Rather than occurring for all sensory or motor events that could structure a task over time, nonperceptual boosts of visual cortical activity appear to be specific to events that require temporal selection. Moreover, the present study suggests that these nonperceptual modulations may be more general than previously understood. They occur in early auditory cortex and when visuo-spatial attention is directed to concurrent images or central visual targets.

The target-mediated boost may be related to findings from a single-unit and multiunit recording study on nonhuman primates (Brosch et al. 2005). For that study, macaques were

trained to release a bar when the pitch of a tone sequence decreased. The firing rate of neurons in auditory cortex increased in response to visual and behavioral events that occurred as part of the task. Other neuroimaging work in humans has also found that activity in extrastriate visual areas increases when a target sound previously associated with a visual image is expected (Bueti and Macaluso 2010). The data from *experiments 1–4* suggest that similar modulations occur in humans and in early visual and auditory cortex. However, unlike earlier data, the target-mediated boost was observed in participants with little previous experience in the detection task. In addition, the effect occurred in visual cortex when the task was purely auditory (*experiment 3*) and in auditory cortex when the task was purely visual (*experiment 2*). It is therefore unlikely that the target-mediated boost reflects a learned association between visual, auditory, and behavioral events.

Potential Cognitive and Neural Sources of the Target-Mediated Boost

Temporal selection has been conceptualized as a gate that increases the likelihood that the selected input enters working memory (Bowman and Wyble 2007; Chun and Potter 1995; Olivers and Meeter 2008). However, in these models temporal selection's facility effects are constrained to the selected item and to later perceptual areas. The data presented here suggest that, at the very least, current models of temporal selection are incomplete. Because temporal selection diffusely enhances activity in early visual areas, whatever mechanism underlies it must have effects that extend beyond late perceptual regions representing the selected stimulus. Rather, the fact that increased activity in early visual cortical areas is associated with temporal selection and changes in task structure (Jack et al. 2006) is consistent with a different set of models: those that describe how the cognitive system represents goals and external events. In these models, changes in context or the completion of a goal can trigger a gating mechanism that updates neural representations to better reflect the new situation (Aston-Jones and Cohen 2005; Bouret and Sara 2005; Frank et al. 2001; O'Reilly et al. 1999; Zacks et al. 2007). The broad early visual cortical activity corresponding to these moments in time reported here is consistent with such an updating mechanism (cf. Jack et al. 2006).

The fact that nonperceptual modulations of activity in early visual cortex are strongest in V1 suggests that they do not arise from indirect feedback from late visual or frontoparietal attentional regions. Rather, the relationship between the early visual cortical boost and task structure and attention suggests two potential subcortical sources. The first is the dopamine-based gating system in the basal ganglia. According to one model, the basal ganglia act as a gate that protects representations of goals and context from disruption by new input from other cortical regions. When goals are completed or the context changes, the gating mechanism is triggered to allow active memory updating and to initiate motor actions (Frank et al. 2001; O'Reilly et al. 1999). The release of dopamine from the basal ganglia is also associated with expectancy violations, facilitating reinforcement learning by signaling unexpected rewards (Schultz and Dickinson 2000).

A second potential source of the boost is the phasic release of norepinephrine (NE) from the LC, which has been characterized as a temporal attentional filter (Nieuwenhuis et al. 2005). The LC-NE response is thought to facilitate the updating of neuronal representations in response to external cues by enhancing their responsivity to new input (Aston-Jones and Cohen 2005; Bouret and Sara 2005). It has been proposed that the LC-NE response to targets in continuous detection tasks like those used here may give rise to the P3b (Nieuwenhuis et al. 2005), which is positively correlated with activity in pericalcarine cortex (Mantini et al. 2009, supplementary material).

Functional Consequences

Although the timing and nature of the target-mediated boost are consistent with a role in context updating, there was little evidence in the present study that it interacted with the visual processing of attended, suprathreshold images. Moreover, the target-mediated boost was present in the periphery even when visuo-spatial attention was allocated to centrally presented visual stimuli. Directing attention to a central visual stimulus neither limited the boost to regions representing the stimulus nor increased the magnitude of the boost in later visual areas. Although the data demonstrate an effect of temporal selection on early visual cortical activity, they provide no clear answers regarding the functional consequences of this activity.

