Activity in the Visual Cortex is Modulated by Top–Down Attention Locked to Reaction Time

Farshad Moradi[1], Constanze Hipp[2,3], and Christof Koch[1]

Abstract

■ We studied the correlation between perception and hemodynamic activity in the visual cortex in a change detection task. Whenever the observer perceived the location of a change, rightly or wrongly, the blood oxygenation level-dependent signal increased in the primary visual cortex and the nearby extrastriate areas above the baseline activity caused by the visual stimulation. This non-sensory-evoked activity was localized and corresponded to the perceived location of the change. When a change was missed, or when observers attended to a different task, the change failed to evoke such a response. The latency of the nonsensory component increased linearly with subjects' reaction time, with a slope of one, and its amplitude was independent of contrast. Control experiments are compatible with the hypothesis that the nonsensory hemodynamic signal is mediated by top-down spatial attention, linked to (but separate from) awareness of the change. ■

INTRODUCTION

A central question in sensory neuroscience is the search for neuronal populations that directly contribute to phenomenal experience (Koch, 2004; Crick & Koch, 1995). Within this context, the extent to which the early visual cortex correlates with conscious perception, with selective visual attention, a combination of both, or with something else, and whether this requires feedback from higher regions, has been vigorously pursued (Kamitani & Tong, 2005; Ress & Heeger, 2003; Lee & Blake, 2002; Tong & Engel, 2001; Polonsky, Blake, Braun, & Heeger, 2000; Ress, Backus, & Heeger, 2000; Chen et al., 1998).

Thus, the seminal experiments (Ress & Heeger, 2003; Tong & Engel, 2001; Polonsky et al., 2000) stress a connection to conscious perception and awareness on the basis of binocular rivalry or a pattern detection task. This suggests that the amplitude of V1 activity determines which stimuli reach phenomenal awareness. It is also possible, though, that the operations that underlie awareness occur at higher processing stages, such as the frontal lobes, which then feed back signals to the occipital lobe (Dehaene & Changeux, 2005; Lamme & Roelfsema, 2000). In addition, it is well documented that selective visual attention—directed by an explicit external cue—can modulate activity in the early visual cortex in a topographic manner (Ress et al., 2000; Brefczynski & DeYoe, 1999; Tootell et al., 1998), and enhances visual processing (Liu, Pestilli, & Carrasco, 2005; Ress et al., 2000). Nonetheless, it is not clear whether top-down attentional feedback is intrinsically deployed when a stimulus reaches awareness.

To investigate the nature of the correlation between activity in the early cortex and subject's behavior further, we designed a novel change detection/blindness display that allows us to manipulate stimulus parameters as well as the attentional state of the subject. In particular, our task is characterized by long and variable reaction times, and is therefore ideally suited to dissociate purely stimulus-driven from response-related modulations of the BOLD signal. We find that BOLD activity in striate and extrastriate regions can be decomposed into a large sensory and a small nonsensory component, and that the second component can be activated by spatial, top-down attention. The timing of the peak of this response correlates strongly with the subject's reaction times. Control experiments indicate that the nonsensory fMRI component originates outside the visual cortex and expresses an attentional, spatially specific, signal.

METHODS

Two authors and 20 paid and naïve volunteers from the Caltech campus, aged 20–32 years, with normal or corrected-to-normal vision, participated in one or more experiments. All experiments were conducted according to the guidelines of the institute's committee for protection of human subjects. The stimuli were generated on a PC using the Psychophysics Toolbox (Brainard, 1997) and presented to the subject via optical goggles (VisuaStim XGA; Resonance Technology, Northridge, CA; 800×600 resolution at 60 Hz; visual field = $30° \times$

[1]California Institute of Technology, [2]University of Zürich, [3]Swiss Federal Institute of Technology, Zürich

24°; maximum contrast ratio = 18:1). Responses were collected using a five-button right-hand keypad. Reaction times were measured from the onset of the change until the subject released the button. In two subjects wherein the onset of each response was also recorded, subjects released the key, on average, 200–300 msec after pressing it.
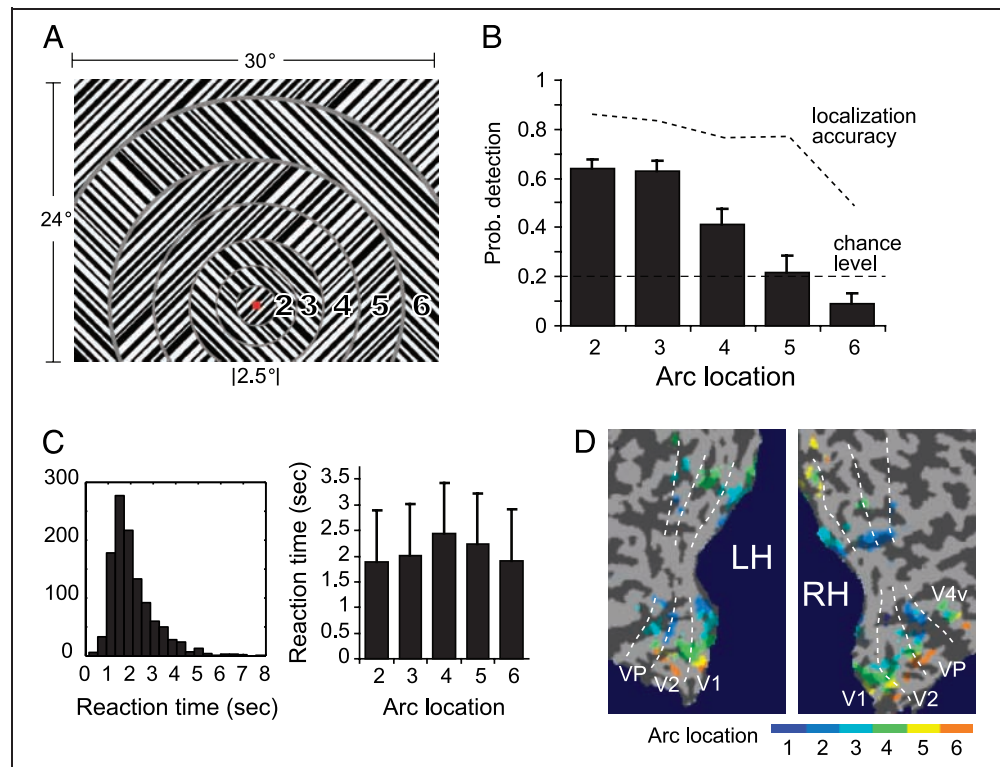
Visual Stimuli

