**You might find this additional information useful...**

Supplemental material for this article can be found at:

http://jn.physiology.org/cgi/content/full/00441.2007/DC1

This article cites 63 articles, 31 of which you can access free at:

http://jn.physiology.org/cgi/content/full/98/6/3436#BIBL

Updated information and services including high-resolution figures, can be found at:

http://jn.physiology.org/cgi/content/full/98/6/3436

Additional material and information about *Journal of Neurophysiology* can be found at:

http://www.the-aps.org/publications/jn

This information is current as of February 12, 2008 .

# Dynamics of Response to Perceptual Pop-Out Stimuli in Macaque V1

**Matthew A. Smith,[1] Ryan C. Kelly,[1,2] and Tai Sing Lee[1,2]**
[1]*Center for the Neural Basis of Cognition and* [2]*Computer Science Department, Carnegie Mellon University, Pittsburgh, Pennsylvania*

**Smith MA, Kelly RC, Lee TS.** Dynamics of response to perceptual pop-out stimuli in macaque V1. *J Neurophysiol* 98: 3436–3449, 2007. First published September 26, 2007; doi:10.1152/jn.00441.2007. Contextual modulation due to feature contrast between the receptive field and surrounding region has been reported for numerous stimuli in primary visual cortex. One type of this modulation, iso-orientation surround suppression, has been studied extensively. The degree to which surround suppression is related to other forms of contextual modulation remains unknown. We used shape-from-shading stimuli in a field of distractors to test the latency and magnitude of contextual modulation to a stimulus that cannot be distinguished with an orientation-selective mechanism. This stimulus configuration readily elicits perceptual pop-out in human observers and induces a long-latency contextual modulation response in neurons in macaque early visual cortex. We found that animals trained to detect the location of a pop-out stimulus were better at finding a sphere that appeared to be lit from below in the presence of distractors that were lit from above. Furthermore, neuronal responses were stronger and had shorter latency in the condition where behavioral performance was best. This asymmetry is compatible with earlier psychophysical findings in human observers. In the population of V1 neurons, the latency of the contextual modulation response is 145 ms on average (ranging from 70 to 230 ms). This is much longer than the latency for iso-orientation surround suppression, indicating that the underlying circuitry is distinct. Our results support the idea that a feature-specific feedback signal generates the pop-out responses we observe and suggest that V1 neurons actively participate in the computation of perceptual salience.

## INTRODUCTION

Neurons in primary visual cortex (V1) respond to stimuli presented in a defined region of space, known as the classical receptive field (CRF). Stimuli placed outside this region do not elicit a response on their own, yet can profoundly modulate the CRF-driven response (for a review see Albright and Stoner 2002). This phenomenon of contextual modulation from the area outside the CRF has been studied with many different forms of feature contrast–oriented line segments (Gilbert and Wiesel 1990; Knierim and Van Essen 1992; Nothdurft et al. 1999), sinusoidal gratings (Bair et al. 2003; Cavanaugh et al. 2002; Marcus and Van Essen 2002; Sceniak et al. 1999; Smith et al. 2006), texture (Kastner et al. 1997; Lamme 1995; Lee et al. 1998; Rossi et al. 2001), binocular disparity (Zipser et al. 1996), color or luminance (Zipser et al. 1996), and shape-from-shading (Lee et al. 2002). The earliest studied of these types of modulation, surround suppression to extended iso-orientation contours, was originally thought to be mediated by horizontal axonal projections within V1 (DeAngelis et al. 1994; Gilbert and Wiesel 1983; Knierim and Van Essen 1992; Nelson and

Frost 1978). However, studies of other forms of contextual modulation have favored feedback from higher cortical areas as an explanation, partially based on the delay of the modulation relative to the initial response (Knierim and Van Essen 1992; Lamme 1995; Lee et al. 1998, 2002; Zipser et al. 1996).

The latency of contextual modulation in V1 has been observed to occur as soon as 8 ms after the CRF-driven response in the case of surround suppression (Bair et al. 2003), but can also be delayed by $\geq$60 ms (Lee and Nguyen 2001; Lee et al. 2002; Roelfsema et al. 2003; Rossi et al. 2001), with many intermediate values reported (for a review see Alexander and Wright 2006). Typically, orientation-based contextual effects occur with delays in the shorter end of this range (Knierim and Van Essen 1992; Lamme 1995; Marcus and Van Essen 2002; Nothdurft et al. 1999; Rossi et al. 2001), whereas modulation to more complex stimuli occurs with longer delays (Lee and Nguyen 2001; Lee et al. 2002; Roelfsema et al. 2003). This difference in latency might lead us to expect that surround suppression underlies the former but not the latter. However, feedback from extrastriate cortex has been proposed as the underlying mechanism for both surround suppression (Angelucci et al. 2002; Bair et al. 2003; Cavanaugh et al. 2002; Levitt and Lund 2002) and other forms of contextual modulation (Lamme 1995; Lee et al. 1998, 2002; Zipser et al. 1996). Furthermore, most stimuli used to measure contextual modulation have some orientation structure, which makes it difficult to avoid involving some measure of surround suppression. Finally, nearly all studies of contextual modulation have used a population average response to measure latency, which can be insensitive to the response dynamics of individual neurons. A more thorough examination of the magnitude and latency of contextual modulation is necessary to determine the extent to which it is related to surround suppression.

Lee et al. (2002) used a shape-from-shading stimulus that is well suited to such a study. It consisted of an array of shaded Lambertian spheres, in which one of the spheres (the oddball) was lit from a direction opposite to all the others (Fig. 1, *C* and *D*). Alternatively, assuming a single lighting source from above, observers tend to perceive the spheres as convex or concave. This stimulus readily induces *perceptual pop-out* (Bergen and Julesz 1983a; Treisman and Gelade 1980), a phenomenon in which an oddball stimulus is easily and automatically identified among a field of identical distractors. It also evokes a long-latency contextual modulation response in V1 and V2 neurons, which can be modified by experience (Lee et al. 2002). An orientation-selective and phase-insensitive mechanism, such as surround suppression (DeAngelis et al. 1994; Webb et al. 2005), cannot distinguish the shading direc-

---

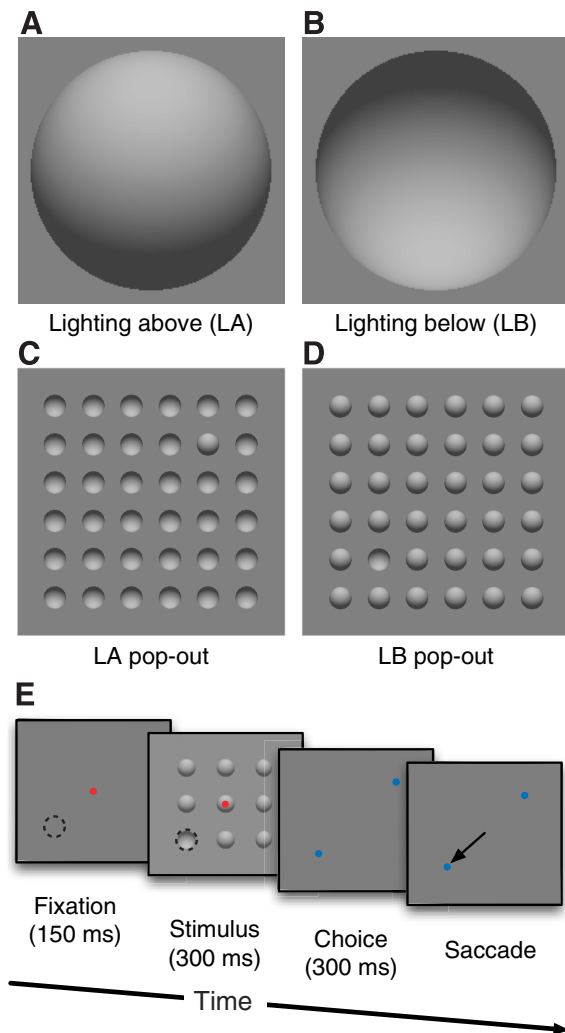0022-3077/07 $8.00 Copyright © 2007 The American Physiological Society          www.jn.org

FIG. 1. Examples of the stimulus and the task. Our stimulus consisted of an array of Lambertian spheres with lighting from above (LA, A) or from below (LB, B), which appear to be convex and concave, respectively. C: when a sphere that is lit from above appears among an array of spheres lit from below, the LA sphere readily pops out (LA pop-out). D: similarly, a sphere lit from below readily pops out among an array of spheres lit from above (LB pop-out). E: task was a 2-alternative forced choice (2AFC) design in which the animal was required to make a saccade to the pop-out stimulus to receive a reward. Receptive field is shown with a dashed circle. Pop-out stimulus could appear in one of 2 locations, either in the receptive field or in the opposite quadrant. In this case the pop-out stimulus appeared in the receptive field and a saccade to the lower left target was rewarded.

nomenon of perceptual pop-out, then we might expect them to display a comparable asymmetry.

Here we report the results of behavioral and physiological experiments in macaque monkeys performing a pop-out or oddball detection task based on shape-from-shading stimuli. Animals were better able to detect the pop-out of spheres lit from below, and V1 neuronal responses to such stimuli were stronger and occur sooner than those to the pop-out of spheres lit from above. Furthermore, contextual responses to both types of pop-out stimuli occur with a long delay after the stimulus onset, much longer than that reported for iso-orientation surround suppression.

## METHODS

### Animals

We examined the responses of single units in V1 of one male and one female Rhesus macaque monkeys (*Macaca mulatta*). Under general anesthesia, each animal was implanted with a head fixation post, scleral search coils (Robinson 1963), and a recording chamber (Crist et al. 1988) overlying the operculum of area V1. All procedures were approved by the Institutional Animal Care and Use Committee of Carnegie Mellon University and were in compliance with the guidelines set forth in the United States Public Health Service Guide for the Care and Use of Laboratory Animals.

### Electrophysiology

We inserted electrodes through the dura into the operculum of V1 using a hydraulic microdrive (Narishige, Tokyo, Japan). After we passed through the dura, we waited for 30 min before searching for isolated units. We recorded extracellular activity from single units with tungsten electrodes coated with epoxy (FHC, Bowdoinham, ME). The impedance of these electrodes after passing through the dura was typically 300–800 kOhm. Signals from the electrode were amplified (FHC, Bowdoinham, ME) and band-pass filtered (typically 500 Hz to 10 kHz). Action potentials were detected on the basis of height using a window discriminator, and the time of each spike was stored with 1-ms resolution. Once we isolated a single unit, we mapped its receptive field using the minimum response field technique with a small oriented bar. We used the center position of the receptive field (RF) for all subsequent stimuli (a sphere was centered on the RF in each condition). Receptive field positions ranged from 2 to 5° eccentricity in the lower visual field, and diameter ranged from 0.5 to 1.0° (determined by hand-mapping with a bar stimulus while the animal fixated). We measured eye position with a scleral search coil system (Riverbend Instruments, Birmingham, AL) and sampled at 200 Hz during the experimental sessions. Animals were required to maintain eye position within 0.8° of the central point during fixation. In a previous study we found that animals typically maintained fixation within a window less than half this diameter, and there was no difference in the animals' fixation pattern between LA ("lighting above") and LB ("lighting below") stimuli (Lee et al. 2002).