The recognition data from these experiments were unusual in that they showed a relatively small memory advantage for images presented at the same time as targets relative to those presented with distractors (Lin et al. 2010; Swallow and Jiang 2010). We can only speculate as to why this attentional boost effect was small in these experiments. However, an obvious difference is that the stimuli appeared at regular and predictable intervals in previous behavioral studies. In contrast, the present study used no-tone intervals to jitter the detection stimuli. The regular presentation of the detection stimuli and images in previous studies may have facilitated discrimination of the targets and distractors by making the stimuli more predictable. Unpredictable stimuli could induce a less efficient mode of attention than the rhythmic and predictable stimuli used in earlier experiments (Schroeder and Lakatos 2009).

On the surface, these data suggest that nonperceptual modulations of V1 activity may be epiphenomenal, having no effect on visual processing. The small memory effect as well as the fact that the FFA and PPA showed similar responses to targets and distractors, even for their preferred stimuli, are consistent with this possibility. However, in addition to the tenuousness of conclusions based on null effects, the conclusion that the nonperceptual V1 modulations are epiphenomenal is premature for several reasons. First, the present studies used stimuli that were ideal for examining the effects of temporal selection on activity in category-selective visual regions (the FFA and the PPA). However, because the boost was strongest in V1, its effects on perceptual processing might be strongest for visual features that are represented in V1. Second, it is possible that temporal selection has its greatest effects on visual processing when the visual input is degraded (cf. Noesselt et al. 2010). In addition, the functional consequences of the boost on perceptual processing may not be immediately observable. Indeed, it is possible that the target-mediated boost facilitates perceptual learning of visual features that coincide with goal-relevant stimuli, as in task-irrelevant perceptual learning (Seitz and Watanabe 2003).

A final possibility is that the target-mediated boost does not directly enhance perceptual processing. Rather, it could act as an entrainment signal to synchronize periodic fluctuations in the neuronal sensitivity of perceptual regions representing various visual features, modalities, and stimulus locations (Engel and Singer 2001; Lakatos et al. 2008; Schroeder and Lakatos 2009). Although the stimuli used in our experiments were presented at variable and unpredictable intervals, each instance of a behaviorally relevant stimulus could produce a signal that entrains neural processing when it occurs with sufficient regularity.

Conclusions

Spatial selection and increased attentional demands tend to increase cortical responses in regions that represent the selected stimuli, while decreasing activity in regions that do not (Luck et al. 1997; Silver et al. 2007). In contrast, the data presented here demonstrate that temporal selection of auditory and visual stimuli increases activity in early visual cortical regions. Nonvisual stimuli were found to enhance activity in early visual cortical areas when they were selected. These modulations diverge from those of spatial selection in two critical ways: they are not constrained to the spatial location or modality of the target, and they decrease, rather than increase, along the ventral visual processing stream. These differences underscore the divergent computational demands of spatial and temporal selection. They also join a growing body of evidence that suggests that current models and understanding of temporal selection need to be extended to account for its effects on context processing.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: K.S. and Y.V.J. conception and design of research; K.S., T.M., and Y.V.J. performed experiments; K.S. and Y.V.J. analyzed data; K.S., T.M., and Y.V.J. interpreted results of experiments; K.S. prepared figures; K.S. and Y.V.J. drafted manuscript; K.S., T.M., and Y.V.J. edited and revised manuscript; K.S., T.M., and Y.V.J. approved final version of manuscript.

REFERENCES

Anderson AK, Christoff K, Panitz D, de Rosa E, Gabrieli JD. Neural correlates of the automatic processing of threat facial signals. *J Neurosci* 23: 5627–5633, 2003.

Aston-Jones G, Cohen JD. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu Rev Neurosci* 28: 403–450, 2005.

Baier B, Kleinschmidt A, Müller NG. Cross-modal processing in early visual and auditory cortices depends on expected statistical relationship of multisensory information. *J Neurosci* 26: 12260–12265, 2006.