The display comprised seven concentric ring segments (arcs), as illustrated in Figure 1A. To compensate for the cortical magnification, the width of each arc was systematically increased from center (1.4°) to the periphery (6.1°). Each arc contained alternating black and white stripes (gratings) with random widths (between 0.1° and 1°) that were tilted 45° left or right, and were separated from other arcs by a thin gray area. Every 400 msec, the stripes were replaced with a new, randomly generated set of stripes, thus the whole display appeared as flickering at 2.5 Hz. The new stripes had the same orientation as the old ones, except that every 4–8 sec, the orientation of the stripes flipped by 90° in one of the arcs. Thus, the change in orientation and local phase/

spatial frequency occurred at the same time. To avoid subjective grouping effects, we made sure that at any time during the experiment no more than two consecutive arcs have the same orientation. A red fixation mark was continuously displayed at the center of the rings, 6° below the center of the display. Observers were instructed and trained to monitor arcs 2–6 and ignore the innermost and outermost ones.

Change Detection Experiments

Nine subjects performed 8–11 change detection runs. Behavioral responses were recorded inside the scanner. Before the session, subjects were trained outside the scanner with a similar task. To familiarize observers with the task, an easy version of the experiment was used outside the scanner for training. In the main experiment, the transient associated with the change was masked by the synchronous changes in local phase and spatial frequency of other gratings. Seven participants underwent additional sessions, similar to the main experiment, but using a low-contrast (20% of the maximum contrast) display. In all other experiments, the high-contrast (100% of the maximum) gratings were used.

Figure 1. Stimulus and behavioral results. (A) A schematic of the display used in the main experiment. Observers monitored arcs 2–6 while fixating the central red square and indicated if there was any change in the orientation by pressing one of the five keys. Every 0.4 sec, each strip within the grating was varied randomly in width (range = 0.1–1°). Every 4–8 sec, the orientation of one of the gratings changed by 90°. Observers did not know when to expect the change (noncued design). The orientation in each arc flipped every 30–50 sec in random order. (B) Probability of the detection of the change as a function of the location (eccentricity) of the arc across observers. Dashed line depicts localization accuracy [Hit/(Hit + False alarm) for the location of the change]. Error bars indicate *SEM*. (C) Reaction-time histogram and reaction times as a function of the arc (mean ± *SD* across trials). (D) Voxels corresponding to each arc (based on localizer scans) are represented on the flattened cortex for one observer.

Additionally, each observer conducted three localizer runs, in which arcs were sequentially presented for 5 sec, followed by a 5-sec blank period (repeated six times). In another two to three runs, we mapped a 30° rotating wedge (9 deg/sec) of a flickering checkerboard pattern, and used the polar component of the resulting retinotopic map to identify V1–V2 borders (Engel, Glover, & Wandell, 1997; Sereno et al., 1995).

In a subset of participants, we made sure that horizontal eye movements (measured outside the scanner during training) were confined to 1° from the fixation. Additionally, two observers performed the experiment while we tracked their eye movements inside the scanner (MREye eyetracking system; Resonance Technology).

Data Acquisition

T2*-weighted echo-planar images (repetition time = 2 sec, echo time = 30 msec, flip angle = 90°, FOV = 210 mm, 31 × 3 mm interleaved axial slices) were acquired in the Caltech 3.0-Tesla Trio whole-body scanner (Siemens) using the whole-head coil and an in-line motion correction sequence (Thesen, Heid, Mueller, & Schad, 2000). The slices were positioned to cover the occipital lobe. Each functional run comprised 152 volumes (304 sec); the first seven volumes were discarded. After the initial 14 sec, a fixation point was presented on a homogenous gray screen for 20 sec, followed by the dynamic gratings whose widths changed every 400 msec (2.5 Hz). Another 20 sec of gray screen appeared at the end of the run. A 12-min high-resolution (1 × 1 × 1 mm³) T1-weighted magnetization prepared rapid gradient echo sequence (T1 = 1.5 sec, TI = 0.8 sec, TE = 3.05 msec, 160 sagittal slices) was acquired for cortical flattening.

Distracted Attention Control

We presented a stream of digits at the fovea (every 1.6 sec) and asked observers to perform a two-back task and ignore the peripheral change. Subjects had to quickly press a single button whenever the first and third one of the last three digits seen were identical. The display and data acquisition was the same as for the high-contrast change detection experiment.

Top-Down Attention Controls

Transient and sustained endogenous attention controls were conducted on naïve subjects who did not participate in the main experiment. In the transient endogenous attention control, a semantic cue (a digit between 2 and 6) was displayed every 3–9 sec (uniformly distributed) at fixation, and observers had to report the orientation of the grating inside the corresponding arc by pressing one of the two buttons. The orientation of the grating in the

cued arc was not changed for at least 2 sec after the cue. In the sustained attention control, a target arc was displayed in isolation at the beginning of each scan. In the subsequent 5 min, observers had to report orientation changes in the target arc (high-contrast display identical to the main experiment) by pressing a single key as fast as they could while ignoring changes in other arcs.

Data Acquisition

T2*-weighted echo-planar images (FOV = 192 mm, 30 × 3 mm interleaved axial slices) were acquired using a phase-array surface coil. The flickering display was presented for 5 min after the 4-sec initial scans were discarded. Other acquisition parameters were the same as for the rest of the experiment.

Analysis

Linear and low-frequency (<0.01 Hz) temporal drift was removed for each slice in the Fourier space. Student's t maps for each localizer and retinotopic-mapping run were computed separately. For each subject, images were coregistered between runs using a linear image registration tool (Jenkinson, Bannister, Brady, & Smith, 2002). Retinotopic maps were projected onto flattened cortical surface maps created in *BrainVoyager* QX (Brain Innovation B.V., Maastricht), and the V1/V2 boundaries were identified following a well-established method (Engel et al., 1997; Sereno et al., 1995). These boundaries were used in conjunction with the localizer data to define the regions of interest (ROI) for each arc. Statistical t maps for each arc were thresholded at $t < 3$ ($p < .006$, uncorrected), and voxels that were significantly activated by more than one arc were excluded. Figure 1D illustrates ROIs for one observer depicted on the flattened cortex. Our analysis is robust with respect to the exact value of the threshold for defining ROIs and we replicated all findings using a lower threshold ($t < 2.5$) as well.