### Pop-out detection task

The animals performed a two-alternative forced choice (2AFC) pop-out–detection task. Each animal was required to fixate a 0.15° red dot for 150 ms to begin the trial. The stimulus array appeared for 300 ms (or 500 ms for some trials in one of the animals). After the last video frame containing the stimulus and fixation spot, the two target dots appeared on the following frame. They were located roughly equidistant from fixation, one in the receptive field and one in the opposite quadrant. The animal was required to leave the central fixation window within 300 ms, perform a saccade to one of the two

tion of the spheres. Furthermore, the perceptual pop-out of shape-from-shading stimuli depends on the three-dimensional (3D) interpretation of the scene, such as a convex object popping out from a field of concavities (Ramachandran 1988), which is likely accomplished by extrastriate cortex. In addition, human psychophysical studies with shape-from-shading pop-out stimuli have revealed an asymmetry in pop-out detection— objects shaded to appear lit from below in a field of objects lit from above pop out better than the opposite configuration (Sun and Perona 1996a,b, 1998). Because our visual system tends to assume lighting from above in ambiguous scenes (Ramachandran 1988), one explanation is that unfamiliar objects pop out more strongly than those with a familiar or simple interpretation (Sun and Perona 1996b; Treisman 1985). If contextual modulation responses in V1 neurons are related to the phe-

target dots within 100 ms, and maintain fixation within a 1.0° window around that target for 250 ms. If the animal failed to meet any of the preceding conditions within the given time allotment, the trial ended and there was a timeout of 1 s before the next trial could begin. On correct trials, ones in which the animal successfully met the preceding conditions and reached the target dot at the location of the pop-out stimulus, the animal received a liquid reward and an audible beep. On incorrect trials no reward was given. The two types of pop-out stimuli (LA and LB) were shown at five contrasts at each of the two target locations. The two stimuli and two target locations are shown in Supplemental Fig. 1.[1] This resulted in 20 total conditions in each block, which was randomized and repeated 10–50 times. In all of the experiments described herein as well as during initial training, the LA and LB stimuli were presented in equal proportion.

### Visual stimulus generation

Stimuli were presented on a Dell Ultrascan monitor operating at a resolution of 1,024 × 768 pixels and a video frame rate of 100 Hz. The monitor was positioned 58 cm from the animal and subtended 34 × 27°. We used look-up tables to correct for nonlinearities in the relation between input voltage and phosphor luminance in the monitors. The mean luminance of the display was about 32 cd/m$^2$. The maximum luminance was 65 cd/m$^2$ and the minimum was near 0.0 cd/m$^2$. The stimulus consisted of an array of shaded Lambertian spheres with lighting from above (LA) or from below (LB). This array extended across the entire monitor (typically 24 spheres horizontally and 18 spheres vertically), with one of the spheres in the array shaded with lighting in the opposite direction of the others (the "oddball"). Each shaded sphere in the stimulus display was 1° in diameter and the center-to-center spacing was 1.6°. The stimulus was created by making an array of spheres (larger than the video display) that was shifted so that one was precisely centered on the RF of the neuron under study. The sphere locations were thus set based on the RF location and only the direction of shading was changed between trials. On each trial, the oddball sphere either appeared in the RF, or in the opposite quadrant sphere location that was nearest to equally distant from fixation. This ensured that the stimulus array had the same precise sphere positions in all conditions, and thus avoided the possibility that the subject could use cues from the edge of the video monitor to determine the correct choice. Before inclusion in this study, each neuron was tested to ensure it did not respond to stimulation at the location of the adjacent spheres in the array (a receptive field >2.2° would be required to respond to the adjacent spheres). The Michelson contrast of the spheres was adjusted by scaling the luminance values about the mean (3, 6, 12, 25, or 50%). Stimulus presentation and data collection were under the control of an Intel x86–based host computer running NIMH Cortex software.

### Choice probability

We calculated the trial-to-trial correlation between neuronal firing and behavioral choice using the "choice probability" (CP) metric (Britten et al. 1996). We converted all the firing rates at each contrast into z-scores by subtracting the mean response and dividing by the SD at that contrast. We then combined the data from the 3 and 6% contrast trials, and divided it into two groups based on whether the animal made a saccade into or away from the RF of the recorded neuron. We included data only for which the animal made at least five saccades in each direction (this eliminated only one neuron). From the two firing rate distributions, we calculated a receiver-operating characteristic (ROC) curve. The CP was computed by integrating the area under this ROC curve. The chance value for CP is 0.5 if there is no correlation between the neuronal firing and behavioral choices. For each neuron, we tested the significance of its CP using a permutation

test (Britten et al. 1996; Nienborg and Cumming 2006; Uka and DeAngelis 2003). For 1,000 permutations, we randomly assigned firing rates to the two choices (while maintaining the distributions of both firing rates and choices). We compared the distribution of these permuted CPs with our original CP measurement. We counted each CP as significant only if it fell outside the 95% confidence interval of the distribution of permuted CPs.

### Response modulation

For each neuron, we presented four types of trials: LA or LB pop-out with the oddball stimulus either in the receptive field or in the diagonally opposite quadrant (see Supplemental Fig. 1). The oddball response was determined from the trials in which the oddball stimulus was in the receptive field. For LA pop-out, if the receptive field was up and to the right, the stimulus would look like the one shown in Fig. 1C. The uniform or baseline response was determined from the trials in which the other type of oddball was located away from the receptive field. Using the same example with the receptive field up and to the right, the LA uniform response would be measured with a stimulus that looks like the one in Fig. 1D. The LA pop-out response was then taken as the difference between the response to the stimuli in Fig. 1, C and D. In all trials, an oddball stimulus was present in the display. The use of an oddball stimulus on every trial, rather than a separate uniform stimulus, was necessary to maximize the number of trials from which to compute the response latency. In a previous study (Lee et al. 2002), we used a completely uniform array of spheres for the uniform condition and found response modulation to the pop-out stimulus was similar to our present results. For all of the work described herein, we restricted our analysis only to trials on which the animal made a correct response. However, we found no significant difference in the neuronal response on correct and error trials.

Based on our measurements of response latency and magnitude, we found that the time period from 100 to 300 ms after stimulus onset contained most of the response modulation due to the pop-out stimulus. We therefore restricted our analysis only to that time period. We determined the spike rate computed within the 200-ms window across all of the trials for each of the stimulus conditions and computed a modulation ratio for LA and LB pop-outs. This index was calculated by taking the difference between the pop-out and uniform condition divided by their sum, as shown here

$$Modulation \; ratio = \frac{(Response_{oddball} - Response_{uniform})}{(Response_{oddball} + Response_{uniform})}$$

$Response_{oddball}$ is measured in the time window from 100 to 300 ms after the onset of the stimulus in which the oddball is centered on the receptive field. $Response_{uniform}$ is measured in the same time window for the condition in which the shading of the sphere in the receptive field matches its neighbors (the oddball is located in the opposite hemifield). Thus the receptive field stimuli in the two conditions are identical. The only difference between the two conditions is >1° from the RF center, larger than our measurements of RF size (0.5–1.0°).

### Determination of latency

We counted the number of spikes in response to the stimulus on each trial in a 10-ms window. We then moved this window in 1-ms increments across the entire stimulus period (stimulus onset to 300 ms), and performed a Wilcoxon sign-rank test on the list of spike counts (baseline response subtracted). The baseline was calculated as the average response across all conditions in the 100-ms period preceding stimulus onset (while the animal fixated). Using a criterion of $P < 0.01$, we looked for the earliest time at which the response exceeded baseline in 10 consecutive windows. The first of these consecutive windows was defined as the onset latency.

---

[1] The online version of this article contains supplemental data.

After using this automated procedure, we performed a bootstrap analysis (Efron and Tibshirani 1993) to test the significance of our latency measurements. From the set of correct trials, we chose an equal number of bootstrap trials randomly with replacement. We then used the same method described earlier to determine the latency for that bootstrap sample. This procedure was repeated 1,000 times to generate a bootstrap distribution of latencies. We compared this distribution to our original latency measurement (determined using all of the trials). We kept this original latency only if it fell within the 95% confidence interval of the bootstrap distribution, and if the bootstrap distribution had SD <25 ms and SE <5 ms. Using this procedure, we were able to determine the onset latency of 89 cells to the LA stimulus and 83 cells to the LB stimulus (out of 114 total neurons).

In determining the latency of the pop-out response, we used the same method with a window size of 20 ms starting at response onset. In this case we used the Wilcoxon rank-sum test to determine whether the number of spikes in the window on each trial in the pop-out condition was significantly different from the number of spikes in the uniform condition. We performed the same bootstrap analysis, this time keeping only latencies that fell within the 95% confidence interval of the bootstrap distribution, and for which the distribution had SD <50 ms and SE <10 ms. With this method, we were able to determine the modulation latency of 37 cells to the LA stimulus and 38 cells to the LB stimulus. The difficulty in determining latency for the onset and modulation is likely due to the relatively small magnitude of the contextual response, the stringency of the requirements of our bootstrap test, and the fact that we recorded from all visually responsive neurons without excluding those that did not respond well to the pop-out stimuli.

We considered two ways in which our latency calculation might be biased. First, we used the onset latency as our start time when using the sliding window to determine the modulation latency. It is possible that this led to neurons with a late onset time having longer modulation latencies. However, the distribution of onset latencies and modulation latencies overlapped by only a few neurons. An inspection of these few revealed no early modulation effects missed due to the later start window. Second, our method of determining latencies was based on finding statistically significant deviations in firing rate. Using this kind of method, neurons with smaller modulation effects might tend to have artificially longer latencies due to an increased time to reach statistical significance. For this reason, we analyzed the relationship between modulation ratio and latency. There was no significant correlation between the two (Pearson's $r = -0.06$, $n = 94$, $P = 0.56$).

## RESULTS

We recorded single-unit neuronal activity in area V1 of two macaque monkeys while measuring their behavioral performance in a pop-out detection task. The task involved detection and saccade to an oddball target in a array of distractors, using the same shape-from-shading stimuli as in a previous study (Lee et al. 2002). Each object in the array is generated by shading a Lambertian sphere with lighting from above (LA, Fig. 1A) or from below (LB, Fig. 1B). If the observer assumes a single lighting source from above, the spheres are perceived as convex or concave. The array then consists of a group of spheres lit from below with an oddball lit from above (LA pop-out, Fig. 1C) or the reverse (LB pop-out, Fig. 1D).

