

Visual Responses in FEF, Unlike V1, Primarily Reflect When the Visual Context Renders a Receptive Field Salient

Wilsaan M. Joiner,^{1,2} James Cavanaugh,¹ Robert H. Wurtz,¹ and  Bruce G. Cumming¹

¹Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892, and ²Department of Bioengineering, George Mason University, Fairfax, Virginia 22030

When light falls within a neuronal visual receptive field (RF) the resulting activity is referred to as the visual response. Recent work suggests this activity is in response to both the visual stimulation and the abrupt appearance, or salience, of the presentation. Here we present a novel method for distinguishing the two, based on the timing of random and nonrandom presentations. We examined these contributions in frontal eye field (FEF; $N = 51$) and as a comparison, an early stage in the primary visual cortex (V1; $N = 15$) of male monkeys (*Macaca mulatta*). An array of identical stimuli was presented within and outside the neuronal RF while we manipulated salience by varying the time between stimulus presentations. We hypothesized that the rapid presentation would reduce salience (the sudden appearance within the visual field) of a stimulus at any one location, and thus decrease responses driven by salience in the RF. We found that when the interstimulus interval decreased from 500 to 16 ms there was an approximate 79% reduction in the FEF response compared with an estimated 17% decrease in V1. This reduction in FEF response for rapid presentation was evident even when the random sequence preceding a stimulus did not stimulate the RF for 500 ms. The time course of these response changes in FEF suggest that salience is represented much earlier (< 100 ms following stimulus onset) than previously estimated. Our results suggest that the contribution of salience dominates at higher levels of the visual system.

Key words: context; FEF; receptive field; salience; V1; visual response

Significance Statement

The neuronal responses in early visual processing [e.g., primary visual cortex (V1)] reflect primarily the retinal stimulus. Processing in higher visual areas is modulated by a combination of the visual stimulation and contextual factors, such as salience, but identifying these components separately has been difficult. Here we quantified these contributions at a late stage of visual processing [frontal eye field (FEF)] and as a comparison, an early stage in V1. Our results suggest that as visual information continues through higher levels of processing the neural responses are no longer driven primarily by the visual stimulus in the receptive field, but by the broader context that stimulus defines—very different from current views about visual signals in FEF.

Introduction

Stimuli falling within the visual receptive field (RF) of a neuron, elicit a response that depends on many factors including attention, inhibitory surrounds, distractors, and salience (Motter, 1993; Ito and Gilbert, 1999; McAdams and Maunsell, 1999; Mc-

Adams and Reid, 2005; Joiner et al., 2011; Cavanaugh et al., 2012). Here we use the term salience to mean the extent to which a localized visual stimulus is distinctive (Itti and Koch, 2001; Fecteau and Munoz, 2006). In this usage, salience is defined entirely by the visual input. Nonetheless, the salience of any local feature depends critically on its global context (Itti and Koch, 2001); a bright spot is more salient when it is presented alone than when it is one of many similar spots. Other factors that contribute to salience (e.g., contrast), can be computed locally, within a single RF. Thus, the effect of salience upon visual responses is a complex mixture of local and contextual information. Finally, salient stimuli involuntarily engage attentional mechanisms that in turn influence neuronal responses (Goldberg et al., 2006; Ipata et al., 2006; Arcizet et al., 2011). Thus, even though it is useful to restrict the term salience to image-defined properties, it is difficult to isolate the effects of salience on neuronal processing. One commonly used approach uses arrays of features in which one

Received May 16, 2017; revised Aug. 25, 2017; accepted Aug. 30, 2017.

Author contributions: W.M.J., J.C., R.H.W., and B.G.C. designed research; W.M.J. performed research; W.M.J., J.C., and B.G.C. analyzed data; W.M.J., R.H.W., and B.G.C. wrote the paper.

This work was supported by the National Eye Institute Intramural Research Program at the National Institutes of Health and a Grant from the National Eye Institute (R01 EY021252) to W.M.J. We thank Altah Nichols and Tom Ruffner for machine shop support.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Wilsaan M. Joiner, Department of Bioengineering, George Mason University, Nguyen Engineering Building, 1G5, 4400 University Drive, Room 3800, Fairfax, VA 22030. E-mail: wjoiner2@gmu.edu.

DOI:10.1523/JNEUROSCI.1446-17.2017

Copyright © 2017 the authors 0270-6474/17/379871-09\$15.00/0

“target” differs from the remaining “distractors” in one feature, such as color. This has been a powerful and influential approach, but difficult to use for comparing different cortical areas, because these differ in which feature dimensions are most important. Partly for this reason, it has not been possible to compare directly the effect of saliency in early and late visual areas, using equivalent measures.

Here we describe a new way to estimate the effect of saliency on visual responses, in which the context is defined by time, rather than by spatial layout. We compared the response of neurons to identical visual stimuli in one of two contexts. In the first context, each stimulus was more salient because they appeared between long intervals without visual stimulation [an interstimulus interval (ISI) of 500 ms]. In the second context, identical stimuli were less individually salient because they appeared in rapid succession with a short ISI of only 16 ms. We presented stimuli both within and outside the RF of each neuron, so we can distinguish the global saliency (the sudden appearance of a stimulus anywhere) from the local saliency (abrupt appearance of a stimulus within the RF). Thus, with the short ISI the whole array of flashing stimuli was salient, but saliency was not restricted to the RF. Importantly, we gathered sufficient data with short ISIs that we could identify epochs where the random sequence produced no RF stimulation for 500 ms preceding a 100 ms presentation in the RF. This produced a local history of stimulation (for 600 ms) that was identical in the two contexts, yet the temporal context rendered the stimuli more salient in the long ISI condition. Because our two contexts are defined by the sequence of visual stimuli, it is necessarily possible to differentiate them with an appropriate bottom-up computation. As there is no agreed definition of how saliency is computed, it might be argued that the important difference is not saliency. However, the computation required is sufficiently complex that it fits naturally into most definitions of saliency, so we shall use the label saliency to differentiate the two contexts below. We therefore exploit the fact that this protocol can be applied to early visual areas to perform the first direct comparison of the role of saliency in frontal eye fields (FEFs) and the striate cortex.

Materials and Methods

Four adult male monkeys (*Macaca mulatta*) weighing from 7 to 9 kg, were implanted with scleral search coils for measuring eye position, recording cylinders for accessing neurons in FEF or primary visual cortex (V1), and a post for immobilizing the head during experiments as described previously (Sommer and Wurtz, 2000). All procedures were approved by the Institute Animal Care and Use Committee and complied with Public Health Service Policy on the humane care and use of laboratory animals.

Recording and behavioral procedures for FEF. We implanted recording cylinders over the FEF (the anterior bank of the arcuate sulcus) approximately normal to the cranial surface. Cylinder placement over the sulcus was guided by MR images. We recorded single neuron responses and microstimulated in FEF with tungsten microelectrodes advanced by a stepper microdrive. FEF neurons were verified using two criteria: saccade-related activity and the ability to evoke saccades with stimulation currents $<50 \mu\text{A}$ (Bruce et al., 1985).