The time course of the BOLD signal was calculated by summing all voxels inside each ROI after slice-timing correction and resampling at 1 Hz using cubic-spline interpolation, normalized to the sum of all voxels in the ROI for the first volume in each run. Evoked responses were averaged with respect to the onset of the change (Figures 2–4) or the response (Figure 6). The average BOLD signal 10 sec before the onset was subtracted from each trial to correct for shifts in the baseline. Trials in the first and last 30 sec were excluded. The amplitude of BOLD signal for each condition for each participant was calculated by averaging the responses 6 sec after the change occurred. Averages comprising less than 16 trials were excluded from further analysis. Statistical significance was determined by Student's t test (or analysis of variance when appropriate) on the average amplitude.

The latency of the peak (trough) was calculated after binning trials based on reaction time for the average response of each bin: The response was resampled at 10 Hz (cubic-spline) and the maximum value between 2 and 10 sec (4 and 16 sec) after change onset was identified as the peak (trough). The sensory component (response to the gratings themselves) was calculated by subtracting average activation 10 sec before the onset and 10–20 sec after the offset of the gratings (baseline) from the average BOLD signal 10–20 sec after the onset and 0–10 sec before the offset (stimulation) for each run.

RESULTS

The stimulus display (Figure 1A and supplementary movie[1]) consists of seven concentric arcs of gratings whose width changed every 400 msec (i.e., phase and spatial frequency changed locally), giving the entire display a flickering appearance. In addition, and much less frequently (every 4–8 sec), the orientation of a single grating making up an arc flipped by 90°. The change in orientation was synchronous with the changes in phase and spatial frequency in other gratings. Observers had to report the location of the grating that had changed orientation in arcs 2 to 6 and ignore everything else (Figure 1A).

Subjects easily see the orientation of any of these dynamic gratings when shown in isolation. They can also discern the orientation of gratings inside a particular arc throughout the entire display if the location is cued. However, because of crowding it is not possible to consciously and simultaneously register the orientation of all arcs. More importantly, the visual transients associated with phase and spatial frequency changes effectively mask the transient associated with the synchronous orientation change (change-blindness; Rensink, 2002; Rensink, ORegan, & Clark, 1997). Consequently, in 60 ± 3% of trials, subjects failed to notice the change in orientation (all values indicate mean ± SEM across observers, unless otherwise stated). Changes in the peripheral arcs were missed more frequently than in the more central ones (Figure 1B). When a change was noticed, the reported location was usually correct (80 ± 3%; chance level is 20%). The average reaction time was 2.13 ± 0.12 sec (Figure 1C; see Methods), consistent with spontaneous, serial deployment of attention.
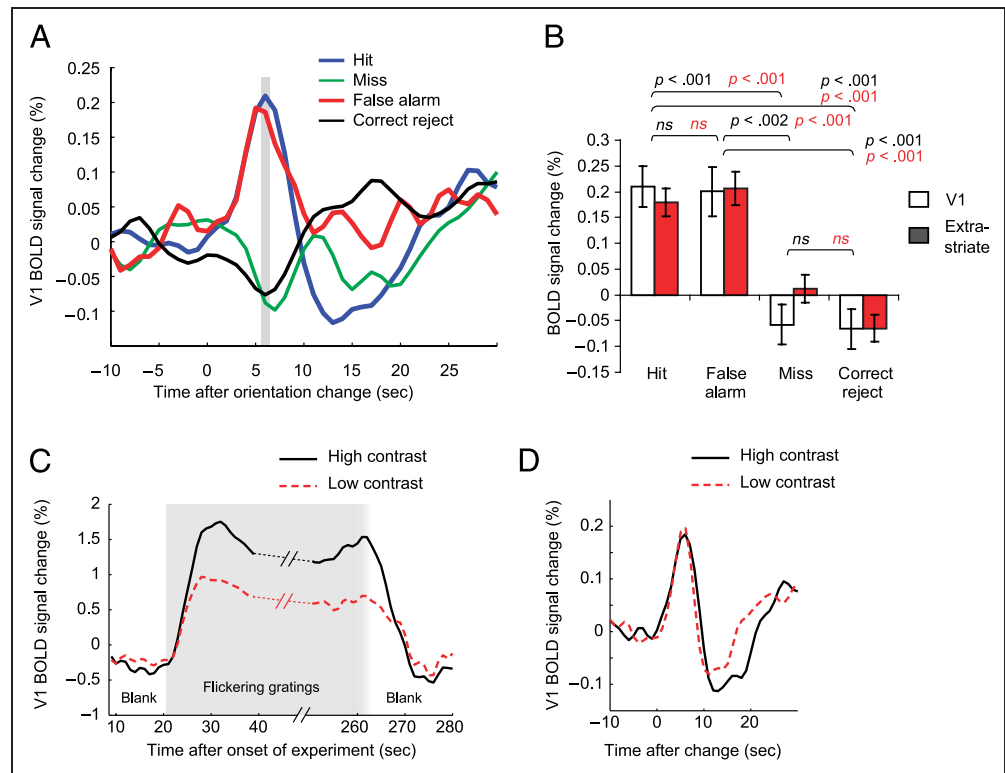
The visual areas in the occipital lobe corresponding to each arc and the borders between these retinotopic areas were identified in separate scans (Figure 1D). In the main change-blindness experiment, the BOLD signal comprised two components: a large, sustained increase (relative to the homogenous gray background) in response to the visual transients, and a much smaller transient component that followed the flip in orientation. Notably, this transient response occurred only when the change was detected by the observer [$F(3,29) = 13.9$,

$p < .00001$, Tukey–Kramer post hoc test: Hit vs. Correct reject: $p < .001$, Miss vs. Correct reject: not significant. Figure 2A and B]. A similar increase in the BOLD signal in false alarm trials suggests that the transient activity in V1 and nearby extrastriate areas (V2 = visual area 2; VP = ventral posterior area) correlates with the subjective percept rather than with the physical change in the display. When the change in the orientation was not perceived, it did not evoke this transient response. In two observers, we tracked eye position inside the scanner and verified that the correlation between BOLD signal and observers' response was not caused by eye movements. Hereafter (and for reasons that shall soon be clear), the sustained response to the changing gratings and the transient response to the perception of the change will be referred to as the sensory and nonsensory components, respectively. In a small subset of trials, observers falsely reported a change at a location that was not an immediate neighbor of the arc that changed. In these trials, the evoked BOLD signal was larger in the reported location than in the actual location of the change (paired $t_8 = 2.69$, $p < .05$).