### Psychometric performance function

Two monkeys performed a two-alternative forced choice (2AFC) pop-out detection task, shown in Fig. 1E. After the animal achieved fixation, a field of spheres appeared on the display. The field was aligned so that a sphere was centered on the receptive field of the neuron under study. This sphere was randomly chosen on half of the trials to be an oddball, and on the other half the oddball was positioned in the opposite quadrant. Thus there are four trial types: LA and LB pop-out positioned either in the receptive field or the opposite quadrant (a pop-out stimulus was present on the display on every trial). After 300 ms, the fixation point disappeared and two target dots appeared. The animal was required to immediately saccade to the target over the oddball sphere (the location of the pop-out) to receive a reward.

After initial training, both animals performed this task with high accuracy for both LA and LB pop-outs at high contrast. Once performance had reached a steady state, we varied the contrast of the spheres to modify the difficulty of the animal's detection of pop-out. We used five contrasts: 3, 6, 12, 25, and 50%. These were designed to vary from nearly imperceptible pop-out at the low end to easy detection at the high end. We calculated the percentage correct for both types of pop-out for each contrast level from a session in which the animal performed an average of 30 trials of each condition (range from 10 to 80). The average percentage correct at each contrast level is shown in Fig. 2A (average across 114 sessions in two subjects). The animals clearly were better at detecting LB pop-out than LA pop-out at contrasts >6%.

For each session, we fit a Naka–Rushton curve to the data. From this fit we obtained a measure of half-saturation ($C_{50}$), the value of contrast at which the performance is half of the maximum. The $C_{50}$ for LA pop-out ($12.4 \pm 4.2\%$) was higher than that for LB pop-out ($9.9 \pm 3.3\%$), indicating an increased difficulty in detecting LA pop-out relative to LB (Fig. 2B). This difference was statistically significant for the combined data (Wilcoxon sign-rank test, $P < 0.0001$) and also for the data from monkey D ($P < 0.0001$) and monkey F ($P = 0.002$) individually.

Our visual system tends to assume objects in the entire visual scene are illuminated in the same way. Lighting from above is more commonly encountered, and thus is assumed by default, than lighting from below (Ramachandran 1988). Thus when the LB stimulus (which can be perceived as a convex sphere lit from below) is juxtaposed with LA stimuli, it is perceived as a concave sphere (a dent) in a field of convex spheres (bumps). Concavity is less common than convexity in natural scenes (Potetz and Lee 2003), and spheres lit from below are more unusual than spheres lit from above. In both interpretations, the LB stimuli are considered less usual or familiar than the LA stimuli. Human psychophysical studies with shape from shape-from-shading pop-out stimuli show that objects shaded to appear lit from below pop out better than those shaded to appear lit from above (Sun and Perona 1996a,b, 1998). This has been interpreted to mean that unusual or complicated objects generate a stronger pop-out percept than objects with a familiar or simple interpretation (Sun and Perona 1996b; Treisman 1985).

### Example responses

For each recording, we used the trials in which the oddball stimulus was presented in the receptive field to determine the pop-out response. The "uniform" (or background) response was measured from the trials in which the pop-out was of the opposite shading and located in the opposite quadrant. Thus the baseline
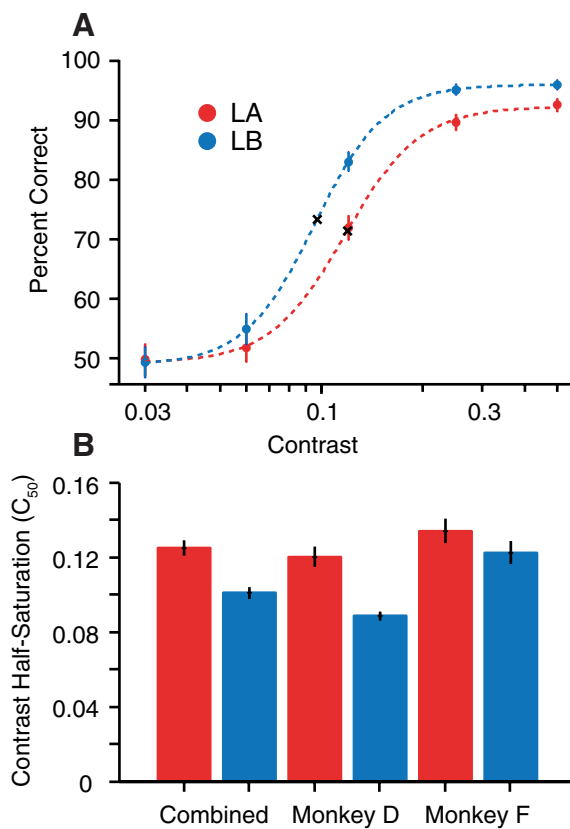
## A



## B

FIG. 2. Psychometric performance function. *A*: contrast of the stimulus (all of the spheres simultaneously) was varied between 3 and 50% randomly on each trial. Dark gray and light gray points indicate the percentage correct for LA and LB pop-outs, respectively, averaged across 114 behavioral sessions from 2 animals. Error bars shown are ±1 SE computed across the same 114 sessions. Data were fit with a Naka–Rushton curve, shown with dashed lines. Half-saturation ($C_{50}$, the value of contrast at which the performance is half of the maximum) is indicated with a black ×. LB pop-out was significantly easier to detect than LA pop-out across a wide range of contrasts. *B*: we fit the behavioral data from each session with a Naka–Rushton curve and extracted the $C_{50}$ value from that fit. Dark gray and light gray bars show the average $C_{50}$ (±1 SE) for all of the data and for monkey D and monkey F individually. $C_{50}$ for LB pop-out was significantly lower than that for LA pop-out in the combined data and for each animal individually.

response for LA pop-out (when the LA oddball is in the receptive field) is the condition when the LB pop-out is located in the opposite quadrant (see Supplemental Fig. 1 for details). In both conditions, the receptive field is presented with the same LA shaded sphere—the difference between the stimuli is evident only at distances of >1° from the receptive field center.

Figure 3 shows the responses of three V1 neurons recorded during the performance of the behavioral task described earlier with the highest contrast stimulus (50%). In each recording, the initial response to the pop-out (solid line) and uniform (dashed line) stimulus are equal. After some time, the two traces diverge, and the response to the pop-out condition exceeds that to the uniform condition, as reported previously by Lee et al. (2002).

### Neuronal contrast response

To determine the magnitude of this effect in the entire population of neurons, we measured the response rate in a window from 100 to 300 ms after stimulus onset. This window was chosen to include the time at which contextual modulation

began in the population response and extends for the duration of the stimulus. Responses computed for all five contrasts are shown in Fig. 4, *A* and *B*. Both the pop-out (solid lines) and uniform (dashed lines) responses grow with contrast in a similar manner for both LA (red lines) and LB (blue lines).

Because the responses to both the pop-out and uniform stimuli grow with contrast, their absolute magnitude may not be the most meaningful measure. We computed a modulation ratio between these two responses (their difference divided by their sum; see METHODS), shown in Fig. 4*C*. This ratio is positive when the response to the pop-out stimulus exceeds that to the uniform stimulus, and negative when the opposite occurs (ranging from +1 to −1). From this analysis it is clear that the increased contrast does not merely act as a better drive for the CRF and increase the firing rate. Instead, the pop-out and uniform responses grow more distinct at higher contrast.

The modulation to the LB pop-out stimuli slightly exceeds that to the LA pop-out stimuli (at all contrasts except 3%, where the small differences make the measure very noisy).
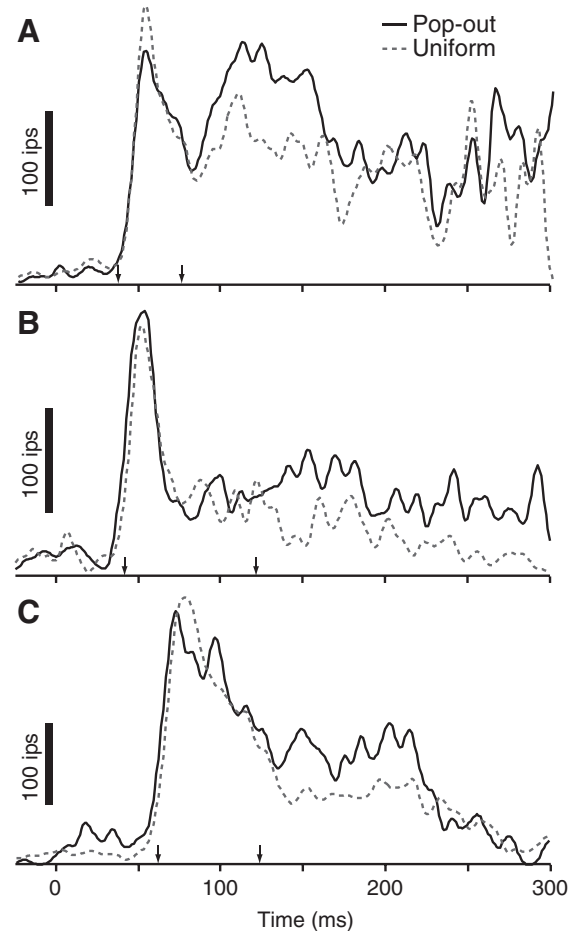
## A