The monkey sat in a primate chair with its eyes 57 cm in front of a tangent screen. The chair was in the center of magnetic field coils in a dark room that was sound attenuated. Computers running REX (Hays et al., 1984) and associated programs controlled stimulus presentation, administration of reward, the recording of eye movements and single neuron activity, and the on-line display of results. We excluded neurons from analysis if we were unable to determine the parameters of the visual RF due to loss of neuronal isolation during the recording session.

Visual stimuli on the screen appeared on a gray background, back-projected by a DPI projector. In each experiment we first determined the center of the RF, by creating a coarse spatial map. While the monkey fixated a central white cross, we sequentially presented spots of light of a fixed diameter (1–5°, depending on RF eccentricity) at nine locations on a 3×3 grid. The exact time of stimulus onset and offset was determined by a photocell on the projector side of the screen that registered the appearance of a white square in the corner of the screen (hidden from the monkey) for a single monitor cycle that was synchronized to both the appearance and disappearance of each stimulus. The stimulus grid center and spacing were adjusted to sample a large part of the RF and form an initial quantitative estimate of RF center and qualitative estimate of the extent. From the 3×3 grid, we determined the RF center as the average of the neuronal responses at nine locations weighted by the magnitude of the visual responses at each location (Joiner et al., 2011; Cavanaugh et al., 2012).

Recording and behavioral procedures for V1. The recordings from V1 were made in two animals. In one animal a 10×10 array of silicon probes (“Utah” array, Blackrock Microsystems) had been surgically implanted on the operculum. In the second animal, a linear electrode array (“V-probe”, Plexon) was introduced transdurally in each recording session. The signal from each electrode was sampled at 30 kHz and saved to disk. Custom software was used to identify well isolated single units. The animal sat in a primate chair and each eye viewed a CRT monitor (Viewsonic P225f, framerate 85 Hz) through a Wheatstone stereoscope. In these experiments both monitors always showed identical images. Stimuli were presented with custom software using OpenGL. Eye position was monitored and recorded (using Spike2, Cambridge Electronic Design), and stimuli were only presented while the animal maintained fixation on a central spot. The stimuli were bars of 99% contrast presented against a gray background.

Saliency experiments. There were seven probe locations for the experiments conducted in FEF (16 for V1 experiments). For experiments in FEF the stimuli were small white spots of fixed diameter (1–5°, depending on RF eccentricity). The orientation (horizontal or vertical) of the spot array and the spacing between the spots were adjusted for each neuron so that spot locations were both within and outside the neuron’s RF. For V1, the stimuli were bars of fixed height and width (3° and 0.13°, respectively). As with the spots, the bar locations were chosen to ensure that stimuli were located both within and outside the RF.

We used exactly the same stimuli in two blocks of trials with the only difference in the trials being the timing of the stimuli. A block consisted of ~50 trials for FEF experiments, and 240 for V1 experiments. Throughout each trial a single spot or bar was presented at one of the stimulus locations for 100 ms, with a new random location chosen for each presentation. Importantly, each stimulus location had an equal probability of being selected. Therefore, within a trial it was possible to find a sequence of presentations where multiple stimulus presentations were outside the RF or that the same stimulus location was presented consecutive times.

To examine the role of saliency on the visual activity, the timing between stimulus presentations (the ISI) was different in the two blocks of trials. In one block the ISI between presentations was a long duration (500 ms) and in the other block (the short ISI condition) the ISI was the duration of a single frame, 16 ms (12 ms for V1). For the short ISI condition a given stimulus presentation could also be a blank screen (no stimulus) that was the same duration (100 ms) and equally probable as any visual stimulus presentation. This was critical for the short ISI condition because it allowed the possibility to have a sequence of blank presentations preceding the presentation of a visual stimulus, thus allowing a direct comparison of the visual stimulation elicited in the long ISI condition. (Note that in some experiments we made the blank screen twice as probable as any stimulus presentation to increase the likelihood of a sequence of blank presentations preceding the presentation of a visual stimulus.)

For FEF experiments, a single trial for the long ISI condition included eight visual stimulus presentations and the trial duration was ~5 s. A single trial for the short ISI condition in FEF included 60 presentations (a random combination of spot presentations or blank screens) and the

trial duration was ~ 7 s. For V1 experiments, a long ISI trial included four visual stimulus presentations and the trial duration was ~ 2.4 s. A single short ISI trial in V1 included 21 presentations (a random combination of bar presentations or blank screens) and the trial duration was ~ 2.1 s.

On every trial the monkey received a liquid reward for maintaining fixation within a 1.5° (1° for V1) window around the fixation point for the duration of the trial. Trials in which the eye left this window were discarded because any eye movement would displace the location of the visual probes on the retina.

Analysis of RF activity. We searched for neurons that had an identifiable visual RF and recorded from a total of 51 visually-sensitive FEF neurons in two monkeys (25 in monkey FlN and 26 in monkey Cap) and V1 neurons in an additional two monkeys (5 in monkey Jbe and 10 in monkey Lem). We recorded until we acquired a certain number of trials or when the response isolation became uncertain. We saw no significant difference between the monkeys and have combined their results. We used results from all FEF neurons that had a clear visual response, whether or not there was accompanying saccade-related activity.

For each condition we examined the neuronal activity from 100 ms before to 300 ms after the stimulus presentation and determined the spike density function across all presentations of a given stimulus location. We also determined the mean spike count within a 100 ms window starting at response onset.

The locations of stimulus presentations for the short ISI condition were drawn randomly from a uniform distribution (see above). We computed mean spike density functions following the presentation of any stimulus by convolving the spike train with a half Gaussian ($\sigma = 20$ ms), because these are causal filters. We also computed mean spike counts for conditions defined by a specific sequence of stimuli. For example, we could determine the mean spike count for the presentation of a stimulus in the center of the RF that followed a 500 ms period during which no stimuli were presented in the RF.

Latencies were estimated by fitting a bilinear function to the mean spike density function for all stimuli within the RF, after subtracting the spontaneous rate.

$$R(t) = [0, t < \textit{latency}; k(t - \textit{latency}), t > \textit{latency}]$$

The fit was constrained to have a value of 0 before the latency, and had a linear slope, k , after the latency. Slope and latency were the only parameters of the fit. Separation time (see Fig. 5C) was estimated with the same procedure, applied to the difference between the mean spike density functions for the two ISI contexts.

Statistical analysis. The spike counts determined for both ISI conditions were dependent variables. Therefore, the correlation of the spike counts across probe locations was accomplished with a type II regression (Armitage et al., 2001). Significances and p values, unless otherwise specified, were determined by paired two-tailed t tests and repeated-measures ANOVA. For all tests the significance level was 0.05.

Results

Effect of ISI on responses in FEF

For each neuron we estimated the RF location and then presented stimuli differing only in position as depicted at the top of Figure 1. Seven locations were uniformly spaced along a row that sliced through the center of the RF, positioned so that some stimuli were within the RF, and some were outside the RF. We also included blank presentations with no stimulus at all. During a single period of fixation, stimuli were flashed for 100 ms sequentially with either a 500 or 16 ms ISI. These presentations were at random locations with equal probability of occurrence. Figure 1 shows the results for an example FEF neuron. The spike density functions for each stimulus location, distinguished by the respective colors, are shown in Figure 1A for the long ISI condition. When the identical probes were presented for the same duration, but at a short ISI (Fig. 1B) the responses were greatly reduced in magnitude compared with the long ISI condition. Importantly, for both presentation conditions the different probe locations

produced visual responses that decrease with distance from the RF center and approach zero, which demonstrates that the probes spanned the RF. Figure 1C plots the average firing rate (determined within a 100 ms window starting at response onset (the vertical dashed black lines in A and B) as a function of probe location for both conditions (producing a map of the RF along one dimension). The response magnitudes were quite different in the ISI conditions, but the shape of the RF was similar. For example, in both cases there is a peak at the -7° probe location, with a similar decrease in the response for adjacent locations. Thus, although the magnitude of the largest responses decreased as the ISI was shortened, the shape of the RF was approximately constant.