In a variant of the main experiment, participants viewed a similar display, but the contrast of the gratings was lowered fivefold to 20%, resulting in a substantially smaller sensory-evoked component (0.94% vs. 1.73%, $t_6 = 11.6$, $p < .0001$; Figure 2C). Despite the reduced contrast, behavioral performance was similar to the high-contrast session (42% vs. 40% correct, low vs. high contrast, respectively, ns). If the peak in Figure 2A reflected modulation of V1 responses by the visual input alone (e.g., by changing the gain of lateral geniculate nucleus or layer 4 neurons), lowering the contrast should have reduced both components. On the contrary, reducing the contrast had little effect on the magnitude of the nonsensory BOLD signal (0.23% vs. 0.22%, low vs. high contrast; $t_6 = 0.31$, $p = .77$; Figure 2D), and the spatial and temporal pattern of the nonsensory component was quite similar to the nonsensory component of the high-contrast stimulus. Average contrast modulation (ratio between low- and high-contrast conditions per subject) were 1 and 0.52 for nonsensory and sensory components of the BOLD signal, respectively ($t_{12} = 3.08$, $p < .01$). Thus, the evoked V1 nonsensory component appears to be additive, not modulatory or multiplicative. A similar trend was observed in V2 (sensory component, low vs. high contrast: 0.93% vs. 1.45%, $p < .001$; nonsensory component: 0.16% vs. 0.22%, $p = .18$) and VP (sensory component: 1.06% vs. 1.56%, $p < .001$; nonsensory component: 0.16% vs. 0.19%, $p = .53$).

Remarkably, the latency of the peak nonsensory BOLD activity in striate and extrastriate areas increased with the reaction time (Figure 3A): that is, the longer the subject took to report a change, the later the peak in the BOLD signal. Sensory activation will be time-locked to the change, although it might be modulated by attention. In contrast, components of activation in the visual

Figure 2. The evoked response depends on perception. (A) Average BOLD response in V1 in correctly detected (Hit), missed (Miss), falsely localized (False alarm), and correctly rejected trials for the high-contrast display (averaged across nine observers. See Analysis in the Methods section for details). The response to the ongoing change in width (every 0.4 sec) of the gratings was discounted by subtracting the average signal in a 10-sec window before the onset of the change. Hit and false alarm traces are based on the activity in the voxels corresponding to the reported location. The miss trace corresponds to the location of the change and the correct reject trace to the average of voxels for arcs 2–6. In correct reject trials, either the first or last arc flipped, but observers were instructed not to report changes in these two locations. (B) Amplitude of the BOLD signal 6 sec after change in orientation (gray bar in A) in the four different types of trials in A. White: V1 activation, Red: V2 and VP activation. Error bars depict *SEM* across observers. The same pattern of results was observed for all arcs. (C) V1 sensory component was higher in the high-contrast display compared with the low-contrast display. (D) The nonsensory component was similar in both the high- and low-contrast displays. Note that the hemodynamic response evoked by the change in orientation adds to the much larger response evoked by the gratings whose width changed every 400 msec. C and D both depict average signal across observers.

areas that are induced by top-down feedback will be more closely related to the response. A highly significant linear relationship with a slope of one (1.03 ± 0.4) exists between reaction time and the time the BOLD response peaked. We did not find any consistent effect of subject, arc location, correctness of the response, or binning parameters on the slope. The evoked BOLD activity shifts with the reaction time without any major change
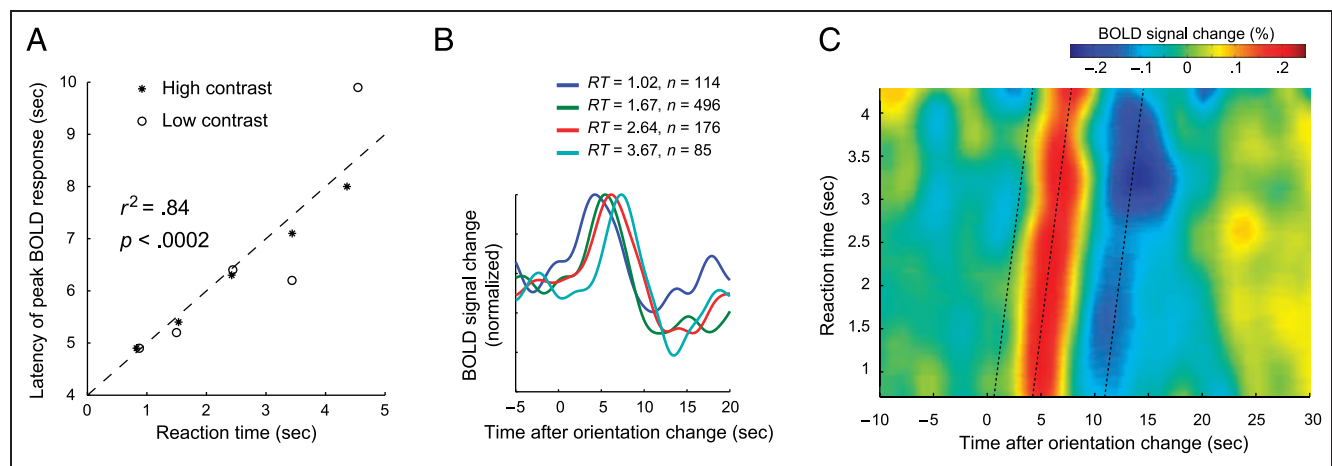
Figure 3. The latency of the V1 nonsensory component increases with reaction time. (A) Time of peak BOLD response with respect to the stimulus onset as a function of observers' reaction times (data pooled within 1-sec bins). Abscissa: average reaction time of trials in each bin; ordinate: time of peak of the average BOLD response. (B) Average V1 hemodynamic response as a function of reaction time. Responses were normalized by dividing by the peak. (C) Moving average of the BOLD activity (Gaussian window, full width at half maximum = 0.6 sec) sorted by reaction time. Note the significant dip following the sharp increase. Dashed lines with a slope of one illustrate reaction time, the positive peak, and the negative peak of response, respectively.

in the shape of the hemodynamic response function (Figure 3B and C). Absence of a nonsensory BOLD component in missed trials, and the linear dependence of the latency of the activity on the reaction time of the observer indicate that the increased BOLD signal is not driven by the change in the image. Further analysis of the average V1 BOLD signal relative to the time of the button press demonstrates that the increase in signal 4–5 sec after reporting the change is followed by a significant (but more variable) decrease for hit and false alarm trials (average trough 8 sec after the response; Hit: $-0.12 \pm 0.02\%$, $p < .01$; False alarm: $-0.11 \pm 0.04\%$, $p < .05$), time-locked to the response ($r^2 = .8$, $p < .001$). There was no indication of such a trough for miss trials.