## B

## C



FIG. 3. Example responses demonstrating response modulation to the pop-out stimulus. Here we show the responses of 3 representative V1 neurons to the pop-out stimuli. All of the responses are aligned to stimulus onset (at *time 0*). Black lines in each panel are the average response traces for the pop-out condition, when the oddball stimulus was in the receptive field (LA for *A* and *B*, and LB for *C*). Dashed lines indicate the reference response, when the receptive field stimulus was the same as the surrounding spheres (uniform condition). Two arrows along the *x*-axis of each panel show the latency to response onset and latency to contextual modulation. Receptive field diameters of the neurons were determined to be 0.75, 0.75, and 0.8°, and the number of trials was 10, 50, and 20, for *A*, *B*, and *C*, respectively.
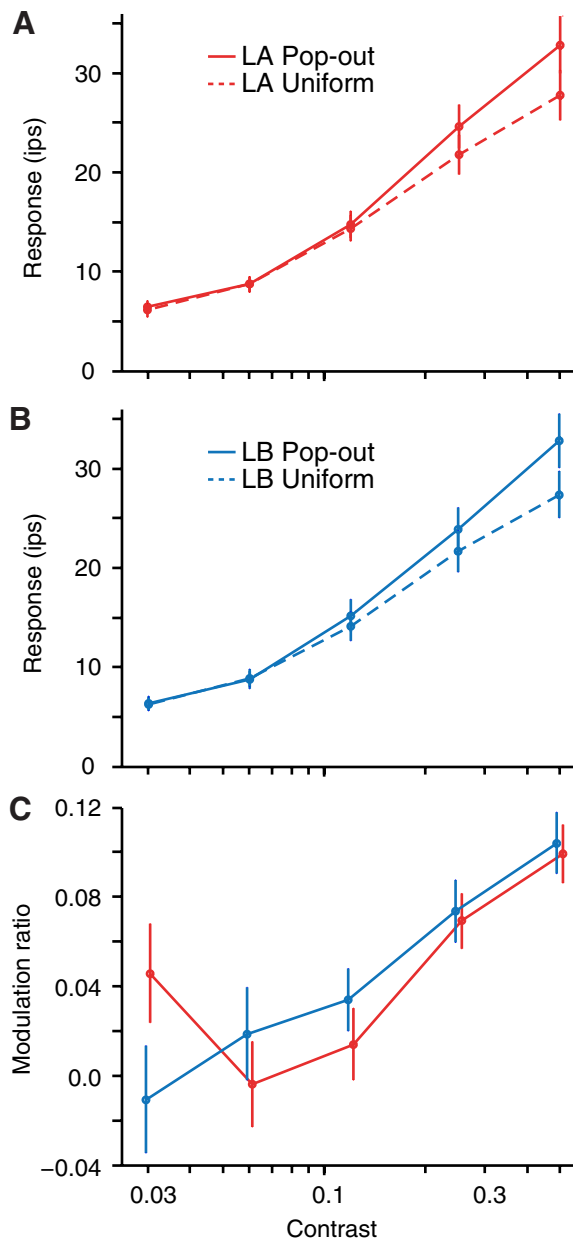
FIG. 4. Contrast response function for pop-out. *A*: for the population of 114 neurons from 2 animals, we show the average response to the LA stimuli in a window from 100 to 300 ms after stimulus onset. Solid lines show the response in the pop-out condition; dashed lines show the response in the uniform condition. *B*: for the population of 114 neurons from 2 animals, we show the average response to the LB stimuli in a window from 100 to 300 ms after stimulus onset. Solid lines show the response in the pop-out condition, and the dashed lines show the response in the uniform condition. *C*: in the same neurons, we computed a modulation ratio between the pop-out and uniform response (their difference divided by their sum; see METHODS). These modulation ratios are plotted here as a function of contrast for LA (red) and LB (blue). Error bars shown in both panels are ±1 SE computed across the 114 neurons.

Although the difference is not statistically significant, this effect is consistent with our observation of a behavioral bias for LB stimuli in the same animals.

*Relationship between neuronal response and behavior*

We have described the behavioral performance of macaque monkeys on a pop-out–detection task, and the contrast re-

sponse of individual neurons to pop-out stimuli. We now turn to analysis of the relationship between behavioral performance and neuronal responses.

Using the same methods described earlier for behavioral performance, we measured the performance of each individual neuron. We used the same data presented in Fig. 4, the response measured from 100 to 300 ms after stimulus onset. We measured the difference in response between the pop-out and uniform conditions, and fit a Naka–Rushton curve to determine the half-saturation contrast ($C_{50}$) for the LA and LB pop-out stimuli. The $C_{50}$ for the LB neuronal responses was lower than that for the LA neuronal responses (19.6 and 23.1%, respectively), but this difference was not statistically significant (Wilcoxon sign-rank test, $P = 0.77$). This was expected based on the analysis we performed earlier (Fig. 4*C*), which showed that the response modulation to LA and LB is similar.

The ability of the subjects to detect the pop-out stimulus varied between sessions. It is possible that those variations would be correlated with similar variations seen in the neuronal contrast response. We analyzed the simultaneously acquired behavioral and neuronal contrast responses to determine whether there was an interaction. The behavioral and neuronal contrast half-saturations are plotted in Fig. 5*A*. We found no significant correlation between these measures for the LA (Pearson's $r = -0.04$, $n = 43$, $P = 0.79$) or LB (Pearson's $r = -0.01$, $n = 47$, $P = 0.946$) pop-out stimuli. Furthermore, the ratio of the neuronal and behavioral $C_{50}$ values was similar for LA and LB (2.07 and 2.06, respectively).

Although we found no relationship between neuronal and behavioral performance across sessions, it is possible that correlation in these two measures might be revealed by analyzing the individual trials. Correlation of this type is frequently quantified with choice probability (CP), a measure of the ability of an ideal observer to predict the animal's behavioral choices from the neuronal firing rate (Britten et al. 1996; Celebrini and Newsome 1994). In extrastriate visual cortex, significant CPs have routinely been observed using a variety of tasks (Britten et al. 1996; Nienborg and Cumming 2006; Parker et al. 2002; Purushothaman and Bradley 2005; Shadlen and Newsome 2001; Uka and DeAngelis 2004).

The measurement of CP requires a large number of trials in which the subject makes each of the possible choices (in our case, saccades to or away from the pop-out stimulus). In the 2AFC pop-out–detection task, the two lowest contrasts (3 and 6%) were well suited to this purpose. After normalizing the responses, we combined the data from these two contrasts and computed a *grand* CP. The frequency histogram in Fig. 5*B* shows the distribution of these CP values across our population of neurons. The average CP was 0.4991, not significantly different from the value of 0.5 predicted by chance. The values for each of the monkeys individually (0.5051 and 0.4955) and for LA and LB stimuli computed separately (0.4956 and 0.5000) were also at chance levels. Using a permutation test, we determined that 15 of the neurons showed a value for CP that was statistically distinguishable from chance ($P < 0.05$, 9 cells had CP values below 0.5 and 6 cells above).

Nienborg and Cumming (2006) reported significant CPs in V2 but not V1 in subjects who performed a task (disparity discrimination of random-dot stereograms) for which substantial CPs have been reported in area MT (Parker et al. 2002; Uka and DeAngelis 2004). Significant CPs have been observed in
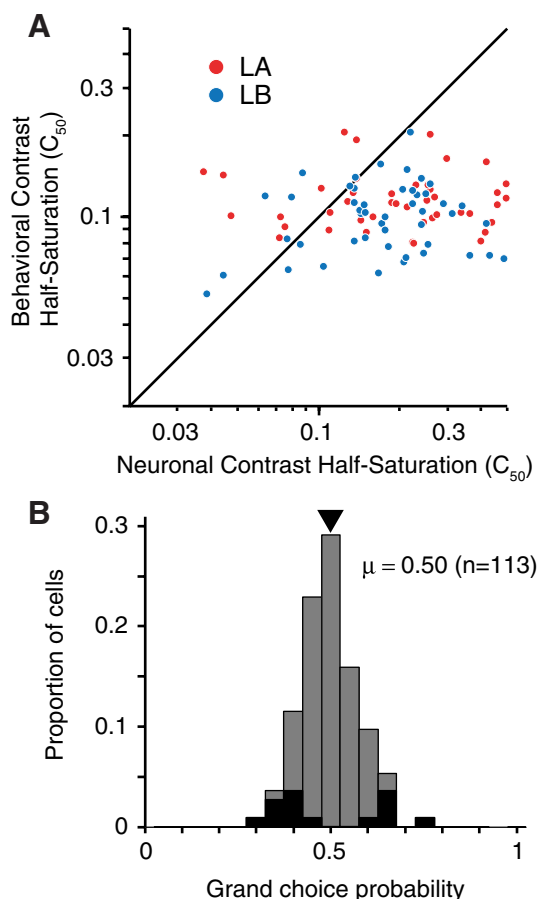
FIG. 5. Comparison between behavioral and neuronal responses. *A*: here we plot the behavioral contrast half-saturation ($C_{50}$) as a function of neuronal contrast half-saturation ($C_{50}$), measured simultaneously while the animals performed a 2AFC pop-out detection task. Most of the points fall below the line of unity slope, indicating that the animals were more sensitive than the neurons. Red and blue points indicate the data for LA and LB pop-out contrast response functions, respectively. Behavioral and neuronal half-saturation values were slightly lower for LB than for LA. *B*: for a population of 113 neurons from 2 animals, this plot is a frequency histogram of the *grand* choice probability. This includes combined data from the 3 and 6% contrast trials. Mean of the distribution was 0.5, which is the same value predicted by chance. Using a permutation test, we determined that 15 of the neurons had a choice probability (CP) that was outside the 95% confidence interval (black bars).

V1, however, in animals performing a reaction time visual-detection task (Palmer et al. 2007). Our data show that the trial-by-trial responses of individual V1 neurons are not well suited to explaining the animal's behavior in a pop-out–detection task. This indicates that the pooling of neuronal responses to produce behavioral output must operate on large groups of V1 or extrastriate neurons. In the next section, we attempt to determine what information about the pop-out stimulus is present in a population of V1 neurons that could be pooled by a later stage of visual processing.

*Population response to pop-out stimuli*

The population average responses to LA and LB stimuli are based on measurements of response in a fixed time window, so transient effects might be neglected. We considered that analysis of the time course of the response might reveal the neural differences between the two types of pop-out that we expected

based on the behavioral difference. The behavioral contrast threshold for LA and LB stimuli is maximally different at the middle range of contrasts (10–30% based on the psychometric curves in Fig. 2*A*). Ideally, we would analyze the neural responses to pop-out stimuli at contrasts in that range. However, we found that the response at intermediate contrast (12 and 25%) was noisier and showed a smaller pop-out modulation. We therefore focused our further analysis on the 50% contrast data.

One means of viewing the effect of the pop-out stimulus across the entire population of neurons is with a *grand* post-stimulus time histogram (PSTH). For each neuron, we found the average responses aligned to stimulus onset for both types of pop-out stimuli at high contrast (50%). The distributions of firing rates evoked by the LA (28.18 spikes/s) and LB (26.81 spikes/s) uniform stimuli were not significantly different (*t*-test, $P = 0.30$). These responses were normalized by the peak firing rate (separately for LA and LB) and averaged across all neurons. This produced the *grand* PSTHs shown in Fig. 6, *A* and *B*. From this analysis, we found that the population response to LB pop-out (Fig. 6*B*) tended to appear sooner than LA pop-out (Fig. 6*A*) by about 28 ms (113 and 141 ms, respectively). This latency was measured as the time at which the pop-out and uniform responses diverged by more than the sum of the SE for the two curves (i.e., when the error bars first do not overlap).

To compare the magnitude of the LA and LB pop-out effects quantitatively, we computed the integrated difference over time between these normalized pop-out and uniform responses for both LA and LB stimuli. These cumulative difference curves are shown in Fig. 6*C*. The response to the LB pop-out stimulus grows faster than that to LA, and reaches a higher level at the end of the fixation period (300 ms). These neuronal results are consistent with our psychophysical observations in the same animals: LB pop-out is easier to detect than LA pop-out.