In Figure 1D we quantified the change in visual response magnitude produced by changing the ISI. We plot the spike count (determined within the same 100 ms window) produced in the short ISI condition against the count produced by the same stimulus with a long ISI. On this plot, if the neuron's discharge were entirely determined by the local stimulus properties (which we call "visual activity"), the ISI would not matter and both the slope and the correlation would be one. If the saliency entirely controlled the neuron's response, only the long ISI would produce a response, so both values would be 0. For the example FEF neuron there was a significant linear relationship between the spike counts (Pearson's $r = 0.85$, $p < 0.05$) with a slope (type II regression) of 0.13 (gray trace), indicating a substantial influence of the saliency. The large correlation suggests that the RF is still clearly defined, its shape is unchanged, and that there was a consistent change in the visual response magnitude across the stimulus locations. The slope estimates the fraction of visual activity remaining with the decrease in ISI from 500 to 16 ms. Thus, the slope and correlation for the example FEF neuron suggests that 13% of the neuronal response resulted from visual activity and the remaining 87% was the effect of the saliency evident only with a long ISI between stimuli.

We performed the same type II regression analysis on a sample of 51 FEF neurons. Figure 2 plots the correlation against the regression slope for each neuron. Many cells show high correlations, but slopes < 1 , like the example cell (filled black circle). Notice that there are also many cells with lower correlations. In these cases the responses to the short ISI presentations are so weak that the RF is poorly defined. Because of this, low correlations are always associated with low slopes. If the response of FEF neurons were more closely related to saliency we would expect low slopes even in neurons with high correlation, exactly as Figure 2 shows.

In Figure 2 the median slope for the sample of FEF neurons was 0.21 (mean 0.22) and the median correlation was 0.85 (mean 0.75), represented in the plot by the unfilled red symbol. These results show that $\sim 21\%$ of the visual response remained when the ISI was reduced across the sample of FEF neurons, indicating that the largest part of the response ($\sim 79\%$) is attributable to saliency.

The role of early visual responses

One possible explanation for this phenomenon is that for some reason (perhaps the allocation of spatial attention), this manipulation of temporal context alters the responses of the afferent visual input to FEF. To control for this, we performed the same experiment while recording a small number (15) of V1 neurons. In this case the stimuli were oriented bars (close to the preferred orientation). Figure 3 shows the results for an example V1 neuron (layout similar to Fig. 1). The different probe locations produced responses that decreased as the stimuli fell farther from the center

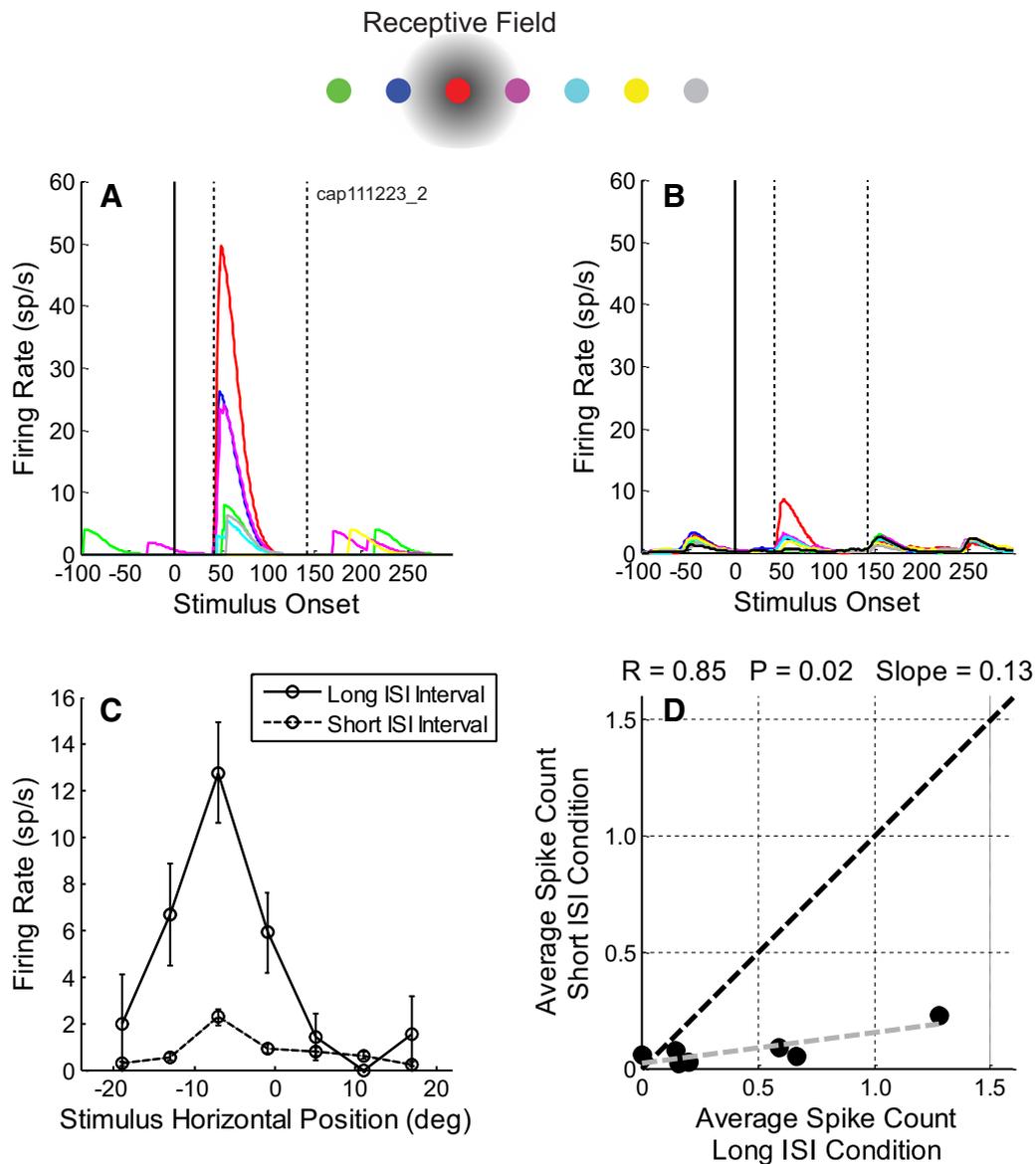


Figure 1. Visual RF responses and the effect of saliency for an example FEF neuron. Visual responses for the (A) long and (B) short ISI conditions for the example neuron. Each of the seven probe locations is depicted in the diagram at the top and represented by a respective color. The total number of presentations for each probe location ranged between 10 and 22 for the long ISI condition and 276–315 for the short ISI condition. The average firing rate and spike count was determined within a 100 ms window (represented by the dashed vertical black lines) beginning at response onset. C, The average firing rate for the example neuron in A is plotted as a function of probe location to create a coarse map of the RF for both ISI conditions. D, The average spike count (determined over the same 100 ms window) for the short ISI condition is plotted as a function of the spike count for the long ISI condition for each probe location for the example neuron. There was a significant linear relationship between the spike counts ($r = 0.85$, $p = 0.02$) revealing an estimated 87% decrease (slope of 0.13) in the response magnitude when the time between stimulus presentations was shortened.

of the RF for both the long and short ISIs (Fig. 3A,B). Similar to the example FEF cell, the shape of the RF was similar (Fig. 3C). However, unlike in FEF, the responses of the V1 cell decreased only slightly at the short ISI (Fig. 3D); there was a strong linear relationship between the spike counts (Pearson's $r = 0.99$, $p < 0.001$) with a slope of 0.96, indicating that the majority of the neuronal response reflects the local visual stimulus.