In both change detection experiments, the nonsensory response in the correctly detected trials was spatially localized: The amplitude of the signal was highest and the latency of the peak BOLD signal in V1 was shortest at voxels that corresponded to the region of the reported arc (Figure 4). Similar results were observed in adjacent extrastriate areas V2 and VP. The amplitude of the nonsensory response was comparable, although V1 activation was slightly stronger than extrastriate activation (pooled data: 0.22% vs. 0.18%, $p < .03$).

A first control experiment, referred to as *distracted attention*, was conducted on five observers to address whether the visual activity correlates with the task or the motor response. Participants were asked to ignore the gratings, and instead, to attend and respond to a foveal stimulus. The display consisted of the same high-contrast gratings as in the main experiment, except that a stream of digits (at 0.625 Hz) was superimposed onto the display at fixation. Observers had to perform a cognitively engaging two-back memory task (see Methods).

Engaging in this attentionally demanding task at the fovea had little effect on the V1 and V2 sensory responses (Figure 5). Yet, the nonsensory BOLD signal

disappeared: We averaged the BOLD activity based on the time of the change and regardless of the response or reaction time. The amplitude of the unsorted average (6 sec after the change) in the distracted attention condition was significantly smaller than the same unsorted average in the main experiment (-0.01% vs. 0.09% BOLD increase, $t_8 = 3.62$, $p < .01$). Thus, neither the visual input nor the motor response per se can explain the nonsensory signal following hit and false alarm trials in the main experiment. Taken together with previous findings, the results demonstrate a double dissociation of sensory and nonsensory components in early visual areas: The sensory component depends on the contrast of the display but is task independent, whereas the nonsensory component is contrast independent but task dependent. This binary distinction into sensory and nonsensory components broke down in area VP. Attending to the foveal distracters and engaging in the memory task significantly reduced the BOLD response to the gratings in VP (Figure 5).

It is possible that the nonsensory BOLD component is mediated entirely by spatial attention. Ress and Heeger (2000) reported localized, stimulus-independent activity in early visual cortices following an attentional cue. In our change-blindness experiments, there is no explicit cue, and deployment of attention is spontaneous. Following the hypothesis that the nonsensory component is due to shifting spatial attention, we predict that if the location of the change is cued (either by increasing the saliency of the change or with a high-level cue), a nonsensory component should occur. We thus conducted control experiments to investigate the effect of top-down attention. Furthermore, if the nonsensory component is solely due to attentional feedback, then perceiving the change at an already attended location should not result in an additional increase in the BOLD signal. In contrast, if perception of the change evokes a localized nonsensory BOLD component at an already