*Modulation of response to pop-out stimuli*

Our analysis of the combined data from two animals performing the 2AFC pop-out–detection task revealed a tendency for modulation due to LB pop-out to occur sooner and with a somewhat larger magnitude. We wanted to determine whether these effects were evident in the analysis of individual neurons. We therefore computed the modulation ratio (as described earlier) for all of the 114 neurons from two animals. For each neuron, we averaged the modulation ratio of its response to the LA and LB stimuli (as we showed in Fig. 4*C*, these ratios tended to be the same at high contrast). Frequency histograms of this measure of response are shown in Fig. 7*A*. The mean ratio across the population was 0.095, and was 0.09 and 0.10 for the LA and LB responses alone, respectively. All of these groups were significantly different from zero (Wilcoxon sign-rank test, $P < 0.0001$). The range of modulation ratios indicates that many neurons did not show significant enhancement for the pop-out stimulus—in fact, some responded less to the pop-out stimulus than the uniform stimulus. We therefore sought to restrict our analysis only to neurons that showed a significant difference in their responses to the two stimuli in the 100- to 300-ms window (Wilcoxon rank-sum test, $P < 0.05$). Of the 114 neurons, we found 38 (33%, black bars in Fig. 7*A*)
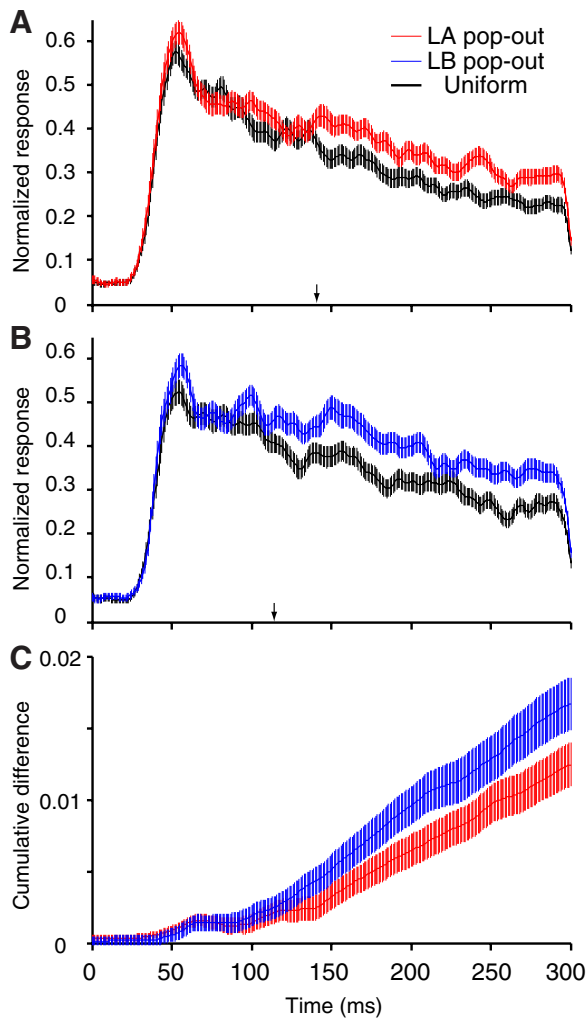
FIG. 6. Grand poststimulus time histogram (PSTH) showing the average timing of the response modulation across the population. *A*: here we show the timing of response modulation to the LA pop-out stimulus (red line) relative to the uniform condition (black line). Response traces from each neuron were aligned to stimulus onset, normalized to a peak value of one, and then averaged pointwise. Error bars are ±1 SE computed for each millisecond time point across the 114 neurons from 2 animals. Two curves diverge (the error bars separate completely) at 141 ms (indicated by the arrow along the *x*-axis). *B*: this panel shows the timing of response modulation to the LB pop-out stimulus (blue line) relative to the uniform condition (black line). Two curves diverge at 113 ms (indicated by the arrow along the *x*-axis). *C*: to highlight the difference in response between the LA and LB pop-out conditions, for each neuron we calculated the cumulative difference between the response to the pop-out and uniform conditions. Average of these cumulative difference traces is shown here for LA (red) and LB (blue) pop-outs. Curves begin to diverge around the time of the initial response modulation, with the LB pop-out response modulation growing faster and to a higher level.

that met this criterion for the LA stimulus (*n* = 23), the LB stimulus (*n* = 25), or both (*n* = 10). Within this group, 36 neurons had an increased response to the pop-out stimulus (average modulation ratio 0.24) and two neurons had a decreased response (average modulation ratio −0.06). In summary, we found that roughly one third of V1 neurons exhibit significant contextual modulation to a shape-from-shading pop-out stimulus. For >90% of these neurons the modulation involves an increased response to the pop-out stimulus.

The distributions of modulation ratio to LA and LB stimuli, computed within the 100- to 300-ms window after stimulus

onset, were not significantly different from each other for all neurons (Wilcoxon sign-rank test, *P* = 0.22) or only for those with significant positive modulation (Wilcoxon rank-sum test, *P* = 0.76). In addition, the modulation of response for a given neuron to the LA and LB stimuli did not tend to be correlated (Fig. 7*B*, *r* = 0.08, *P* = 0.37). Finally, a neuron's initial response to the LA or LB stimuli was not predictive of the magnitude of its later response modulation. The modulation ratio to LA and LB stimuli was not significantly correlated with the early portion of each neuron's response (the first 50 ms after response onset) to that stimulus (*r* = −0.04, *P* = 0.71 and *r* = −0.07, *P* = 0.46). We also computed a preference ratio
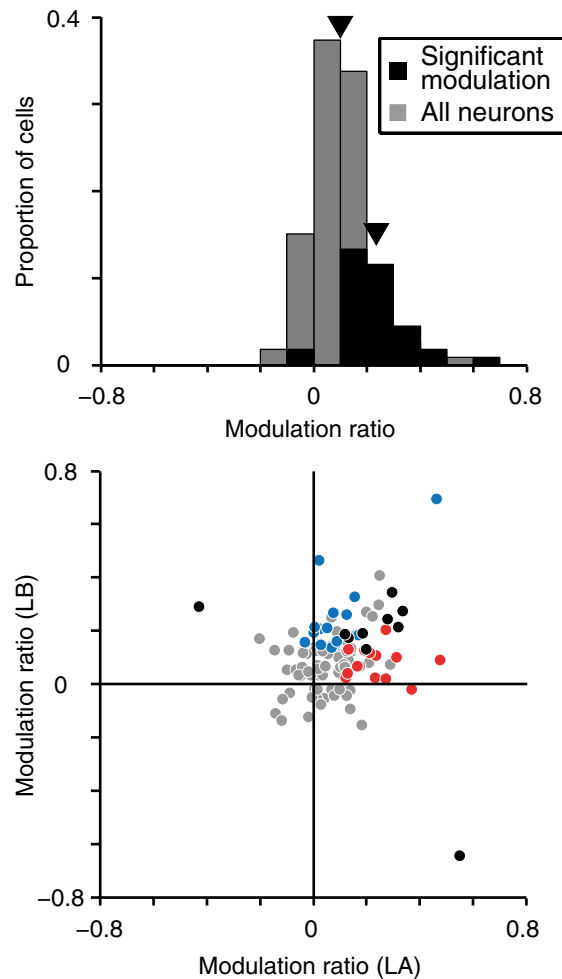


FIG. 7. Strength of the pop-out response. *A*: gray bars are used to plot the frequency histogram of modulation ratios for the entire population of neurons (computed from 100 to 300 ms after stimulus onset) to the LA and LB pop-out stimuli. For each neuron, we averaged the response to the LA and LB stimuli. Black bars represent the neurons that showed a significant difference in spike count between the pop-out and uniform stimulus. Of these 38 neurons, 23 met the significance criterion for the LA stimulus only (average modulation ratio 0.2220), 25 for the LB stimulus only (average modulation ratio 0.2095), and 10 for both stimuli (average modulation ratio 0.1696, for these 10 neurons we averaged the LA and LB response in the histogram shown). Arrows indicate the mean of the 2 distributions (0.0951 and 0.2276). *B*: relationship between each neuron's response to the LA and LB pop-out conditions is shown in this scatterplot. Black circles indicate neurons that showed significant modulation to both LA and LB; the red circles indicate neurons that showed significant modulation to LA but not LB; the blue circles indicate neurons that showed significant modulation to LB but not LA; and the gray circles indicate neurons that did not show significant modulation to either stimulus. We found no significant correlation between modulation to the LA and LB pop-out stimuli.

based on this early response window by computing the difference between the response to LA and LB divided by the sum (analogous to the modulation ratio). There was no significant correlation between a neuron's early preference for LA or LB and its later modulation ratio to either LA ($r = 0.09$, $P = 0.33$) or LB ($r = 0.001$, $P = 0.99$).

Overall, we conclude that responses to these pop-out stimuli are quite strong in >30% of V1 neurons. However, a neuron's initial responses were not predictive of its later modulation to pop-out. Furthermore, neurons that show strong response modulation to one type of pop-out stimulus do not necessarily do so for other stimuli. This result is consistent with the idea that pop-out modulation is overlaid on the relatively simple tuning properties of V1 neurons. Although the LA and LB modulations appear to target distinct populations of V1 neurons, it is unclear at this point which aspect of the tuning properties of the V1 neurons determines whether they will be targeted by LA or LB modulation. It is also tempting to speculate that this may reflect some sort of competitive dynamics between the populations of neurons that respond to the LA and LB stimuli.

*Latency of response to pop-out stimuli*

Most studies of the dynamics of contextual modulation effects in V1 have used population histograms (like those in Fig. 6, *A* and *B*) to determine the latency of the effect. Although this can be an effective method, it is insensitive to the response dynamics of individual neurons. For instance, it can be dominated by cells with large modulation responses or short latencies. We sought to determine the distribution of latencies

of contextual modulation to gain insight into the underlying mechanism.

