Superior colliculus encodes visual saliency during smooth pursuit eye movements

Authors:
Brian J. White[1], Laurent Itti[2], Douglas P. Munoz[1]

Affiliations:
[1]Centre for Neuroscience Studies, Queen's University, Kingston, ON, Canada.
[2]Department of Computer Science, University of Southern California, Los Angeles, CA, USA.

Correspondence to:
Brian J. White
Email: brian.white@queensu.ca

Running title: Saliency coding during smooth pursuit

Keywords: saliency map, priority map, visual coding, visuospatial updating, popout, visuomotor

Abbreviations: SC: superior colliculus; SCs: superior colliculus superficial layers; SCi: superior colliculus intermediate layers; FEF: frontal eye field; LIP: lateral intraparietal area; V1: primary visual cortex; MT: Middle temporal area; MST: medial superior temporal area; RF: receptive field; LCD: liquid crystal display; REX: Real-time experimentation system; DKL color space: Derrington–Krauskopf–Lennie color space; SD: standard deviation; SEM: standard error of the mean.

Manuscript details:
Word count: 7768
Abstract: 175
Intro: 853
Materials and methods: 1388
Results: 1613
Discussion: 1393
Captions: 820
Figures: 4 main, 2 supplementary
References: 58

ABSTRACT

The saliency map has played a longstanding role in models and theories of visual attention, and it is now supported by neurobiological evidence from several cortical and subcortical brain areas. While visual saliency is computed during moments of active fixation, it is not known whether the same is true while engaged in smooth pursuit of a moving stimulus, which is very common in real-world vision. Here, we examined extrafoveal saliency coding in the superior colliculus, a midbrain area associated with attention and gaze, during smooth pursuit eye movements. We found that SC neurons from the superficial visual layers showed a robust representation of peripheral saliency evoked by a conspicuous stimulus embedded in a wide-field array of goal-irrelevant stimuli. In contrast, visuomotor neurons from the intermediate saccade-related layers showed a poor saliency representation, even though most of these neurons were visually responsive during smooth pursuit. These results confirm and extend previous findings that place the SCs in a unique role as a saliency map that monitors peripheral vision during foveation of stationary and now moving objects.

INTRODUCTION

Saliency map theory postulates the existence of a neural map that encodes the visual conspicuity of stimuli/objects across the visual field based on low-level features such as color and motion (**Fig. 1a**, red) (Itti *et al.*, 1998; Itti & Koch, 2001; Borji & Itti, 2013). In contrast, the term priority map has been used to describe a combined representation of visual saliency and behavioral relevancy (**Fig. 1a**, blue), and is theorized to be the core determinant of attention and gaze (Fecteau & Munoz, 2006; Serences & Yantis, 2006). Neural correlates of saliency and/or priority maps have been reported across a network of mostly cortical brain areas (e.g., primary visual cortex, V1 (Li, 2002; Li *et al.*, 2006; Zhang *et al.*, 2012; Yan *et al.*, 2018); V4 (Burrows & Moore, 2009); lateral intraparietal area, LIP (Gottlieb *et al.*, 1998; Bisley & Goldberg, 2010); frontal eye field, FEF (Thompson & Bichot, 2004; Purcell *et al.*, 2012)). However, recent research has shown that the midbrain superior colliculus (SC), which has long been associated with attention (Goldberg & Wurtz, 1972; Krauzlis, *et al.* 2013) and gaze (Gandhi & Katnani, 2011; White & Munoz, 2011a), also plays an important role in an early stage of saliency coding (Veale *et al.*, 2017; White, Berg, *et al.*, 2017; White, Kan, *et al.*, 2017). Specifically, during free viewing of dynamic natural scenes, the response of superior colliculus superficial layer neurons (SCs), whose dominant inputs arise from the retina and V1 (Lock *et al.*, 2003; Cerkevich *et al.*, 2014), was predicted by a computational saliency model (White, Berg, *et al.*, 2017) that has been validated on the free viewing behavior of humans (Itti, 2005) and rhesus monkeys (Berg *et al.*, 2009). In addition, the SCs was shown to signal saliency earlier than V1 (White, Kan, *et al.*, 2017), the dominant gateway to the visual system. In agreement with these results, it has also been shown that gaze patterns during free viewing remain correlated with model predicted

saliency in the absence of V1 (Yoshida *et al.*, 2012), further implicating the SCs saliency map and the retinotectal pathway when geniculo-striate inputs are disrupted.