In Figure 2, the gray symbols summarize the data for V1 for comparison with FEF. The median slope for the sample of V1 neurons was 0.83 (mean 0.85) and the median correlation was 0.97 (mean 0.96), represented in the plot by the unfilled blue circle. These results show that for the sample of V1 cells ~83% of the visual response remained when the ISI was reduced across the sample of V1 neurons, indicating that the largest part of the response is attributable to local visual stimulation.

Saliency and adaptation

An alternative explanation for the decreased response at the short ISI is that rapid stimulation produces local visual adaptation. This possibility is particularly relevant in light of a recent study by Mayo and Sommer (2008) showing that the FEF visual response to the second of two successive stimulus presentations at the same location changed as a function of the time between presentations, the ISI. They found that when the ISI was large (>400 ms) the magnitude of the second response was approximately the same as when the stimulus was presented alone. However, as the ISI decreased to 16 ms, the magnitude of the second response decreased to ~10% of the response to the lone stimulus. Mayo and Sommer (2008) attributed this reduction to neuronal adaptation. The effect of saliency we demonstrate above raises a second possible explanation for this effect. Because Mayo and Sommer (2008)

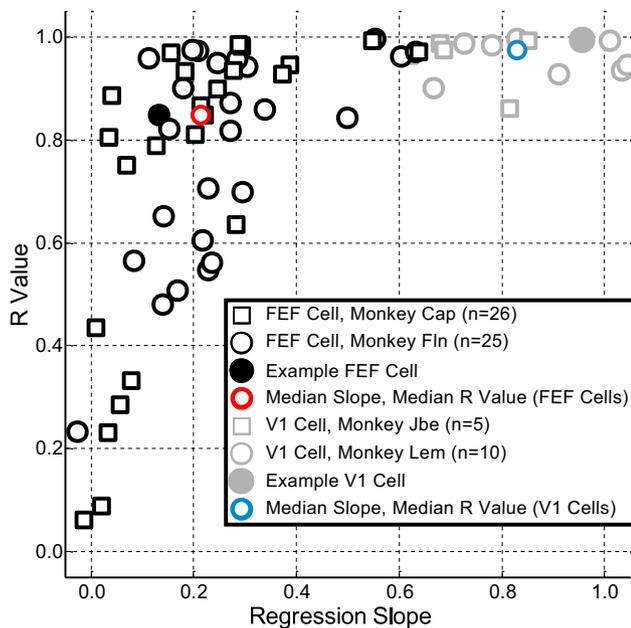


Figure 2. Reduction in response magnitude with ISI for the sample of FEF and V1 neurons. The Pearson correlation is plotted against the slope from the regression analysis of spike counts from the two ISI conditions (as described in Figs. 1*D*, 3*D*) for the 51 FEF (unfilled black symbols) and 15 V1 neurons (unfilled gray symbols). There were 25 FEF neurons from monkey Fln (circles), 26 from monkey Cap (squares), and a total of 15 V1 neurons from monkeys Jbe (5) and Lem (10). The median r value for the sample of FEF neurons was 0.85 (mean 0.75) and the median slope was 0.21 (mean 0.22), signified by the unfilled red symbol. The median r value for the sample of V1 neurons was 0.97 (mean 0.96) and the median slope was 0.83 (mean 0.85), indicated by the unfilled blue symbol. The example cells described in Figures 1 and 3 are represented by the filled black and gray circles, respectively.

presented stimuli against a blank screen, the first presented stimulus was most salient, and its presentation reduces the saliency of the second stimulus (which was presented at the same location). Thus, changes in saliency alone could explain their result, although this does not exclude local visual adaptation. Our random stimulation sequence with short ISIs provides a unique opportunity to differentiate these two explanations. Because there is nothing distinctive about the appearance of any one stimulus in the sequence, they are all equally salient. Nonetheless, we can identify sequences in our random presentations which produced a period of 100–500 ms during which there was no RF stimulation, followed by a stimulus in the RF. This then matches the local stimulation conditions of the long ISI used by Mayo and Sommer (2008), and should not produce adaptation. Comparing responses to these different sequences allows us to estimate the effects of local visual adaptation, without any difference in saliency.

We performed this comparison for responses to the probe location in the center of the RF to look at the responses that had the largest changes. We identified responses that were preceded by a period with no RF stimulation lasting 100–500 ms, the latter matching the 500 ms between stimulus presentations in the large ISI condition. In Figure 4 we plot the magnitude of this response as a percentage of the response observed with the large ISI. This period without RF stimulation could be any combination of one to five consecutive presentations outside the RF or a blank screen. Because only a small fraction of presentations meet the conditions for 500 ms without stimulation, only a subset of FEF neurons ($n = 28$) provided sufficient data for this test. As shown in Figure 4, the response for the FEF neurons gradually increased from 31 to 45% as the preceding period of no RF stimulation

increased from 0 to 500 ms. The response when preceded by 500 ms of no RF stimulation was significantly greater than when preceded by 0 ms (paired two-tailed t test, $t_{(27)} = 3.36$, $p = 0.002$). We checked that the results were similar if we used only the blank condition to define no RF stimulation; i.e., we require that no stimulus was on the screen for a period of 100–300 ms. These results are shown in unfilled black circles in Figure 4 (for a smaller subset of neurons that provided sufficient data for this test), and are not significantly different from the results in which we allow stimuli outside the RF to be included in the nonstimulated condition (two-way ANOVA, $F_{(2,105)} = 0.88$, $p = 0.42$ for the main effect of time duration; $F_{(1,105)} = 0.39$, $p = 0.54$ for the main effect of nonstimulation type; $F_{(1,88)} = 0.7$, $p = 0.93$ for the interaction). In both cases, there is a gradual increase in the response as the duration with no preceding RF stimulation increases. This change was small compared with the difference between short ISI and long ISI conditions; even following 500 ms of no RF stimulation, the visual response in FEF only reached 45% of the response at the large ISI. That is, when the local stimulation history is matched between the two conditions, the response for the short ISI condition was only 45% of the response observed during the large ISI condition for the same RF stimulus. The only difference between these conditions is the global context established by the stimulus timing. In the short ISI condition stimuli are constantly being flashed every 100 ms, and are therefore less salient than the long ISI condition where flashes are occurring 500 ms apart. This context alone has a profound effect on the visual responses of neurons in FEF.