Becker MW, Pashler H, Anstis SM. The role of iconic memory in change-detection tasks. *Perception* 29: 273–286, 2000.

Bledowski C, Prvulovic D, Goebel R, Zanella FE, Linden DE. Attentional systems in target and distractor processing: a combined ERP and fMRI study. *Neuroimage* 22: 530–540, 2004.

Bouret S, Sara SJ. Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. *Trends Neurosci* 28: 574–582, 2005.

Bowman H, Wyble B. The simultaneous type, serial token model of temporal attention and working memory. *Psychol Rev* 114: 38–40, 2007.

Brefczynski JA, DeYoe EA. A physiological correlate of the "spotlight" of visual attention. *Nat Neurosci* 2: 370–374, 1999.

Breitmeyer BG, Ganz L. Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychol Rev* 83: 1–36, 1976.

Brockmole JR, Castelhano MS, Henderson JM. Contextual cuing in naturalistic scenes: global and local contexts. *J Exp Psychol Learn Mem Cogn* 32: 699–706, 2006.

Brosch M, Selezneva E, Scheich H. Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J Neurosci* 25: 6797–6806, 2005.

Bueti D, Macaluso E. Auditory temporal expectations modulate activity in visual cortex. *Neuroimage* 51: 1168–1183, 2010.

Buffalo EA, Fries P, Landman R, Liang H, Desimone R. A backward progression of attentional effects in the ventral stream. *Proc Natl Acad Sci USA* 107: 361–365, 2010.

Chun MM, Potter MC. A two-stage model for multiple target detection in rapid serial visual presentation. *J Exp Psychol Hum Percept Perform* 21: 109–127, 1995.

Chun MM, Jiang YV. Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cogn Psychol* 36: 28–71, 1998.

Cichy RM, Heinzle J, Haynes JD. Imagery and perception share cortical representations of content and location. *Cereb Cortex* 22: 372–380, 2012.

Corbetta M, Patel G, Shulman GL. The reorienting system of the human brain: from environment to theory of mind. *Neuron* 58: 306–324, 2008.

Culham JC, Cavanagh P, Kanwisher NG. Attention response functions: characterizing brain areas using fMRI activation during parametric variations of attentional load. *Neuron* 32: 737–745, 2001.

Davenport JL, Potter MC. Scene consistency in object and background perception. *Psychol Sci* 15: 559–564, 2004.

Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RT. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31: 968–980, 2006.

Desimone R, Duncan J. Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18: 193–222, 1995.

Donchin E, Coles MG. Is the P300 component a manifestation of context updating? *Behav Brain Sci* 11: 357–374, 1988.

Doty RW. Nongeniculate afferents to striate cortex in macaques. *J Comp Neurol* 218: 159–173, 1983.

Driver J, Noesselt T. Multisensory interplay reveals cross-modal influences on "sensory-specific" brain regions, neural responses, and judgments. *Neuron* 57: 11–23, 2008.

Duncan J. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn Sci* 14: 172–179, 2010.

Engel AK, Singer W. Temporal binding and the neural correlates of sensory awareness. *Trends Cogn Sci* 5: 16–25, 2001.

Engel SA, Glover GH, Wandell BA. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cereb Cortex* 7: 181–192, 1997.

Enns J, Di Lollo V. What's new in visual masking? *Trends Cogn Sci* 4: 342–352, 2003.

Falchier A, Clavagnier S, Barone P, Kennedy H. Anatomical evidence of multimodal integration in primate striate cortex. *J Neurosci* 22: 5749–5759, 2002.

Fan J, McCandliss BD, Fossella J, Flombaum JI, Posner MI. The activation of attentional networks. *Neuroimage* 26: 471–479, 2005.

Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis. II. Inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9: 195–207, 1999.

Frank MJ, Loughry B, O'Reilly RC. Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cogn Affect Behav Neurosci* 1: 137–160, 2001.

Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RS. Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* 2: 189–210, 1995.

Genovese C, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15:8 870–878, 2002.

Ghazanfar AA, Schroeder CE. Is neocortex essentially multisensory? *Trends Cogn Sci* 10: 278–285, 2006.

Gold JI, Shadlen MN. The neural basis of decision making. *Annu Rev Neurosci* 30: 535–574, 2007.