Figure 4. Location specificity of the evoked BOLD response. (A) The nonsensory BOLD signal increased mainly in the region of the visual cortex that corresponded to the perceived location of the change (and its immediate neighboring arcs). Pooled high- and low-contrast data. The amplitude of the BOLD signal 6 sec after change in orientation (Hit) is compared between the reported location, its immediate neighbors, and other (distant) arcs. The activity in miss trials is shown for comparison. Error bars indicate *SEM* across sessions (seven observers carried out both high- and low-contrast sessions. Two individuals carried out the high-contrast session only). (B) Time of the peak nonsensory BOLD signal following the change (pooled data across all observers. Latencies were estimated based on bins of 100 trials sorted by reaction time). Error bars depict standard error estimates (Tukey–Kramer post hoc test).
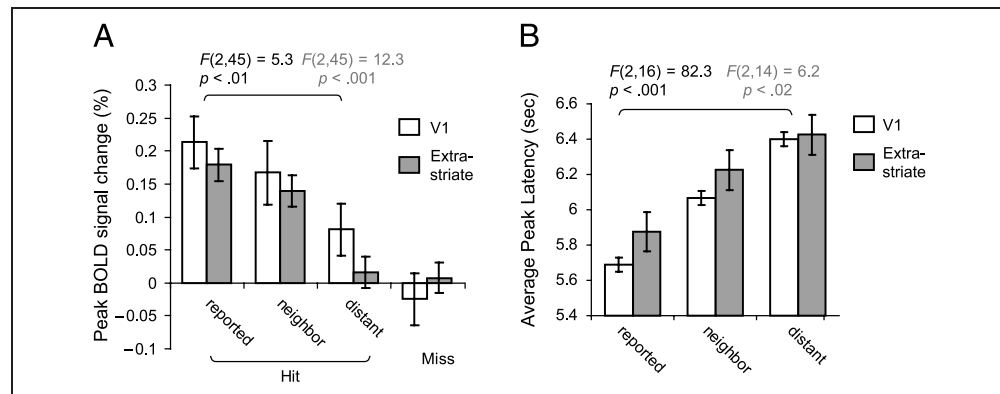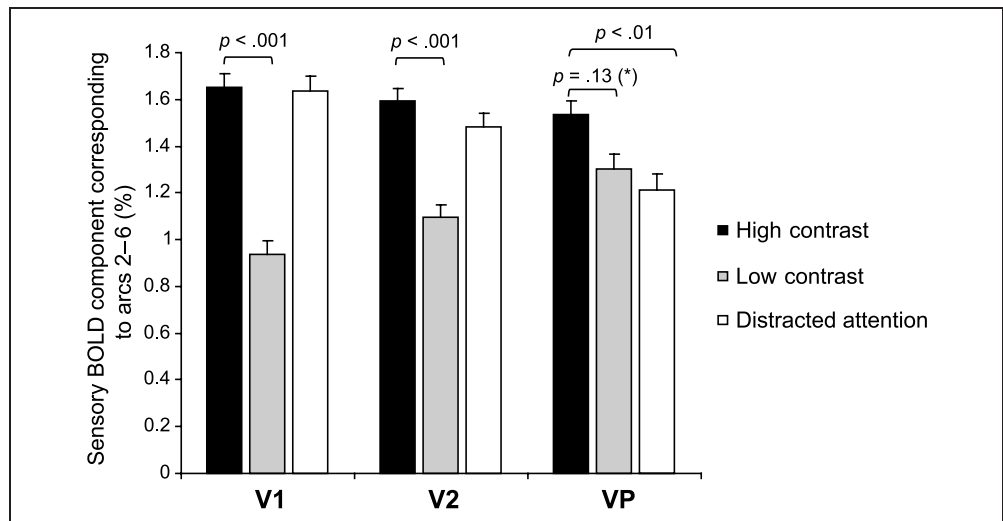
Figure 5. Modulation of sensory component by reducing contrast versus withdrawing top-down, spatial attention. The activation depicts the average BOLD signal increase compared with a blank screen (sensory component) in voxels of the visual cortex that correspond to arcs 2–6. In both high- and low-contrast change detection conditions, observers attended to the arcs. In all visual areas, activation increased with contrast, but the effect was stronger in V1 than in VP. In the distracted attention control, high-contrast grat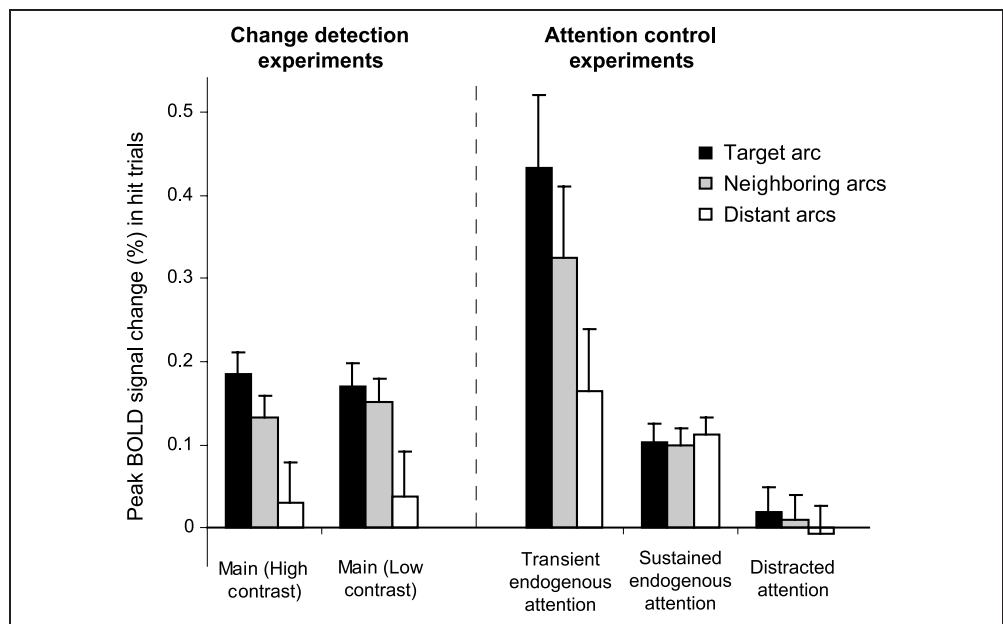ings were used, but observers were instructed to attend to a stream of digits appearing at the fovea and to carry out a two-back memory task. The difference between the high-contrast and distracted attention control, which reflects modulation by top-down and spatial attention, is significant in VP and virtually nonexistent in V1 [pooled data from four observers who participated in all three conditions, $F(8,1243) = 19.32, p < .0001$]. Error bars indicate *SEM* across observers. *The difference reaches significance if all seven observers who participated in high- and low-contrast experiments are included (data given in the text).

attended location, then mechanisms other than spatial attention need to be invoked.

In a second control experiment, the *transient endogenous attention control*, we transiently shifted spatial attention to a distinct location. A cue (arc number) presented at fixation instructed subjects where to attend; observers had to immediately report the orientation (binary choice) of the gratings inside the corresponding arc. In this experiment, changes in orientation were completely irrelevant to the task and they did not coincide with the cues at fixation. Because the cue is a high-level one, any localized signal increase would be due to the top-down shift of spatial attention. Four naïve participants judged the orientation at 91.6 ± 1.5% correct and all showed significant increases of the amplitude of the BOLD signal at the attended location (average increase = 0.4 ± 0.08%, $t_3 = 4.9, p < .02$; Figure 6). As expected, this attentional BOLD signal was

Figure 6. Comparing visual activity following hit events in change detection experiments versus attentional controls. Bars depict the peak of the average BOLD signal increase following the response in the target arc, its immediate neighbors, and distant (nonneighboring) arcs. Error bars depict *SEM* across observers. The peaks of the transient BOLD signal in transient condition were larger than the nonsensory component in the main experiments with high- and low-contrast gratings, although they all show similar location specificity. In comparison, in the sustained attention condition, the hemodynamic signal time-locked to the response is weak and not spatially localized. This transient nonspecific BOLD response cannot account for the localized nonsensory component in the main experiment. For the distracted attention condition, hit refers to correct responses in the memory task, and target refers to the arc that changed at the same time as the target digit appeared at the fovea.

localized (paired $t_3 = 9.8$, $p < .01$, target vs. distant arcs); its amplitude was almost double the nonsensory component in the main experiment.

We further studied the nature of the nonsensory hemodynamic component to orientation changes in a final *sustained endogenous attention control*. Subjects had to continuously monitor orientation changes at a single, prespecified location during the entire experiment and press a button every time a change occurred. This control is complementary to the transient endogenous attention control, as the location of spatial attention remains at one location; consequently, any observed signal changes will be related to the task-relevant changes rather than shifts of spatial attention. Five participants performed at $90.5 \pm 3.3\%$ correct (indicating that they successfully attended the target). The amplitude of the BOLD signal increased following the change in hit trials ($0.11 \pm 0.02\%$, $t_4 = 6.4$, $p < .001$; Figure 6), but not in miss trials ($0 \pm 0.03\%$, $p = .94$). However, the amplitude of the increase in hit trials was about half of the amplitude of the nonsensory component in the main experiment. Furthermore, unlike in the change detection experiment or the transient endogenous attention control, this signal was not spatially localized. That is, when observers already attended to a particular location, there was no further localized response to a reported change.