[Fig. 1 here]

Because the fixation and smooth pursuit systems share a common function and underling neural substrates (Krauzlis, 2003), we asked whether the saliency map operates in a similar manner during foveation of stationary or moving objects. This is important because smooth pursuit is an essential part of natural gaze behavior and is common during free viewing of dynamic natural scenes (White, Berg, *et al.*, 2017). Furthermore, while much is known about visual processing during fixation, extrafoveal processing during smooth pursuit is certainly less understood. Behavioral studies have examined aspects of visual perception (Schütz *et al.*, 2008; Braun *et al.*, 2017) and attention (Lovejoy *et al.*, 2009; Khan *et al.*, 2010; Chen *et al.*, 2017) during smooth pursuit, as well as contextual visual effects on pursuit metrics (Spering & Gegenfurtner, 2007; Kreyenmeier *et al.*, 2017). A few studies have also examined how visually responsive neurons encode extrafoveal stimuli during pursuit (Erickson & Thier, 1991; Ilg, 1996; Inaba *et al.*, 2007; Chukoskie & Movshon, 2009; Dash *et al.*, 2015, 2016), but most of these focused on how motion sensitive areas (e.g., Middle temporal area, MT, and medial superior temporal area, MST) process pursuit-induced retinal motion. For these reasons, understanding saliency coding during pursuit is important, not only for computation saliency models that do not currently distinguish between foveation of stationary or moving objects, but also for a broader understanding of the visual processes that operate during natural gaze behavior.

One might predict an attenuation of salient extrafoveal stimuli during pursuit given its reliance on foveal inputs (Khijrana & Kowler, 1987; Kerzel *et al.*, 2008), and the fact that attention tends to be focused around the pursuit target (Lovejoy *et al.*, 2009; Khan *et al.*, 2010;

Chen *et al.*, 2017). Such a result might be at odds with a pure bottom-up saliency map, but would be in agreement with a behavioral priority map, proposed to exist in SCi. Another possibility is that because smooth pursuit produces retinal motion in the opposite direction of the movement (Erickson & Thier, 1991; Ilg, 1996; Inaba *et al.*, 2007; Chukoskie & Movshon, 2009), this self-induced global motion might dominate local saliency signals. It is therefore conceivable that peripheral saliency may be attenuated during pursuit. Alternatively, it would seem advantageous for the brain to monitor peripheral saliency during pursuit, and we know that SC neurons show visual-related responses to isolated peripheral stimuli (Dash *et al.*, 2015, 2016). Moreover, recent studies support the role of SCs in saliency coding during active fixation in complex and dynamic natural scenes (White, Berg, *et al.*, 2017; White, Kan, *et al.*, 2017). Based on that research, we postulated that the SCs would show an extrafoveal saliency representation during smooth pursuit, whereas the SCi was predicted to show an attenuated saliency response to such goal-irrelevant stimuli, given the long-standing role of SCi in the control of top-down attention (Goldberg & Wurtz, 1972; Krauzlis *et al.*, 2013), and goal-directed target selection (McPeek & Keller, 2002; White & Munoz, 2011b; Shen & Paré, 2014).

MATERIALS AND METHODS

Animal preparation

Data were collected from two male Rhesus monkeys (*Macaca mulatta*; monkey I and monkey U, 11-12kg each). Surgical procedures and extracellular recording techniques have been detailed previously (Marino *et al.*, 2008). All animal care and experimental procedures were approved by the Queen's University Animal Care Committee in accordance with the guidelines of the Canadian Council on Animal Care.

Equipment

Visual stimuli were presented on a high-definition LCD video monitor (Sony Bravia 55", Model KDL-46XBR6) at a screen resolution of 1920 x 1080 pixels (60Hz noninterlaced, 16 bit color depth). Viewing distance was 70cm resulting in a viewing angle of 82° horizontally and 52° vertically. The viewing area that extended beyond the monitor was blackened using black non-reflective cloth.

The tasks were controlled by a Dell 8100 computer running a UNIX-based real-time data control system (REX 7.6 (Hays *et al.*, 1982)), which communicated with a second computer running in-house graphics software (written in C/C++) for presentation of stimuli. Stimulus timing was controlled using a photodiode placed at the left lower corner of the monitor and hidden by non-reflective tape. The photodiode measured the onset of a stimulus (20 x 20 pixels) that pulsed for one frame simultaneously with the onset of the main stimuli (i.e. the photodiode stimulus turned white for one frame then returned to black). The real-time control system (REX) was synchronized to the timing of the photodiode pulse by holding the current state until the pulse was detected.