As a control analysis, we also examined any response interference from presenting objects after the RF visual stimuli (backward masking). To test for this we specifically compared the FEF RF response when the visual stimulation was followed by blank to the RF response when followed by a nonblank stimulus. The geometric mean of the ratio of these two responses was 1.04, and a paired t test showed no significant difference between the two conditions ($p = 0.52$), demonstrating that backward masking cannot explain the low responses we see with short ISIs.

In contrast to FEF, the same analysis for the V1 cells demonstrated that the response for the two ISI conditions was substantially closer, consistent with the results presented in Figure 2. In this case the percentage of the response for the short compared with the long ISI was >87%, even when not preceded by a period of no RF stimulation (0 ms). Unlike FEF, the response did not increase significantly when preceded by 500 ms without RF stimulation, compared with a 0 ms gap (paired two-tailed t test, $t_{(14)} = 2.16$, $p = 0.05$). In addition, preceding blank screens (unfilled gray circle) produced very similar results to preceding stimuli outside the RF (filled gray circles), with no significant difference for only blank screens (two-way ANOVA, $F_{(2,83)} = 0.83$, $p = 0.44$ for the main effect of time duration; $F_{(1,83)} = 1.67$, $p = 0.20$ for the main effect of nonstimulation type; $F_{(2,83)} = 0.12$, $p = 0.89$ for the interaction). Finally, a period of 300 ms without RF stimulation is sufficient for responses to recover to the same value as in the long ISI condition. In summary, these results show that (1) there is modest adaptation in FEF with the rapid presentation of the stimuli during the short ISI condition (an ~14% decrease in the response), (2) a larger fraction of this response reduction (~55%) is due to the decreased saliency of the presentation, and (3) in V1 there is similar local adaptation, but no evidence for an effect of saliency.

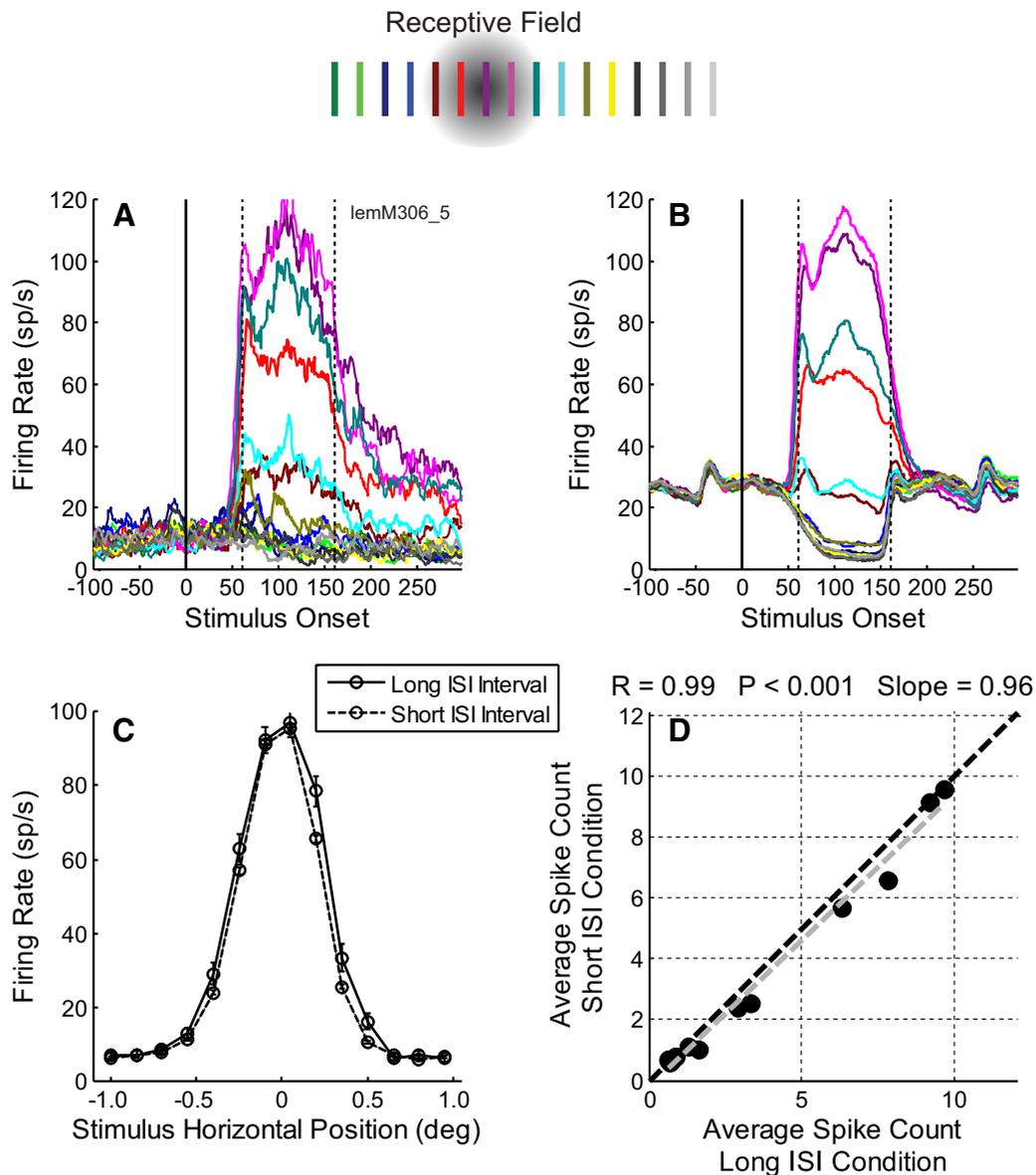


Figure 3. Visual RF responses and the effect of saliency for an example V1 neuron. Layout is the same as Figure 1. Visual responses for the (A) long and (B) short ISI conditions for the example neuron. The 16 probe locations with respect to the RF center are depicted in the diagram at the top and represented by a respective color. The total number of presentations for each probe location ranged between 97 and 134 for the long ISI condition and 2132–2186 for the short ISI condition. The average firing rate and spike count was determined within a 100 ms window (represented by the dashed vertical black lines) beginning at response onset. C, The average firing rate for the example neuron in A is plotted as a function of probe location to create a coarse map of the RF for both ISI conditions. D, The average spike count (determined over the same 100 ms window) for the short ISI condition is plotted as a function of the spike count for the long ISI condition for each probe location for the example neuron. There was a significant linear relationship between the spike counts ($r = 0.99$, $p < 0.001$) revealing an estimated 4% decrease (slope = 0.96) in the response magnitude when the interval between stimulus presentations was shortened.