We determined the onset latency by measuring when the spike count in a sliding 10-ms window significantly exceeded the baseline response. The contextual modulation latency was determined with the same procedure and a 20-ms window (see METHODS). Using this method, we were able to find an onset latency for 75% and modulation latency for about 30% of the neurons for each of the types of pop-out stimuli. We found that the onset latency was about 50 ms after stimulus onset and nearly identical for LA and LB (Fig. 8, *A* and *C*). Furthermore, this value is similar to the onset latency reported previously in V1 for sinusoidal grating stimuli (Bair et al. 2003; Smith et al. 2006). The contextual modulation effect was evident at a much longer delay, >130 ms after stimulus onset (Fig. 8, *B* and *E*). In addition, the modulation to the LB pop-out stimulus occurred about 19 ms earlier than that to the LA pop-out stimulus (*t*-test, $P = 0.04$). Because the measurement of modulation latency might be affected by variation in the onset latency, we calculated the difference in these two latencies (modulation delay, Fig. 8, *C* and *F*). In comparing this modulation delay between the two pop-out stimuli, we found that the latency for LB pop-out stimuli was less than that for LA pop-out stimuli by about 19 ms (*t*-test, $P = 0.05$). In the 17 neurons for which we were able to determine latency for both stimuli, the latency for LB pop-out stimuli was similarly earlier than that for LA pop-out stimuli (Supplemental Fig. 2). We found no correlation, however, between the animal's behavioral performance on the task in a given session ($C_{50}$) and the modulation latency
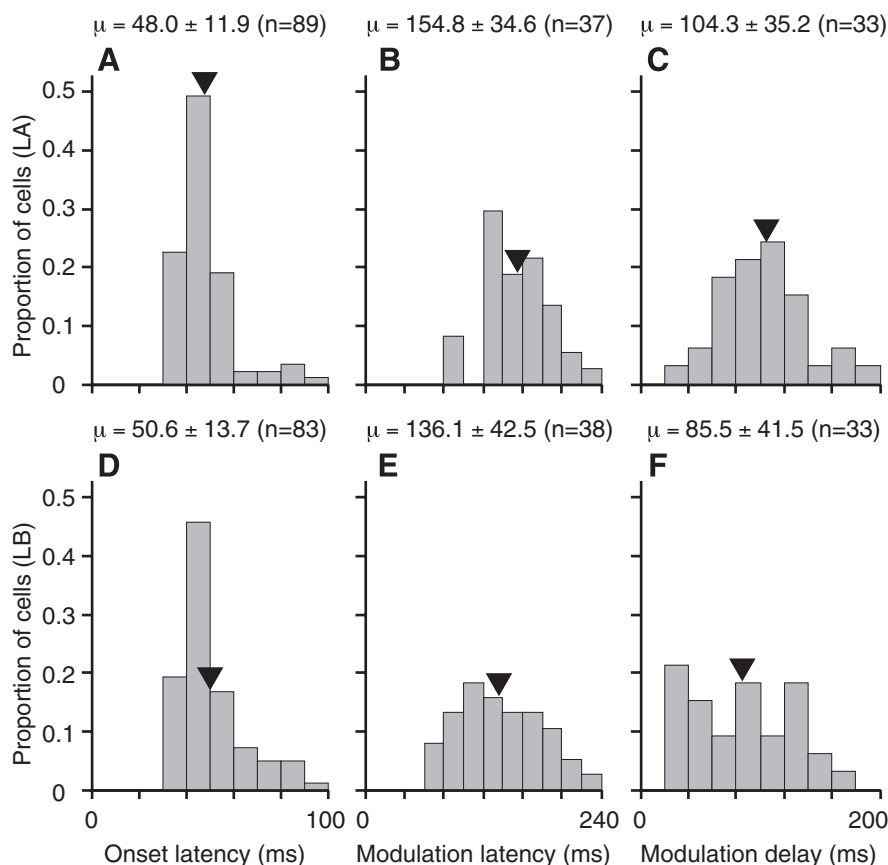


FIG. 8. Latency of response. *A*: here we show a frequency histogram of the onset latency to the LA pop-out stimulus. We attempted to measure latency in the responses of our population of 114 neurons. Arrow indicates the mean of the distribution. Values for the mean, SD, and number of neurons are indicated at the top of each panel. *B*: this is the frequency histogram for the modulation latency, determined by finding the time of the difference between the pop-out and uniform conditions. *C*: for each of the neurons in which we were able to determine the modulation latency, we subtracted the onset latency from that value. This produced a measure of the modulation delay, or the time between the response onset and the modulation induced by the pop-out stimulus. *D–F*: these frequency histograms show the same 3 latencies determined for the responses to the LB pop-out stimulus.

of the recorded neuron for LA ($r = 0.08$, $P = 0.66$) or LB ($r = 0.19$, $P = 0.25$).

We also analyzed data collected from three additional animals before the start of these experiments (111 neurons). Two of these animals were used in a previous study (Lee et al. 2002). Each of them performed a pop-out–detection task, although it was performed only at high contrast and the animals were allowed to saccade to the pop-out stimulus as soon as they desired (a reaction time task). The LA and LB pop-out stimuli were randomly interleaved among other arrays of pop-out stimuli. Two of the animals also were trained exclusively with LA or LB pop-out stimuli for a portion of the previous study (Lee et al. 2002). These animals were not included in the present study because of the task differences and their training history. However, we were able to determine the latency of pop-out modulation in >30 neurons from this population, and found that the response to LB pop-out stimuli occurred roughly 10 ms before that to LA pop-out stimuli. These results from individual neurons, along with the grand PSTH shown in Fig. 6, provide evidence for a neuronal bias for LB pop-out that matches the behavior effect.

Previous studies of contextual modulation effects in V1 have implicated surround suppression as a possible mechanism. Our stimuli are designed to try to avoid the effects of iso-orientation surround suppression because the orientation structure in the stimulus was the same in the receptive field and the surround. The only difference between the center and surround was polarity of shading (phase), and surround suppression was previously shown to be largely insensitive to phase (DeAngelis et al. 1994; Webb et al. 2005). In Fig. 9, we plot the modulation latency as a function of onset latency for the neurons in which we were able to determine both values. These two values were not significantly correlated ($r = 0.10$, $n = 66$, $P = 0.42$) across the full population. For comparison, we have also plotted the latencies reported by Smith et al. (2006) for surround suppression (open circles) and cross-orientation suppression (filled gray circles). Latencies of both of these suppressive effects show a dependence on the onset latency ($r = 0.69$ and $r = 0.45$). However, although the onset latencies were similar to these very different stimuli, the delay to see the effect of the pop-out stimuli was much longer than the latency for surround suppression (145.3 vs. 59.4 ms; $t$-test, $P < 0.0001$). The mechanism that supports contextual modulation due to shape-from-shading pop-out stimuli appears to occur on a vastly different timescale than that which generates iso-orientation surround suppression.

DISCUSSION

We have measured a behavioral bias for greater ease of detecting unusual pop-out items among a field of distractors in macaque monkeys, consistent with human psychophysical literature. This effect is paralleled by a tendency for V1 neurons to show earlier and somewhat larger contextual modulation effects for the same stimuli. Overall, the contextual modulation due to pop-out stimuli occurs quite late, on average 145 ms after stimulus onset and 95 ms after response onset. This is much later than what has been reported for iso-orientation surround suppression and cross-orientation suppression (Bair et al. 2003; Smith et al. 2006). The range of modulation latencies is quite broad, from as early as 70 ms to as late as 230
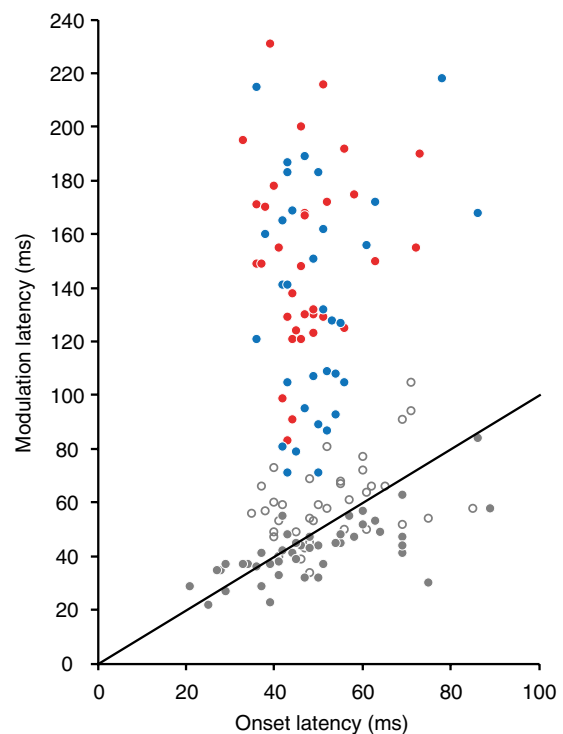


FIG. 9. Relationship between modulation latency and onset latency. Here the modulation latency is plotted as a function of the onset latency for both the LA (red circles) and LB (blue circles) pop-out stimuli. Open gray circles and filled gray circle*s* show the relationship between suppression latency and onset latency for surround suppression and cross-orientation suppression, respectively, using data from Smith et al. (2006) recorded in anesthetized macaque V1. Although the average onset latency is the same, the latency to the contextual modulation due to the pop-out stimulus is much later than that for surround suppression (induced by a large, optimally oriented, drifting sinusoidal grating).

ms, and the effect occurred earlier for stimuli for which the animal's behavioral performance was more accurate. We now consider how these results may shed light on the mechanisms that underlie contextual modulation effects in primary visual cortex.

*Relation to iso-orientation surround suppression*

The phenomenon of surround suppression was first understood as a reduction in response to the lengthening of an optimally oriented bar stimulus (Hubel and Wiesel 1968). This was originally thought to be mediated by horizontal intracortical connections within V1 (DeAngelis et al. 1994; Gilbert and Wiesel 1983; Knierim and Van Essen 1992; Nelson and Frost 1978), although feedback from higher cortical areas has recently been a popular explanation (Angelucci et al. 2002; Bair et al. 2003; Cavanaugh et al. 2002; Knierim and Van Essen 1992; Levitt and Lund 2002; Zipser et al. 1996), based on the temporal delay of the suppression. In a number of recent studies, the response of a V1 neuron has been shown to be influenced by a wide variety of stimuli outside the classical receptive field (Lamme 1995; Lee and Nguyen 2001; Lee et al. 1998, 2002; Roelfsema et al. 2003; Zipser et al. 1996). Some of these effects occur without the presence of orientation contrast, and do not depend strongly on the preferences of the neuron's classical receptive field. The breadth of these influ-

ences brings into question whether a single mechanism can subserve them all.

Recent detailed studies of the dynamics (Bair et al. 2003; Smith et al. 2006) and phase selectivity (DeAngelis et al. 1994; Webb et al. 2005) of classical iso-orientation surround suppression reveal that these properties are incompatible with those we report here for modulation to perceptual pop-out stimuli. Specifically, surround suppression occurs only 8 ms after the initial CRF-driven response and is relatively insensitive to the phase of the surround stimulus. Webb et al. (2005) proposed that the phenomenon of surround suppression might consist of two separate mechanisms—one that is fast and relatively orientation insensitive and another that is slower and orientation selective. Some studies have reported two distinct time courses for orientation texture stimuli—a fast, nonspecific suppression and a slow, orientation-specific response modulation (Knierim and Van Essen 1992; Nothdurft et al. 1999). Similarly, we previously reported that the response to a single sphere (singleton) exceeds that to an array of spheres (Lee et al. 2002). This difference occurred very early in the response time course, far before modulation to the shape-from-shading pop-out stimuli. We report here that modulation to these stimuli is much slower than the effects of iso-orientation surround suppression. Furthermore, an orientation-selective suppressive mechanism would not be able to distinguish between our pop-out and uniform stimuli. Our data provide evidence for the existence of a third form of contextual modulation, distinct from iso-orientation surround suppression, which acts with a significant delay after response onset ($>85$ ms on average).