Eye position was monitored using a 1000 Hz video-based eye tracker (*Eyelink 1000, SR research*). The data were recorded by a third computer running a multi-channel data-acquisition system (*Plexon Inc., Dallas, Texas, USA*). Spike waveforms were sampled at 40 kHz. Eye position, event data, and spike times were digitized at 1 kHz.

Procedure, Stimuli and Task

Monkeys were seated in a primate chair (*Crist Instr., MD, USA*) approximately 70 cm from the LCD video, head restrained. The animals were first trained to perform the smooth pursuit eye movement task in the absence of any other stimuli besides the pursuit target (**Fig. 1b**), which generally moved along a mostly vertical trajectory, but depended upon the receptive field (RF)

location once actual neuronal recording began. The animals were not previously trained to perform visual selection type tasks using arrays of orientated color stimuli. Later in the training phase we began to increase the visibility (contrast) of the goal-irrelevant stimuli associated with the current task (**Fig. 1**). As such, the animals learned early on to disregard the goal-irrelevant stimuli, and to facilitate this, trials were instantly and automatically aborted if gaze slipped from the invisible computer controlled window (~3ºx3º) surrounding the pursuit target. The training took approximately two to three weeks.

During the main experiment, single glass-insulated tungsten microelectrodes (2.0MΩ; *Alpha Omega, GE, USA*) were lowered into the SC through a stainless steel guide tube. The animals viewed a dynamic video, which provided dynamic visual stimulation that facilitated the localization of the visually-responsive dorsal SC surface. In most cases when a neuron was isolated, its visual RF was mapped using a rapid visual stimulation procedure described previously (Marino *et al.*, 2012). If no clear RF emerged during this procedure, we continued to search for another cell. Only visually active neurons were included. Typically, following the mapping procedure we would then run a delayed-saccade task to functionally classify each neuron as visual-only SCs or visuomotor SCi, based on previously established methods (Marino *et al.*, 2012). The RF centers (see **Fig. 2f**) of our sample of SC neurons varied and were not always along the horizontal meridian, which meant that pursuit directions were not strictly vertical, but could often be diagonal and in some cases horizontal.

The main experimental stimuli consisted of a radial arrangement of equally-spaced color bars (210 items) spanning 40°-45° visual angle (**Fig. 1c-f**). The items were horizontally or vertically oriented (typically 0.4°x1.2°, but ranged from 0.3°x0.8° to 0.6°x1.6° for the nearest to furthest RF eccentricities, respectively), and were red or green derived from the red-green cardinal axis in

Derrington–Krauskopf–Lennie (DKL) color space (Derrington *et al.*, 1984), with negative 40% luminance contrast relative to the neutral grey background (65cd/m^2). The main condition consisted of the array with a single oddball whose color and orientation was distinct from the remaining homogenous items (e.g., red horizontally oriented oddball against green vertically oriented background; **Fig. 1c-d**). All combinations of the four stimulus features (red, green, horizontal, vertical) were used such that the oddball could be red or green, horizontal or vertical, and the remaining items consisted of the opposing color and orientation feature. The oddball could appear in (**Fig. 1c**) or opposite (**Fig. 1d**) the RF. This was compared to a singleton control condition in which a single red or green, horizontal or vertical, stimulus appeared in (**Fig. 1e**) or opposite (**Fig. 1f**) the RF. The array/singleton remained stationary and did not move with the pursuit target. All experimental conditions were randomly interleaved.

With respect to the sequence of events (**Fig. 1b, c-f**), first a peripheral fixation point (FP; black Gaussian-windowed spot, SD=0.3°) appeared above or below center, at the same eccentricity as the RF, ±90° radial angle relative to the RF. The animals fixated the FP for a 0.5s-0.7s random period, after which the goal-irrelevant array/singleton appeared. The animals continued fixating the FP for an additional 0.5s-0.7s, after which the FP stepped 1.5º in the opposite direction of center screen, then moved at a constant speed (15º/s) towards then past center screen to the opposite visual field location at an eccentricity equal to the start position (**Fig. 1b**; Rashbass step-ramp procedure to reduce saccades during pursuit initiation (Rashbass, 1961)). Thus, the length of the pursuit target trajectory was twice the eccentricity of the RF center, and consequently the duration of the pursuit trajectory was not constant, but varied with RF eccentricity. The animals were required to smoothly track the moving stimulus within an invisible moving ~3°x3° computer controlled window, which matched the pursuit target speed. If