Time course of saliency contribution to neuronal response

Because we used temporal context to control saliency, we are able to compare responses to salient and non-salient stimuli that are spatially identical. This provides a unique opportunity to examine the temporal evolution of responses to saliency. That is, at what point after stimulus presentation does the saliency modulate the normal response to the visual stimulation. We first determined the population average response for the long ISI condition in the sample of FEF cells (Fig. 5A, black trace). The colored traces represent the average response for different durations of no stimulus in the RF for the short ISI condition. Consistent with Figure 4, the magnitude of the response increases with this duration, but does not reach the magnitude of the long ISI response. Note that as the unstimulated duration increased, fewer cells contributed to

the average response (i.e., fewer cells had sufficient trials to study at these specific sequences of no RF stimulation preceding the stimulus presentation in the RF). Thus, as a control we also determined the average response for the long ISI condition for cells ($n = 28$) that were included in the analysis of 500 ms of no RF stimulation (Fig. 5A, gray trace). Note, these are the same 28 cells depicted in Figure 4. The direct comparison for this subset of cells (gray trace compared with purple trace) shows that for the same preceding context (500 ms of no RF stimulation) the short ISI condition still had a reduced response. Figure 5B plots these average response within the black dashed box in A (between 30 and 60 ms). Note the average response at the long ISI (black and gray traces) separates from the short ISI conditions (colored traces) between 40 and 60 ms. In Figure 5C, we quantify this separation

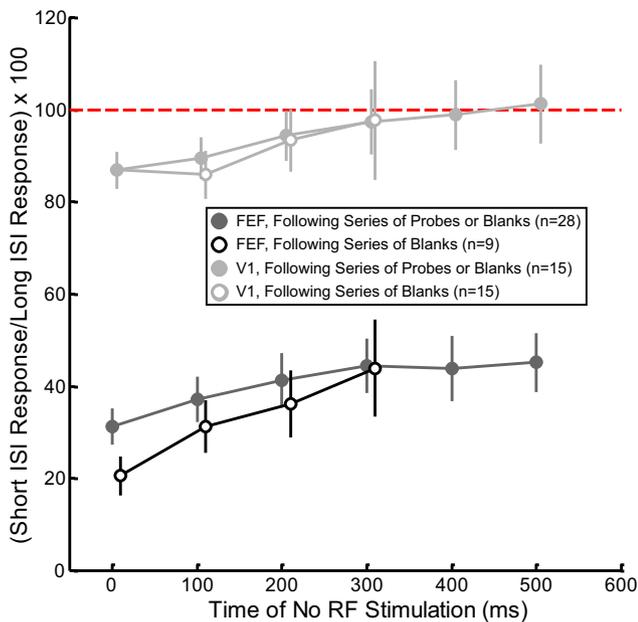


Figure 4. Changes in the RF response for different periods of no RF stimulation. The percentage of the response magnitude (short ISI compared with long) for a stimulus in the center of the RF for FEF and V1 neurons (black and gray symbols, respectively). This percentage is plotted for different durations of no preceding RF stimulation. This period consisted of a combination of sequential presentations of a blank screen or probes outside the RF (filled circles) or only blank screens (unfilled circles). For FEF, 28 cells were used for the no RF stimulation analysis, and only nine cells provided sufficient data for the comparison using only blank screens. A larger group (15) was available for the V1 comparisons.

time for each neuron, and plot it against the response latency. The unfilled blue symbols (circles; $n = 43$) represent 300 ms of no stimulus in the RF. The unfilled purple symbols (squares; $n = 28$) represent 500 ms of no stimulus in the RF. The large black outlined circle and square represent the mean latency and time of separation across all cells for the 300 (43.1 ± 2.0 and 56.0 ± 2.3 ms) and 500 ms of no stimulus in the RF (42.7 ± 2.1 and 59.4 ± 2.8 ms), respectively. Most cells demonstrate a separation of the response very soon after response onset; in both duration cases the majority of the symbols lie above the unity line. Note also that there is very little change in the time of separation between the two durations of no stimulation, and that the average time of the separation (56.0 ± 2.3 and 59.4 ± 2.8 ms) occurs well before stimulus offset at 100 ms.

Discussion

We quantify the extent to which the global context contributing to saliency accounts for the visual response of neurons in FEF. These global factors were isolated by comparing neural responses to stimuli flashed for 100 ms in two different contexts: long (500 ms) and short (16 ms) ISIs. With long ISIs, each stimulus presentation is a substantial change in the global layout; a uniform blank screen suddenly changes to an isolated target. With short ISIs each individual stimulus presentation is much less salient, because a similar stimulus was present somewhere on the screen beforehand. Visual responses in FEF were substantially attenuated by the short ISIs. Thus, the visual response in FEF neurons largely reflects the saliency that results from the global context. It may be that the whole array of rapidly flashing stimuli is salient, but this renders a large region of the visual field outside the RF salient. That the RF no longer represents the most salient locus is sufficient to greatly reduce responses. This in turn implies

that the visual response depends upon integrating information from far outside what is typically considered the “receptive field”. In a control group of neurons recorded from area V1, we show a very different pattern. Although the local stimulation history >500 ms influences response magnitudes, the global context has no impact.

The time course of the responses we observe in FEF suggest that saliency is represented much earlier than in previous studies where spatial context defined saliency, for example by comparing responses to targets versus distractors in a search array. In these studies the initial response is similar for targets and distractors, with a difference emerging only 100 ms or more after stimulus onset (Murthy et al., 2001; Schall, 2002). Because the initial visual response does not differentiate targets from distractors, it was naturally thought that the initial response simply reflects the local stimulation within the RF. Our manipulation of saliency produces a very different pattern, with response differences that are evident only 10 ms after the initial visual response. This implies that even the earliest visual responses in FEF do not simply reflect local stimulation within the RF, but reflect saliency defined by global context. This suggests that the delay seen in spatial studies reflects the time required for the visual system to compute the saliency of different targets after a new image is presented. Our finding that the earliest part of the visual response reflects saliency profoundly alters the interpretation of visual responses in FEF. For example, recent studies have investigated the neural mechanisms that facilitate the stability of vision across the saccadic eye movements used to sample the environment (Sommer and Wurtz, 2006; Zirnsak et al., 2014; Cavanaugh et al., 2016). There are changes in the RF accompanying these saccades in gaze control neural areas, such as FEF. In light of the current results, it is possible that these changes in the RF are due to saliency rather than the result of local visual processing. Combining the methods we present here with goal directed saccadic eye movements will aid characterizing the predictive neural mechanisms that stabilize vision during these frequent disruptions to visual input.

In line with a number of recent authors, we differentiate saliency (a property only of the visual image) from “priority” (which encompasses both saliency and effects such as spatial attention that depend upon behavioral relevance). Salient stimuli engage attentional mechanisms involuntarily (Goldberg et al., 2006; Ipata et al., 2006; Arcizet et al., 2011), and so the modulation of FEF responses probably reflects a combination of these two. For example, in the long ISI condition there is ample time for spatial attention to shift between the briefly presented stimuli, whereas this shift cannot match the frequency of scene changes during the short ISI condition (Zhang et al., 2011). That we see no effect of saliency in V1 suggests that these stimuli do not engage the form of exogenous attention that influences FEF (Buffalo et al., 2010; Wang et al., 2015). However, because the effects of attention in V1 are modest, it is hard to exclude them with confidence. For FEF, one possibility is that the target selection response seen in search tasks (after 100 ms) primarily reflects the attentional component, whereas the early effects of saliency we demonstrate reflect purely bottom-up processes. Our demonstration that single flashes (500 ms ISI) are more effective than flashes within a random sequence in which there happened to be no preceding stimulation for 500 ms (Fig. 4) emphasizes that even these early effects require complex processing. It cannot be explained simply by suggesting that FEF has a longer integration time (or latency), because this would produce profound changes as the period with no stimulation changes from 16 to 500 ms. In both cases we examine the neuronal response to a single flash