In a variant of the sustained endogenous attention control, every 4–14 sec, a high-level cue at fixation indicated the location of the subsequent change. After a variable interval (2–13.6 sec), the orientation of that arc flipped. Four observers correctly reported $87.6 \pm 1.9\%$ of the changes. The V1 BOLD signal increased significantly following the cue ($0.3 \pm .013\%$, $p < .001$). The effect of top-down attention was spatially localized [$F(2,9) = 25.9$, $p < .001$] and remained high until the change occurred. Therefore, the absence of a localized transient BOLD signal cannot be interpreted as an absence of attentional effects in V1 in the sustained attention control. A similar sustained increase occurred in miss trials ($0.32 \pm .09\%$, $p < .05$). Comparing hit and miss trials did not reveal any BOLD response to the perception of the change, confirming the hypothesis that the transient nonsensory component in the change detection experiment is caused by a shift of spatial attention to a nonattended location.

DISCUSSION

Our findings can be summarized as follows: In the main experiment, perception of the change correlated with a small but highly significant and localized increase in hemodynamic activity in the early visual cortex. This non-sensory signal component was independent of the contrast of the display and was time-locked to subjects' response. When the change was not perceived, it did not evoke any measurable hemodynamic signal. The distracted attention control demonstrated that neither the motor response per se nor the same stimulus coupled to a different task that forced subjects to attend to the center of the display induced a similar transient BOLD signal increase in the visual cortex. Top-down attentional cues induced a localized transient increase. However, when observers already attended to a task-relevant location (sustained endogenous attention control), the increase in the hemodynamic signal following the perception of the change was small and not spatially localized. This suggests that shifts of visuospatial attention—and not the awareness of the change at a particular location—underlie the nonsensory component in the change blindness experiment. This component is dissociable from the stimulus-driven activity in early—but not late—visual areas. Its amplitude ($\sim 0.22\%$) is only a fraction of the response to the dynamic gratings ($\sim 1.7\%$ for high-contrast display), and small compared with baseline fluctuations or the effect of contrast. Given its small size, it may be overlooked in studies that focus primarily on the sensory component.

It is conceivable that the long reaction times in our main experiment are due to the physical change in the orientation of the arc not being perceived until random fluctuations of the geniculate input or of intrinsic activity in V1 (Kenet, Bibitchkov, Tsodyks, Grinvald, & Arieli, 2003) temporarily increase the neuronal signal in the visual cortex, thereby enhancing the chance that this larger signal is detected in higher regions and made accessible to conscious perception and motor control. This feed-forward account is compatible with the observed linear relationship between the timing of the peak BOLD activity and reaction time. It does, however, also predict that the same response (on average) should follow the change regardless of the task, whereas no nonsensory response occurred in our distracted attention control. This finding is unlike that reported by Muckli, Kohler, Kriegeskorte, and Singer (2005), studying V1 activity without stimulation of the corresponding retinal location in apparent motion. Their signal was not reduced when attention was distracted. Thus, Muckli et al. (2005) did not rule out that such activity is driven by the sensory input (via horizontal connections from stimulated V1 neurons) and is not directly and immediately linked to the percept. Similarly, Beck and Kastner (2005) recently reported a popout effect in the early visual cortex that did not require top-down attention. In contrast, absence of the nonsensory signal in our distracted attention control indicates that sensory input by itself is insufficient to drive this component.

This nonsensory component could be akin to the attentional increase in visual activity in the absence of an image (Kastner, Pinsk, De Weerd, Desimone, & Ungerleider, 1999), or unrelated to the visual stimulus (Ress et al., 2000)—except that in our experiments the activity does not follow an explicit external cue. Thus,

the same top-down signal in V1 is observed whether the input has high, low, or even zero contrast (Kastner et al., 1999), and whether or not an explicit attentional cue is present (main vs. transient endogenous attention experiments). Our results extend these studies by demonstrating a link between nonsensory attentional component and perception that is not mediated by external cues.

Ress and Heeger (2003) used near-threshold stimuli and found V1 activation following the stimulus when the stimulus was seen, and no activation when it was missed. Their result at the surface appears to be the same as our correlation between V1 activity and change detection. However, their paradigm was designed to engage attention continuously and to minimize attentional shifts; that is, it was designed to insure sustained attention. Under those conditions, the activation for seen targets was retinotopically localized (Ress & Heeger, 2003). In our sustained attention control, the transient signal increase was not localized. Therefore, in contrast to Ress and Heeger, we interpret our control experiments as indicating that the localized nonsensory activity in V1 and V2 is mediated by top-down attentional mechanisms. Remarkably, the nonsensory component in the change detection experiment correlates with perception of a change that could precede deployment of attention by a few seconds. Therefore, the link between perception and V1 activity cannot be attributed to a facilitation of the response to the change in hit trials due to the trial-to-trial variability in the deployment of attention.

In the transient endogenous attention control, the nonsensory signal that followed the cue was robust and localized, indicating that top-down activation can evoke a BOLD response similar to the one in the main experiments. Indeed, it was about twice the size. The nonsensory component in the main experiment could be smaller because a change may occur at an already attended location, with no need for a shift of spatial attention. Top-down feedback, linked to spatial attention and possibly originating in parietal or frontal sites, is the most parsimonious account explaining the correlation between reporting the change and the nonsensory response in all conditions that we examined, including the reaction times data.

Our change detection task reveals concurrent and independent sensory and nonsensory signals in the early visual cortex. The dominant, sensory component depends on contrast and is largely unaffected by attention, whereas the nonsensory, attentional component is smaller, independent of contrast, depends on the task, and reflects whether the orientation change is reported. The nonsensory BOLD component could mediate different, nonexclusive functions. It could act as a ''read-out'' signal (Kamitani & Tong, 2005), permitting conscious access to the neural information in V1 (Ress et al., 2000). Or, it may relate to the precise localization of the change or it may be necessary to initiate long-term plasticity and perceptual learning.

Previous studies of visual change blindness have revealed a network of parieto-frontal regions associated with visual attention (Beck, Muggleton, Walsh, & Lavie, 2005; Pessoa & Ungerleider, 2004; Beck, Rees, Frith, & Lavie, 2001; Huettel, Guzeldere, & McCarthy, 2001), correlated with detection of the change. Those findings provide indirect evidence of involvement of attention in change detection. None of the above studies carried out a detailed analysis of attentional enhancement of early visual activity as we have done here. Thus, our results directly demonstrate for the first time that shift and deployment of attention in V1 and nearby visual areas is tightly linked to perceiving and reporting the change among several items.