Hochstein S, Ahissar M. View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron* 36: 791–804, 2002.

Hon N, Thompson R, Sigala N, Duncan J. Evidence for long-range feedback in target detection: detection of semantic targets modulates activity in early visual areas. *Neuropsychologia* 47: 1721–1727, 2009.

Howard MA, Volkov IO, Mirsky R, Garell PC, Noh MD, Granner M, Damasio H, Steinschneider M, Reale RA, Hind JE, Brugge JF. Auditory cortex on the human posterior superior temporal gyrus. *J Comp Neurol* 416: 79–92, 2000.

Jack AI, Shulman GL, Snyder AZ, McAvoy MP, Corbetta M. Separate modulations of human V1 associated with spatial attention and task structure. *Neuron* 51: 135–147, 2006.

Jäncke L, Mirzazade S, Shah NJ. Attention modulates activity in the primary and secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. *Neurosci Lett* 266: 125–128, 1999.

Jiang Y, He S. Cortical responses to invisible faces: dissociating subsystems for facial-information processing. *Curr Biol* 16: 2023–2029, 2006.

Johnson JA, Zatorre RJ. Attention to simultaneous unrelated auditory and visual events: behavioral and neural correlates. *Cereb Cortex* 15: 1609–1620, 2005.

Johnson JA, Zatorre RJ. Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. *Neuroimage* 31: 1673–1681, 2006.

Kastner S, de Weerd P, Desimone R, Ungerleider LG. Mechanisms of directed attention in the human extrastriate cortex as revealed by function MRI. *Science* 282: 108–111, 1998.

Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE. Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320: 110–113, 2008.

Lin JY, Pype AD, Murray SO, Boynton GM. Enhanced memory for scenes presented at behaviorally relevant points in time. *PLoS Biol* 8: e1000337, 2010.

Luck SJ, Chelazzi L, Hillyard SA, Desimone R. Neural mechanisms of selective attention in areas V1, V2, and V4 in macaque visual cortex. *J Neurophysiol* 77: 24–42, 1997.

Mantini D, Corbetta M, Perrucci MG, Romani GL, Del Gratta C. Large-scale brain networks account for sustained and transient activity during target detection. *Neuroimage* 44: 265–274, 2009.

Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE, Foxe JJ. Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res Cogn Brain Res* 14: 115–128, 2002.

Naue N, Rach S, Strüber D, Huster RJ, Zaehle T, Körner U, Herrmann CS. Auditory event-related response in visual cortex modulates subsequent visual responses in humans. *J Neurosci* 31: 7729–7736, 2011.

Neisser U. *Cognition and Reality: Principles and Implications of Cognitive Psychology.* New York: Freeman, 1976.

Nieuwenhuis S, Aston-Jones G, Cohen JD. Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychol Bull* 131: 510–532, 2005.

Noesselt T, Tyll S, Boehler CN, Budinger E, Heinze HJ, Driver J. Sound-induced enhancement of low-intensity vision: multisensory influences on human sensory-specific cortices and thalamic bodies relate to perceptual enhancement of visual detection sensitivity. *J Neurosci* 30: 13609–13623, 2010.

Oliva A, Torralba A. The role of context in object recognition. *Trends Cogn Sci* 11: 520–527, 2007.

Olivers CN, Meeter M. A boost and bounce theory of temporal attention. *Psychol Rev* 115: 836–863, 2008.

O'Reilly RC, Braver TS, Cohen JD. A biologically based computational model of working memory. In: *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control,* edited by Miyake A, Shah P. Cambridge, UK: Cambridge Univ. Press, 1999, p 375–411.

Pashler H. Dual-task interference in simple tasks: data and theory. *Psychol Bull* 116: 220–244, 1994.

Pinsk MA, Doniger GM, Kastner S. Push-pull mechanism of selective attention in human extrastriate cortex. *J Neurophysiol* 92: 622–629, 2004.

Ploran EJ, Nelson SM, Velanova K, Donaldson DI, Peterson SE, Wheeler ME. Evidence accumulation and the moment of recognition: dissociating perceptual recognition processing using fMRI. *J Neurosci* 27: 11912–11924, 2007.