Does this third form of contextual modulation occur through a separate mechanism from iso-orientation surround suppression? The effect we observe is smaller in magnitude and occurs in fewer neurons than surround suppression. Cavanaugh et al. (2002) reported a suppression index (the magnitude of the suppressed response normalized by the peak response) of 0.38 in V1 neurons. Computed in the same way, we find an index of 0.16, less than one half of the magnitude. Iso-orientation surround suppression (suppression index of $\geq 0.10$) has been observed in 85% of V1 neurons (Cavanaugh et al. 2002), whereas we observe significant modulation to pop-out stimuli in 30% of neurons (however, 69 of 114 neurons had a suppression index of $\geq 0.10$). Because latency is known to vary with stimulus contrast (Albrecht 1995; Carandini et al. 1997; Gawne et al. 1996), the longer modulation latencies we observe might be due to decreased effectiveness at driving the underlying mechanism. However, the match in onset latency between the pop-out stimulus and sinusoidal gratings indicates that the effective contrast is similar. Nonetheless, it is possible that the long latencies we observe for contextual modulation to pop-out stimuli are not due to a separate mechanism from iso-orientation surround suppression. Instead, a single mechanism might underlie these various phenomena, but respond with differing latency based on stimuli of differing complexity and processing requirements. This is consistent with the model proposed by Angelucci et al. (2002), in which a V1 neuron's RF is determined by feedback connections from multiple extrastriate areas and lateral connections within V1, each of which has a different spatial extent and stimulus selectivity. In this model, the different numbers of neurons that respond to various contextual modulation stimuli could be related to the strength and amount of connections from different areas in extrastriate cortex.

## Spatial origins of contextual modulation

Our stimulus was designed to activate the classical receptive field of a V1 neuron with a single shaded sphere, isolating the remainder of the array to the extraclassical surround. We carefully measured the receptive field size and position with a small light bar (using the minimum response field technique) to determine the positioning of the stimulus array, and also ensured that the neuron did not respond to the light bar in the surrounding sphere locations. In a previous study, we tested neuronal responses to an array of spheres without the element covering the RF (Lee et al. 2002) and saw little or no significant change in response. In the neurons that did have a small response it occurred at or shortly after the onset latency for the neuron, much earlier than the response modulation to the pop-out stimulus. These results are a very close match with previous studies using a pop-out stimulus composed of small line segments (Knierim and Van Essen 1992; Nothdurft et al. 1999). The responses of neurons in these two studies to the "surround only" condition, without any stimulus in the area determined to be the RF, might be due to known changes in spatial integration based on stimulus contrast (Cavanaugh et al. 2002; Kapadia et al. 1999; Sceniak et al. 1999). This might indicate that some of the modulation we observe is due to stimulation of the outer regions of the RF, and not the extraclassical surround.

Although this remains a possibility, two lines of evidence indicate that the long-latency modulation we report here is not due to direct stimulation at the edges of the CRF. First, the onset latency that we observe to pop-out stimuli is very similar to that reported for high-contrast gratings (Fig. 9), indicating that the effective contrast of these stimuli is similar. If this is true, then previously observed changes in spatial integration with contrast (Cavanaugh et al. 2002; Kapadia et al. 1999; Sceniak et al. 1999) may not play a significant role for our stimulus. Second, studies of the time course of suppression and facilitation with stimuli of multiple sizes and distances from the RF have not reported a latency of modulation as long as that reported here. Suppression latencies ranging from 10 to 20 ms after response onset are typical with grating stimuli (Bair et al. 2003; Xing et al. 2005), but a delay of $>30$ ms is rare even when the suppressive surround stimulus is withdrawn far away from the RF (Bair et al. 2003). Hupé et al. (2001) reported facilitation and suppression of V1 neuronal responses by a static texture surround with typical latencies of, respectively, 20 and 40 ms after response onset. The modulation latencies we report, 95 ms after response onset, are much later than the suppressive or facilitatory effects observed in areas ranging from the edges of the RF to the distant surround. Nonetheless, the array of spheres we used for this study, a relatively sparse field of elements not optimized for each individual neuron, is quite different from the typical stimulus used to probe surround suppression. More detailed study of the modulation induced by various spatial configurations and sphere sizes will be necessary to definitively determine the spatial origins of the signals we observe.

*Effects of attentional modulation*

On possible factor underlying these contextual effects might be a top-down attentional signal. The study of attention in primary visual cortex has yielded varying results and, although some effects have been found in V1 under certain conditions (Ito and Gilbert 1999; Motter 1993; Roelfsema et al. 1998), others have found no effect at all (Grunewald et al. 2002; Luck et al. 1997; Marcus and Van Essen 2002). Attentional modulation, when it is present in a V1 neuron, can be evident even in the early phase of the response (before the peak) (McAdams and Reid 2005) when the animal is cued to a particular location. However, it can also occur with a substantial delay after stimulus onset in a more demanding task where the animal does not know in advance where the relevant stimulus will appear (Roelfsema et al. 1998). In V4, attention effects are known to be highly specific for stimulus features (Bichot et al. 2005) and are strongest when competing stimuli are placed inside the receptive field (Luck et al. 1997). We previously have shown that modulation to the pop-out stimuli used here was reduced when the animals performed an attentionally demanding distraction task (Lee et al. 2002). Furthermore, training the animal to detect the location of the pop-out stimulus increased the magnitude of the modulation effect in both V1 and V2 (Lee et al. 2002). In V4, however, attentional modulation is stronger for low-contrast stimuli (Reynolds et al. 2000), whereas we find that the effect of contextual modulation is strongest with high-contrast stimuli. We also find no correlation between neural responses and the animal's behavioral choice, which might be expected if attention played a role. Thus although the spatial and temporal profile of attentional effects is potentially broad enough to influence neuronal responses in our task, a nonspecific top-down spatial attention signal is insufficient to explain the contextual modulation we observe.

*Asymmetry of contextual modulation*

We report here that neuronal contextual modulation is larger and evident sooner for the lighting from below (LB) than the lighting from above (LA) pop-out stimulus. This matches our observation that lighting from below pop-out is easier to detect, based on psychophysical observations in the same animals. Thus some of the contextual modulation we observe is feature specific, rather than a nonspecific enhancement effect. If attention plays a role in these modulation responses, it is not through a nonspecific spatial enhancement signal. The process of exogenous or reflexive attention (Jonides and Yantis 1988; Posner 1980; Yantis and Jonides 1984), reacting to salience signals computed by an automatic process (e.g., pop-out), might well play a role in contextual modulation.

Our physiological and psychophysical results are consistent with previous work on perceptual pop-out in human subjects. Treisman (1985) reported that the pop-out effect is strong for a tilted line segment in the presence of vertical distractors, whereas it is weak for the opposite configuration. Sun and Perona (1996a,b, 1998) found a similar result for shaded cube and sphere stimuli, where detection was best for lit-from-below shapes among lit-from-above distractors. A simple interpretation of these results is that pop-out effects are strongest for an unfamiliar or unusual target stimulus in the presence of an array of familiar distractors. Lighting from above is indeed the default assumption for ambiguous visual stimuli (Ramachandran 1988). Furthermore, Potetz and Lee (2003) analyzed the statistical correlation between luminance and range images of natural scenes and found that the mean luminance image associated with a convex object indeed resembles the LA pattern, suggesting convexity is more prevalent than concavity and lighting tends to come from above. In our paradigm, the concave "dent" (LB pop-out stimulus) was unusual and generated a stronger pop-out response. Thus our behavioral finding of a lower-contrast threshold for LB pop-out extends previous human psychophysical findings to include nonhuman primates, and our matching physiological results demonstrate a neural basis for this phenomenon.

*Is it really the neural correlate of pop-out?*

Shape-from-shading stimuli, like those used in this study, readily evoke a perception of 3D shape in human observers. We are not aware of any published study in nonhuman primates that demonstrates that they perceive these same stimuli as 3D shapes, as do humans. One way to test for such an effect would be to measure whether shading influences the perception of stereoscopic depth. Our laboratory has recently undertaken such a study, and preliminary results show that the shading used in our LA stimulus biases a monkey to view the object as convex, whereas the shading in our LB stimulus produces a concave bias (unpublished data). This is the expected result if monkeys, like humans, have a default assumption of lighting from above (Ramachandran 1988) and the shape-from-shading stimuli evoke an actual perception of 3D shape. The behavioral and physiological asymmetry that we observe between the LA and LB pop-out conditions also provides some evidence that monkeys perceive these stimuli similarly to humans. Even if the monkeys do not experience the perception of 3D shape from the shaded spheres, the pop-out stimulus elicits a long-latency contextual response in V1 neurons, and the interpretation of our data does not depend solely on the perception of 3D shape by the subjects.

One of the hallmarks of the pop-out–detection task is that reaction time and accuracy are relatively insensitive to the number of distractors (Treisman 1985). Using a reaction time task with stimuli identical to those used here, we previously reported saccadic reaction times of 225–250 ms in two animals (Lee et al. 2002). Bichot and Schall (2002) reported a median saccadic latency of about 225 ms in two monkeys performing a pop-out–detection task with one target and three distractors. They further found that the time of neuronal target discrimination in the frontal eye fields preceded saccade latency by 80 ms. If we consider 225 ms as a typical saccadic reaction time for monkeys performing a pop-out task, then the average modulation latency in our V1 population also precedes saccade latency by about 80 ms, similar to the modulation latency of frontal eye field (FEF) neurons. This seems to rule out a simple process in which V1 passes information about the detected pop-out target to FEF, leading to movement of the eyes. In such a scenario the modulation latencies of FEF neurons would lag behind those of V1 neurons. We are left to consider two interpretations of the data. First, it is possible that the contextual modulation signals we observe in V1 are not involved in the decision to move the eyes. Instead, they may play a role in

computing perceptual salience on a slightly longer timescale than is required for the behavioral choice, perhaps by refining the initial evaluation of the scene. We favor a second interpretation, in which contextual modulation signals emerge as a result of a recurrent interactive computation between multiple levels of the visual hierarchy. If this is the case, it may be that neurons in V1 pass information to their neighbors through horizontal connections and to extrastriate cortex. Long-latency contextual modulation signals, such as those we report here, emerge over time as the network computes perceptual salience through recurrent connections within and between cortical areas. It is possible for a contextual pop-out signal to emerge simultaneously in FEF and V1 through such a recurrent interaction. Additional experiments will be necessary to help distinguish which of these interpretations is a better description of visual processing.

Hegdé and Felleman (2003) compared the responses of V1 neurons to pop-out stimuli and "conjunction-target" stimuli, which do not pop out for human observers. They found that V1 neurons were largely unable to distinguish between these two types of stimuli and concluded that they were coding for center–surround feature discontinuities rather than representing the neural correlates of perceptual pop-out. Our results here do not make this distinction—indeed, we have been careful to indicate that the responses we observe are a contextual modulation due to perceptual pop-out stimuli, and not necessarily the neural correlates of pop-out per se. Nevertheless, the match between our neural and behavioral data on pop-out asymmetry provides strong evidence that V1 is involved in the neural computation of pop-out. The computational processes for extracting pop-out stimuli likely involve recurrent interaction between V1 and extrastriate visual cortex.