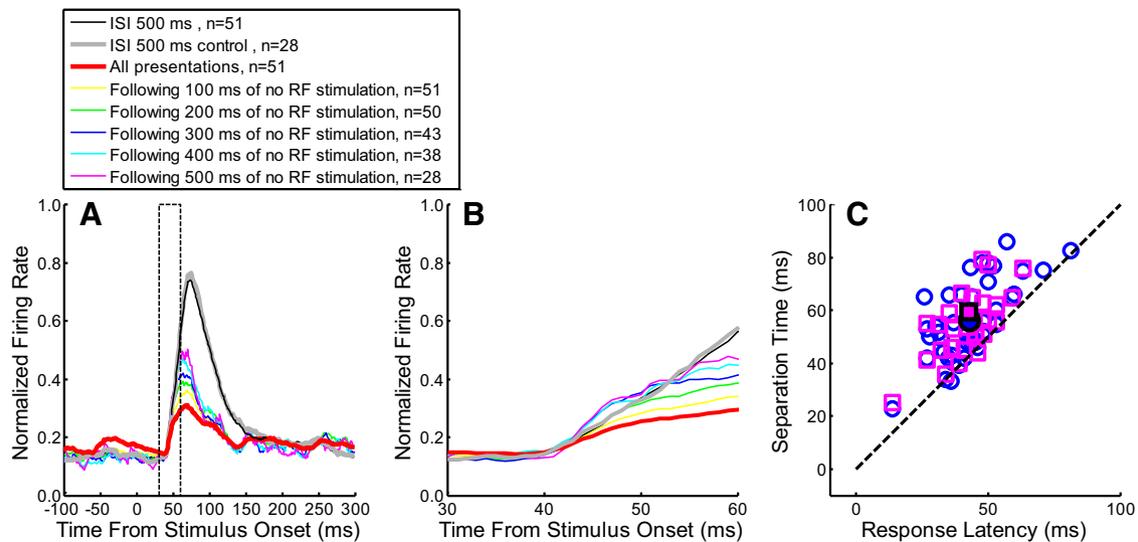


Figure 5. Timing of saliency contribution for the sample for FEF neurons. **A**, The average visual response at the long ISI condition for all FEF cells ($n = 51$) is represented by the black trace. The colored traces represent the average response for different durations of no stimulus in the RF for the short ISI condition. The thick gray trace represents a control analysis that shows the response for the long ISI condition for the 28 cells that had a duration of no RF stimulation that matched the long ISI (purple trace). **B**, The average responses within the black dashed box in **A** (between 30 and 60 ms). Note the rise of the response for the long ISI conditions (black and gray traces) compared with the short ISI conditions (colored traces). **C**, The time of response separation between the long and short ISI conditions as a function of the response latency at the long ISI condition. As in **A**, the unfilled blue symbols (circles; $n = 43$) represent 300 ms of no stimulus in the RF, and the unfilled purple symbols (squares; $n = 28$) represent 500 ms of no stimulus in the RF. The large black outlined circle and square represent the mean latency and time of separation across all cells within each respective group.

preceded by a 500 ms blank. In one context, these flashes always occur after 500 ms without stimulation. In the other context these are rare events in continuous sequence. That they produce such different responses means that information about retinal stimulation >600 ms previously, and far outside the RF, profoundly influences responses in FEF (whereas changes in the 500 ms preceding the flash have little impact). In neurons reflecting such complex properties, the use of a minimum response field to define the area that receives retinal input may not be a helpful simplification. It is even possible that top-down mechanisms contribute even to the early effects of saliency on FEF neurons that we observe; it may be that the enhanced responses are partly the result of exogenous attention being drawn to the location of the flashed object in the long ISI condition.

Visual search studies in both primates (Thompson and Bichot, 2005; Thompson et al., 2005) and humans (Zenon et al., 2010) suggest that FEF contains a map of saliency and priority; a topographical spatial map that represents the activity related to the distribution of visual saliency and behavioral relevance within the environment. This existence of a map for priority is supported by several studies demonstrating that during the search for specific stimuli, the majority of FEF visual activity represents the behavioral relevance of an object rather than specific features (Schall et al., 1995; Thompson et al., 1996, 1997). Our results strengthen this framework, indicating that the quantitative effect of saliency is even larger than previously thought. In other words, saliency greatly increases the sensory signal in FEF, whereas in V1 the visual response is not changed by saliency. This difference may be explained from a computational perspective. For example, the limited capacity of neural processing constrains the amount of visual space that can be managed; higher processing levels should limit their resources to behaviorally relevant stimuli—a distinction that may be facilitated by the saliency. Neural mechanisms in FEF might select and focus processing preferentially toward salient information. This increased sensitivity to salient stimuli may allow precise encoding of information from

this spatial location, enabling efficient processing of the potentially most relevant information in the sensory input.

Our protocol also provides a method that differentiates saliency from local visual adaptation. Typical adaptation paradigms present an adaptor followed by a probe. Changes in responses to the probe might then be caused either by local adaptation, or by the fact that the adaptor renders the probe less salient. In our context with short ISIs, the saliency of each presentation is nearly constant. Nonetheless, because the sequence of locations is random, it is possible to compare the effect of repeated presentations of nearby stimuli with cases where a stimulus is presented at a location that has not been stimulated for several hundred milliseconds. This comparison then quantifies the effect of local adaptation without a contribution from saliency. Strikingly, FEF neurons show only modest adaptation under these circumstances (Fig. 4). Comparison with a recent study of adaptation in FEF neurons (Mayo and Sommer, 2008) highlights the importance of saliency. They used a traditional adapting stimulus (which was itself salient), and found that this substantially reduced responses to a subsequent probe. When we present an almost identical sequence of local stimulation, in a context where adaptor and probe are equally salient, this effect is greatly attenuated. This raises the possibility that saliency, rather than local visual adaptation, is primarily responsible for the changes seen by Mayo and Sommer (2008). To some extent this distinction is semantic; the phenomenon they report is certainly a form of adaptation. But the conclusion that the adaptation is an effect of saliency rather than reflecting local visual processing changes the possible set of neural substrates. Visual processing that integrates over long time periods (>600 ms) might explain some of the effects we describe as saliency, because our long ISI and short ISI conditions (compared in Fig. 4) do differ in the visual stimulation >600 ms before the response we report. This difference might drive a form of adaptation, although it would be surprising that it has so little effect over the 500 ms we examine in Figure 4. It is also noteworthy that our data from V1 show little impact of stimulus history

over this timescale. Thus, any such adaptation is not present in the afferent input, and depends on integrating information in complex ways over large spatial and temporal windows, in a way that identifies the saliency of stimuli defined by the temporal context we used. This certainly requires a more complex form of adaptation than that seen in earlier visual areas.

Saliency has been studied in various visual areas (V1: Smith et al., 2007; Jingling and Zhaoping, 2008; lateral intraparietal cortex: Gottlieb et al., 1998; Kusunoki et al., 2000; Ipata et al., 2006; Arcizet et al., 2011; Foley et al., 2014; shifting RFs in FEF: Joiner et al., 2011; pulvinar: Robinson and Petersen, 1992). These studies have all used spatial context to produce saliency, presenting several stimuli, one of which is made distinctive by some feature (e.g., color or shape). Comparing results across these visual areas is difficult, because they differ in their selectivity for visual features. Because the method we employ here uses only timing to control saliency, it can be used with any spatial feature. This will allow it to be used in a wide range of visual areas, with (spatially) optimal stimuli. The profound difference we show between V1 and FEF suggests this technique applied to other cortical areas offers important insights into the neural substrate of saliency.