Polich J. Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol* 118: 2128–2148, 2007.

Raymond JE, Shapiro KL, Arnell KM. Temporary suppression of visual processing in an RSVP task: an attentional blink? *J Exp Psychol Hum Percept Perform* 18: 849–860, 1992.

Reynolds JH, Chelazzi L. Attentional modulation of visual processing. *Annu Rev Neurosci* 27: 611–647, 2004.

Reynolds JH, Heeger DJ. The normalization model of attention. *Neuron* 61: 168–185, 2009.

Rockland KS, Ojima H. Multisensory convergence in calcarine visual areas in macaque monkey. *Int J Psychophysiol* 50: 19–26, 2003.

Schira MM, Tyler CW, Breakspear M, Spehar B. The foveal confluence in human visual cortex. *J Neurosci* 29: 9050–9058, 2009.

Schroeder CE, Lakatos P. Low-frequency neural oscillations as instruments of sensory selection. *Trends Neurosci* 32: 9–18, 2009.

Schultz W, Dickinson A. Neuronal coding of prediction errors. *Annu Rev Neurosci* 23: 473–500, 2000.

Schwartz S, Vuilleumier P, Hutton C, Maravita A, Dolan RJ, Driver J. Attentional load and sensory competition in human vision: modulation of fMRI responses by load at fixation during task-irrelevant stimulation in the peripheral visual field. *Cereb Cortex* 15: 770–786, 2005.

Seitz AR, Watanabe T. Is subliminal learning really passive? *Nature* 422: 36, 2003.

Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268: 889–893, 1995.

Shinoda H, Hayhoe MM, Shrivastava A. What controls attention in natural environments? *Vision Res* 41: 3535–3545, 2001.

Shomstein S, Yantis S. Control of attention shifts between vision and audition in human cortex. *J Neurosci* 24: 10702–10706, 2004.

Shulman GL, Astafiev SV, Franke D, Pope DLW, Snyder AZ, McAvoy MP, Corbetta M. Interaction of stimulus-driven reorienting and expectation in ventral and dorsal frontoparietal and basal ganglia-cortical networks. *J Neurosci* 29: 4392–4407, 2009.

Silver MA, Ress D, Heeger DJ. Neural correlates of sustained spatial attention in human early visual cortex. *J Neurophysiol* 97: 229–237, 2007.

Sirotin YB, Das A. Anticipatory haemodynamic signals in sensory cortex not predicted by local neuronal activity. *Nature* 457: 475–479, 2009.

Slotnick SD, Thompson WL, Kosslyn SM. Visual mental imagery induces retinotopically organized activation of early visual areas. *Cereb Cortex* 15: 1570–1583, 2005.

Sokolov EN, Nezlina NI, Polyanskii VB, Evtikhin DV. The orientating reflex: the "targeting reaction" and "searchlight of attention." *Neurosci Behav Physiol* 32: 347–362, 2002.

Spence C, Driver J. On measuring selective attention to an expected sensory modality. *Percept Psychophys* 59: 389–403, 1997.

Swallow KM, Jiang YV. The attentional boost effect: transient increases in attention to one task enhance performance in a second task. *Cognition* 115: 118–132, 2010.

Swallow KM, Zacks JM, Abrams RA. Event boundaries in perception affect memory encoding and updating. *J Exp Psychol Gen* 138: 236–257, 2009.

Thiel CM, Zilles K, Fink GR. Cerebral correlates of alerting, orienting and reorienting of visuospatial attention: an event-related fMRI study. *Neuroimage* 21: 318–328, 2004.

Tootell RB, Hadjikhani N, Hall EK, Marrett S, Vanduffel W, Vaughan JT, Dale AM. The retinotopy of visual spatial attention. *Neuron* 21: 1409–1422, 1998.

Yovel G, Kanwisher N. Face perception: domain specific, not process specific. *Neuron* 44: 889–898, 2004.

Zacks JM, Speer NK, Swallow KM, Braver TS, Reynolds JR. Event perception: a mind/brain perspective. *Psychol Bull* 133: 273–293, 2007.