The linear relationship between the timing of the fMRI peak activity and reaction time suggests that attention shifts around the time that the subjects committed to a response. If attention to the change itself determined how fast a change was registered, then we would have expected a correlation between the amplitude of the BOLD activity—rather than its latency—with reaction time (i.e., a weaker BOLD activity for longer reaction time than for shorter ones). Alternatively, the activity in V1 could have increased from the time of the change and saturated until the time the observer responded. This was the case in one of the attention controls, but not in the change detection experiment. Another possibility is that trials with short reaction time correlate with the activity in visual areas, but as time passes, this correlation disappears due to the involvement of other areas and processes. Other studies have reported a correlation between timing of brain activity and perception in parietal and frontal areas, but not in primary sensory areas (Formisano et al., 2002; Menon, Luknowsky, & Gati, 1998). Such a lack of a meaningful relationship between the timing of peak V1 BOLD activity and reaction time could have suggested that change is represented entirely outside the primary visual cortex.

The nature of our display insures considerable variability in reaction times, with subjects reporting a change occasionally 3 sec or more after it occurred. Our interpretation of the nonsensory BOLD component as reflecting top-down attention rather than change awareness per se leaves open the fundamental question of the causal relationship between attention and awareness, and the interesting possibility that the change is perceived much sooner but is not reported until later, due to a variety of confounding factors (attention, sensory–motor mapping, visual memory etc.). That is, the nonsensory BOLD component represents an attentional feedback after subjects have perceived the change. If true, it would provide further evidence in favor of a dissociation between visual attention and visual awareness (Dehaene & Changeux, 2005; Koch & Tsuchiya, in press; Lamme, 2003).

Acknowledgments

Note

1. Movie: An example of the display used in the main experiment. Observers were instructed to fixate at the red square while monitoring any changes in the orientation of rings 2–6. The change in the orientation is perceived only by scrutinizing the display. In this example, the orientation of rings 1 (the central disk), 6, and 3 changes sequentially (www.klab.caltech.edu/~farshadm/sfn2004/demo.avi, www.klab.caltech.edu/~farshadm/sfn2004/demo.mov).

REFERENCES

Beck, D. M., & Kastner, S. (2005). Stimulus context modulates competition in human extrastriate cortex. *Nature Neuroscience, 8,* 1110–1116.

Beck, D. M., Muggleton, N., Walsh, V., & Lavie, N. (2005). Right parietal cortex plays a critical role in change blindness. *Cerebral Cortex*.

Beck, D. M., Rees, G., Frith, C. D., & Lavie, N. (2001). Neural correlates of change detection and change blindness. *Nature Neuroscience, 4,* 645–650.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the "spotlight" of visual attention. *Nature Neuroscience, 2,* 370–374.

Chen, W., Kato, T., Zhu, X. H., Ogawa, S., Tank, D. W., & Ugurbil, K. (1998). Human primary visual cortex and lateral geniculate nucleus activation during visual imagery. *NeuroReport, 9,* 3669–3674.

Crick, F., & Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature, 375,* 121–123.

Dehaene, S., & Changeux, J. P. (2005). Ongoing spontaneous activity controls access to consciousness: A neuronal model for inattentional blindness. *PLoS Biology, 3,* e141.

Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex, 7,* 181–192.

Formisano, E., Linden, D. E., Di Salle, F., Trojano, L., Esposito, F., Sack, A. T., et al. (2002). Tracking the mind's image in the brain: I. Time-resolved fMRI during visuospatial mental imagery. *Neuron, 35,* 185–194.

Huettel, S. A., Guzeldere, G., & McCarthy, G. (2001). Dissociating the neural mechanisms of visual attention in change detection using functional MRI. *Journal of Cognitive Neuroscience, 13,* 1006–1018.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage, 17,* 825–841.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience, 8,* 679–685.

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron, 22,* 751–761.

Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A., & Arieli, A. (2003). Spontaneously emerging cortical representations of visual attributes. *Nature, 425,* 954–956.

Koch, C. (2004). *The quest for consciousness: A neurobiological approach*. Englewood, CO: Roberts & Company Publishers.

Koch, C., & Tsuchiya, N. (in press). Attention and consciousness: Two distinct brain processes. *Trends in Cognitive Sciences.*

Lamme, V. A. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences, 7,* 12–18.

Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences, 23,* 571–579.

Lee, S. H., & Blake, R. (2002). V1 activity is reduced during binocular rivalry. *Journal of Vision, 2,* 618–626.

Liu, T., Pestilli, F., & Carrasco, M. (2005). Transient attention enhances perceptual performance and fMRI response in human visual cortex. *Neuron, 45,* 469–477.

Menon, R. S., Luknowsky, D. C., & Gati, J. S. (1998). Mental chronometry using latency-resolved functional MRI. *Proceedings of the National Academy of Sciences, U.S.A., 95,* 10902–10907.

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *PLoS Biology, 3,* e265.

Pessoa, L., & Ungerleider, L. G. (2004). Neural correlates of change detection and change blindness in a working memory task. *Cerebral Cortex, 14,* 511–520.

Polonsky, A., Blake, R., Braun, J., & Heeger, D. J. (2000). Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nature Neuroscience, 3,* 1153–1159.

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology, 53,* 245–277.

Rensink, R. A., ORegan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8,* 368–373.

Ress, D., Backus, B. T., & Heeger, D. J. (2000). Activity in primary visual cortex predicts performance in a visual detection task. *Nature Neuroscience, 3,* 940–945.

Ress, D., & Heeger, D. J. (2003). Neuronal correlates of perception in early visual cortex. *Nature Neuroscience, 6,* 414–420.

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science, 268,* 889–893.

Thesen, S., Heid, O., Mueller, E., & Schad, L. R. (2000). Prospective acquisition correction for head motion with image-based tracking for real-time fMRI. *Magnetic Resonance in Medicine, 44,* 457–465.

Tong, F., & Engel, S. A. (2001). Interocular rivalry revealed in the human cortical blind-spot representation. *Nature, 411,* 195–199.

Tootell, R. B., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., et al. (1998). The retinotopy of visual spatial attention. *Neuron, 21,* 1409–1422.