References

- Arcizet F, Mirpour K, Bisley JW (2011) A pure saliency response in posterior parietal cortex. *Cereb Cortex* 21:2498–2506. [CrossRef Medline](#)
- Armitage P, Berry G, Matthews JNS (2001) *Statistical methods in medical research*. Oxford, UK: Wiley-Blackwell.
- Bruce CJ, Goldberg ME, Bushnell MC, Stanton GB (1985) Primate frontal eye fields: II. Physiological and anatomical correlates of electrically evoked eye movements. *J Neurophysiol* 54:714–734. [Medline](#)
- Buffalo EA, Fries P, Landman R, Liang H, Desimone R (2010) A backward progression of attentional effects in the ventral stream. *Proc Natl Acad Sci U S A* 107:361–365. [CrossRef Medline](#)
- Cavanaugh J, Joiner WM, Wurtz RH (2012) Suppressive surrounds of receptive fields in monkey frontal eye field. *J Neurosci* 32:12284–12293. [CrossRef Medline](#)
- Cavanaugh J, Berman RA, Joiner WM, Wurtz RH (2016) Saccadic corollary discharge underlies stable visual perception. *J Neurosci* 36:31–42. [CrossRef Medline](#)
- Fecteau JH, Munoz DP (2006) Saliency, relevance, and firing: a priority map for target selection. *Trends Cogn Sci* 10:382–390. [CrossRef Medline](#)
- Foley NC, Jangraw DC, Peck C, Gottlieb J (2014) Novelty enhances visual saliency independently of reward in the parietal lobe. *J Neurosci* 34:7947–7957. [CrossRef Medline](#)
- Goldberg ME, Bisley JW, Powell KD, Gottlieb J (2006) Saccades, saliency and attention: the role of the lateral intraparietal area in visual behavior. *Prog Brain Res* 155:157–175. [CrossRef Medline](#)
- Gottlieb JP, Kusunoki M, Goldberg ME (1998) The representation of visual saliency in monkey parietal cortex. *Nature* 391:481–484. [CrossRef Medline](#)
- Hays AV Jr, Richmond BJ, Optican LM (1984) A UNIX-based multiple process system for real-time data acquisition and control. *WESCON Conference Proceedings* 2:1–10.
- Ipata AE, Gee AL, Gottlieb J, Bisley JW, Goldberg ME (2006) LIP responses to a popout stimulus are reduced if it is overtly ignored. *Nat Neurosci* 9:1071–1076. [CrossRef Medline](#)
- Ito M, Gilbert CD (1999) Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron* 22:593–604. [CrossRef Medline](#)
- Itti L, Koch C (2001) Computational modelling of visual attention. *Nat Rev Neurosci* 2:194–203. [CrossRef Medline](#)
- Jingling L, Zhaoping L (2008) Change detection is easier at texture border bars when they are parallel to the border: evidence for V1 mechanisms of bottom-up saliency. *Perception* 37:197–206. [CrossRef Medline](#)
- Joiner WM, Cavanaugh J, Wurtz RH (2011) Modulation of shifting receptive field activity in frontal eye field by visual saliency. *J Neurophysiol* 106:1179–1190. [CrossRef Medline](#)
- Kusunoki M, Gottlieb J, Goldberg ME (2000) The lateral intraparietal area as a saliency map: the representation of abrupt onset, stimulus motion, and task relevance. *Vision Res* 40:1459–1468. [CrossRef Medline](#)
- Mayo JP, Sommer MA (2008) Neuronal adaptation caused by sequential visual stimulation in the frontal eye field. *J Neurophysiol* 100:1923–1935. [CrossRef Medline](#)
- McAdams CJ, Maunsell JH (1999) Effects of attention on orientation tuning functions of single neurons in macaque cortical area V4. *J Neurosci* 19:431–441. [Medline](#)
- McAdams CJ, Reid RC (2005) Attention modulates the responses of simple cells in monkey primary visual cortex. *J Neurosci* 25:11023–11033. [CrossRef Medline](#)
- Motter BC (1993) 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. [Medline](#)
- Murthy A, Thompson KG, Schall JD (2001) Dynamic dissociation of visual selection from saccade programming in frontal eye field. *J Neurophysiol* 86:2634–2637. [Medline](#)
- Robinson DL, Petersen SE (1992) The pulvinar and visual saliency. *Trends Neurosci* 15:127–132. [CrossRef Medline](#)
- Schall JD (2002) The neural selection and control of saccades by the frontal eye field. *Philos Trans R Soc Lond B Biol Sci* 357:1073–1082. [CrossRef Medline](#)
- Schall JD, Morel A, King DJ, Bullier J (1995) Topography of visual cortical afferents to frontal eye field in macaque: convergence and segregation of processing streams. *J Neurosci* 15:4464–4487. [Medline](#)
- Smith MA, Kelly RC, Lee TS (2007) Dynamics of response to perceptual pop-out stimuli in macaque V1. *J Neurophysiol* 98:3436–3449. [CrossRef Medline](#)
- Sommer MA, Wurtz RH (2000) Composition and topographic organization of signals sent from the frontal eye field to the superior colliculus. *J Neurophysiol* 83:1979–2001. [Medline](#)
- Sommer MA, Wurtz RH (2006) Influence of the thalamus on spatial visual processing in frontal cortex. *Nature* 444:374–377. [CrossRef Medline](#)
- Thompson KG, Bichot NP (2005) A visual saliency map in the primate frontal eye field. *Prog Brain Res* 147:251–262. [CrossRef Medline](#)
- Thompson KG, Hanes DP, Bichot NP, Schall JD (1996) Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. *J Neurophysiol* 76:4040–4055. [Medline](#)
- Thompson KG, Bichot NP, Schall JD (1997) Dissociation of target selection from saccade planning in macaque frontal eye field. *J Neurophysiol* 77:1046–1050. [Medline](#)
- Thompson KG, Bichot NP, Sato TR (2005) Frontal eye field activity before visual search errors reveals the integration of bottom-up and top-down saliency. *J Neurophysiol* 93:337–351. [CrossRef Medline](#)
- Wang F, Chen M, Yan Y, Zhaoping L, Li W (2015) Modulation of neuronal responses by exogenous attention in macaque primary visual cortex. *J Neurosci* 35:13419–13429. [CrossRef Medline](#)
- Zenon A, Filali N, Duhamel JR, Olivier E (2010) Saliency representation in the parietal and frontal cortex. *J Cogn Neurosci* 22:918–930. [CrossRef Medline](#)
- Zhang Y, Meyers EM, Bichot NP, Serre T, Poggio TA, Desimone R (2011) Object decoding with attention in inferior temporal cortex. *Proc Natl Acad Sci U S A* 108:8850–8855. [CrossRef Medline](#)
- Zirnsak M, Steinmetz NA, Noudoost B, Xu KZ, Moore T (2014) Visual space is compressed in prefrontal cortex before eye movements. *Nature* 507:504–507. [CrossRef Medline](#)