In summary, our data provide evidence for a type of contextual modulation that occurs in the absence of orientation contrast between the center and surround and acts with a significant delay after response onset (>70 ms on average). This is a third form of contextual modulation, distinct from the two putative components of surround suppression—one fast and broadly tuned for stimulus properties, another slightly slower and orientation selective. A simple nonspecific top-down signal cannot explain our results. Nonetheless, feedback from extrastriate cortex has been proposed as a mechanism for many types of contextual modulation. Because different extrastriate areas and local circuits are involved in different types of stimulus detection, even if feedback is involved in all contextual modulation we might expect that the latency would vary depending on the stimulus type. Orientation contrast, which can be coded at the lowest level, occurs with the shortest latency. Modulation to the shape-from-shading pop-out stimuli used here requires significant interaction with higher-order areas, and therefore occurs with a longer delay. This involvement of higher-order areas is presumably because of the 3D interpretation of the stimulus that plays a critical role in this pop-out phenomenon (Lee et al. 2002; Ramachandran 1988; Sun and Perona 1996b). Further study of the latency of different forms of contextual modulation, and of the connections between V1 and extrastriate cortex, will be necessary to prove this theory. Nonetheless, our results strongly support the involvement of feedback in contextual modulation due to shape-from-shading pop-out stimuli. The physiological and psychophysical asymmetry that we observe in pop-out responses, as

well as the presence of neurons that respond only to LA or LB pop-out, indicates that this feedback is feature specific. Taken together, our findings suggest that V1 neurons actively participate in a recurrent circuit for the computation and representation of perceptual salience.

REFERENCES

**Albrecht DG.** Visual cortex neurons in monkey and cat: effect of contrast on the spatial and temporal phase transfer functions. *Vis Neurosci* 12: 1191–1210, 1995.

**Albright TD, Stoner GR.** Contextual influences on visual processing. *Annu Rev Neurosci* 25: 339–379, 2002.

**Alexander DM, Wright JJ.** The maximum range and timing of excitatory contextual modulation in monkey primary visual cortex. *Vis Neurosci* 23: 721–728, 2006.

**Angelucci A, Levitt JB, Walton EJ, Hupé JM, Bullier J, Lund JS.** Circuits for local and global signal integration in primary visual cortex. *J Neurosci* 22: 8633–8646, 2002.

**Bair W, Cavanaugh JR, Movshon JA.** Time course and time-distance relationships for surround suppression in macaque V1 neurons. *J Neurosci* 23: 7690–7701, 2003.

**Bergen JR, Julesz B.** Parallel versus serial processing in rapid pattern discrimination. *Nature* 303: 696–698, 1983a.

**Bichot NP, Rossi AF, Desimone R.** Parallel and serial neural mechanisms for visual search in macaque area V4. *Science* 308: 529–534, 2005.

**Bichot NP, Schall JD.** Priming in macaque frontal cortex during popout visual search: feature-based facilitation and location-based inhibition of return. *J Neurosci* 22: 4675–4685, 2002.

**Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA.** A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis Neurosci* 13: 87–100, 1996.

**Carandini M, Heeger DJ, Movshon JA.** Linearity and normalization in simple cells of the macaque primary visual cortex. *J Neurosci* 17: 8621–8644, 1997.

**Cavanaugh JR, Bair W, Movshon JA.** Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J Neurophysiol* 88: 2530–2546, 2002.

**Celebrini S, Newsome WT.** Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. *J Neurosci* 14: 4109–4124, 1994.

**Crist CF, Yamasaki DS, Komatsu H, Wurtz RH.** A grid system and a microsyringe for single cell recording. *J Neurosci Methods* 26: 117–122, 1988.

**DeAngelis GC, Freeman RD, Ohzawa I.** Length and width tuning of neurons in the cat's primary visual cortex. *J Neurophysiol* 71: 347–374, 1994.

**Efron B, Tibshirani RJ.** *An Introduction to the Bootstrap.* New York: Chapman & Hall, 1993.

**Gawne TJ, Kjaer TW, Richmond BJ.** Latency: another potential code for feature binding in striate cortex. *J Neurophysiol* 76: 1356–1360, 1996.

**Gilbert CD, Wiesel TN.** Clustered intrinsic connections in cat visual cortex. *J Neurosci* 3: 1116–1133, 1983.

**Gilbert CD, Wiesel TN.** The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision Res* 30: 1689–1701, 1990.

**Grunewald A, Bradley DC, Andersen RA.** Neural correlates of structure-from-motion perception in macaque V1 and MT. *J Neurosci* 22: 6195–6207, 2002.

**Hegdé J, Felleman DJ.** How selective are V1 cells for pop-out stimuli? *J Neurosci* 23: 9968–9980, 2003.

**Hubel D, Wiesel T.** Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 195: 215–243, 1968.

**Hupé JM, James AC, Girard P, Bullier J.** Response modulations by static texture surround in area V1 of the macaque monkey do not depend on feedback connections from V2. *J Neurophysiol* 85: 146–163, 2001.

**Ito M, Gilbert CD.** Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron* 22: 593–604, 1999.

**Jonides J, Yantis S.** Uniqueness of abrupt visual onset in capturing attention. *Percept Psychophys* 43: 346–354, 1988.

**Kapadia MK, Westheimer G, Gilbert CD.** Dynamics of spatial summation in primary visual cortex of alert monkeys. *Proc Natl Acad Sci USA* 21: 12073–12078, 1999.

**Kastner S, Nothdurft HC, Pigarev IN.** Neuronal correlates of pop-out in cat striate cortex. *Vision Res* 37: 371–376, 1997.

**Knierim JJ, Van Essen DC.** Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J Neurophysiol* 67: 961–980, 1992.

**Lamme VAF.** The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* 15: 1605–1615, 1995.

**Lee TS, Mumford D, Romero R, Lamme VAF.** The role of primary visual cortex in higher level vision. *Vision Res* 38: 2429–2454, 1998.

**Lee TS, Nguyen M.** Dynamics of subjective contour formation in early visual cortex. *Proc Natl Acad Sci USA* 98: 1907–1911, 2001.

**Lee TS, Yang CF, Romero RD, Mumford D.** Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. *Nat Neurosci* 5: 589–597, 2002.

**Levitt JB, Lund JS.** The spatial extent over which neurons in macaque striate cortex pool visual signals. *Vis Neurosci* 19: 439–452, 2002.

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

**Marcus DS, Van Essen DC.** Scene segmentation and attention in primate cortical areas V1 and V2. *J Neurophysiol* 88: 2648–2658, 2002.

**McAdams CJ, Reid RC.** Attention modulates the responses of simple cells in monkey primary visual cortex. *J Neurosci* 25: 11023–11033, 2005.

**Motter BC.** Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J Neurophysiol* 70: 909–919, 1993.

**Nelson JI, Frost BJ.** Orientation-selective inhibition from beyond the classical visual receptive field. *Brain Res* 139: 359–365, 1978.

**Nienborg H, Cumming BC.** Macaque V2 neurons, but not V1 neurons, show choice-related activity. *J Neurosci* 26: 9567–9578, 2006.

**Nothdurft HC, Gallant JL, Van Essen DC.** Response modulation by texture surround in primate area V1: correlates of "popout" under anesthesia. *Vis Neurosci* 16: 15–34, 1999.

**Palmer C, Cheng SY, Seidemann E.** Linking neuronal and behavioral performance in a reaction-time visual detection task. *J Neurosci* 27: 8122–8137, 2007.

**Parker AJ, Krug K, Cumming BG.** Neuronal activity and its links with the perception of multi-stable figures. *Philos Trans R Soc Lond B Biol Sci* 357: 1053–1062, 2002.

**Posner MI.** Orienting of attention. *Q J Exp Psychol* 32: 3–25, 1980.

**Potetz B, Lee TS.** Statistical correlations between 2d images and 3d structures in natural scenes. *J Opt Soc Am A* 20: 1292–1303, 2003.

**Purushothaman G, Bradley DC.** Neural population code for fine perceptual decisions in area MT. *Nat Neurosci* 8: 99–106, 2005.

**Ramachandran VS.** Perception of shape from shading. *Nature* 331: 163–166, 1988.

**Reynolds JH, Pasternak T, Desimone R.** Attention increases sensitivity of V4 neurons. *Neuron* 26: 703–714, 2000.

**Robinson DA.** A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Trans Biomed Eng* 10: 137–145, 1963.

**Roelfsema PR, Khayat PS, Spekreijse H.** Subtask sequencing in the primary visual cortex. *Proc Natl Acad Sci USA* 100: 5467–5472, 2003.

**Roelfsema PR, Lamme VA, Spekreijse H.** Object-based attention in the primary visual cortex of the macaque monkey. *Nature* 395: 376–381, 1998.

**Rossi AF, Desimone R, Ungerleider LG.** Contextual modulation in primary visual cortex of macaques. *J Neurosci* 21: 1698–1709, 2001.

**Sceniak MP, Ringach DL, Hawken MJ, Shapley R.** Contrast's effect on spatial summation by macaque V1 neurons. *Nat Neurosci* 2: 733–739, 1999.

**Shadlen MN, Newsome WT.** Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol* 86: 1916–1936, 2001.

**Smith MA, Bair W, Movshon JA.** Dynamics of suppression in macaque primary visual cortex. *J Neurosci* 26: 4826–4834, 2006.

**Sun J, Perona P.** Early computation of shape and reflectance in the visual system. *Nature* 379: 165–168, 1996a.

**Sun J, Perona P.** Preattentive perception of elementary three-dimensional shapes. *Vision Res* 16: 2515–2529, 1996b.

**Sun J, Perona P.** Where is the sun? *Nat Neurosci* 1: 183–184, 1998.

**Treisman A.** Preattentive processing in vision. *Comput Vis Graphics Image Process* 31: 156–177, 1985.

**Treisman A, Gelade G.** A feature-integration theory of attention. *Cognit Psychol* 12: 97–136, 1980.

**Uka K, DeAngelis GC.** Contribution of middle temporal area to coarse depth discrimination: comparison of neuronal and psychophysical sensitivity. *J Neurosci* 23: 3515–3530, 2003.

**Uka K, DeAngelis GC.** Contribution of area MT to stereoscopic depth perception: choice-related response modulations reflect task strategy. *Neuron* 42: 297–310, 2004.

**Webb BS, Dhruv NT, Solomon SG, Tailby C, Lennie P.** Early and late mechanisms of surround suppression in striate cortex of macaque. *J Neurosci* 25: 11666–11675, 2005.

**Xing D, Shapley RM, Hawken MJ, Ringach DL.** The effect of stimulus size on the dynamics of orientation selectivity in macaque V1. *J Neurophysiol* 94: 799–812, 2005.

**Yantis S, Jonides J.** Abrupt visual onsets and selective attention: evidence from visual search. *J Exp Psychol Hum Percept Perform* 10: 601–621, 1984.

**Zipser K, Lamme VA, Schiller PH.** Contextual modulation in primary visual cortex. *J Neurosci* 16: 7376–7389, 1996.