gaze fell outside this window, the trial was immediately aborted and all stimuli disappeared from the screen within 50ms. This helped to eliminate most saccades that occurred during pursuit. In total, 3.7% (232/6323) of trials were automatically aborted due to break from the computer controlled window after the array/singleton appeared but before the trial was successfully completed (see **supplementary Fig. S1**). Less than 0.8% of these were directed towards the singleton or oddball (defined as trials in which the Euclidean distance between that saccade end point and the singleton/oddball was less than half that eccentricity). After the stimulus reached the end of the movement trajectory, it stopped and the animals continued fixating the target for an additional 200ms after which it disappeared and a liquid reward was given. Visually evoked responses were measured by aligning to the point in the movement trajectory when the pursuit stimulus was at center screen, which corresponded with the time the RF was aligned with the salient oddball/singleton.

Single units were isolated online using a window discriminator, and confirmed offline using spike sorting software (Plexon Inc., Dallas, Texas, USA). Spikes were convolved with a Gaussian function (SD=10ms). A total of 55 neurons were isolated, and 4 neurons were excluded because they did not yield visual activity during the RF mapping procedure or during the delayed-saccade task, leaving a total of 51 (32 visual SCs [16 each from monkey I and U], 19 visuomotor SCi [11 and 8 from monkey I and U, respectively).

Velocity traces were smoothed using a 20 point moving average, and any trials with unsigned velocity (i.e., speed) exceeding 60^{o}/s during the test epoch (±200ms around the time the pursuit target crossed center screen) were removed offline. In total, 8.8% (541/6091) of the trials were removed, leaving between 5 to 44 trials per condition across all recordings. The majority of catch-up saccades are forward, and while some backward catch-up saccades might be missed by

our simple speed threshold approach, this is a rare occurrence. Averaging and statistical analyses were performed within this 400ms epoch because it represents a time period during which the salient oddball/singleton was within the RF border for all neurons. This allowed us to adequately address the main hypotheses without the removal of too many trials due to minor catch-up saccades at other points in the pursuit trajectory.

RESULTS

SCs neurons signal saliency during smooth pursuit eye movements

To examine saliency coding during smooth pursuit, we used a task designed to measure the difference in visually-evoked activation between salient and non-salient items across a wide-field array (White, Kan, *et al.*, 2017) (**Fig. 1;** see Materials and Methods for details). The array consisted of oriented color bars with a salient oddball (**Fig. 1c-d**). A singleton control condition (**Fig. 1e-f**) was included for comparison of visual responses with no surrounding context. On a given trial, the animals smoothly tracked the pursuit target whose trajectory was calculated to move along a trajectory that was 90^{o} relative to the radial direction of the RF, such that the RF was drawn over the salient oddball (**Fig. 1c-d**) or singleton (**Fig. 1e-f**). This allowed us to differentiate a 1st order saliency response (evoked by a local luminance difference between a singleton and the background) from a higher order saliency response, which relies on feature contrasts between salient and non-salient items across the visual field. We compared visually-evoked responses when the goal-irrelevant oddball/singleton appeared *in* versus *opposite* the RF, representing high versus low saliency regions of the display, the difference of which was taken as a measure of saliency coding.

Figure 2a-b shows mean vertical eye position as a function of time from two recording sessions. Average pursuit gain ranged from 0.61 to 1.05 across recordings, but did not differ

significantly between the singleton (black traces) and array (colored traces) conditions (Z_{42} = 0.09, P = 0.93, *Wilcoxon signed rank test* across n=43 sessions; **Fig. 2c**). The pursuit stimulus could move in either direction along a trajectory that was 90º relative to the radial direction of the RF. It should be noted that RFs were not always on the horizontal meridian (see **Fig. 2f** for distribution of RFs), resulting in pursuit directions that were sometimes diagonal or even horizontal, but pursuit was usually mostly upward and downward (**Fig. 2a-b**, upper and lower panels, respectively), and these trials were randomly interleaved. The results were qualitatively similar for either pursuit direction so the data were collapsed in terms of pursuit direction to maximize statistical power. A portion of the total trials (541/7268 or 7.4%) were removed because eye velocity exceeded a saccade threshold criterion of 60º/s during the epoch denoted by the grey vertical shading in **Fig. 2** (see Materials and Methods). Averaging and statistical analyses were performed within this 400ms epoch because it represents a time period during which the salient oddball/singleton was within the RF border across all neurons, and allowed us to adequately address the main hypotheses without the removal of too many trials due to catch-up saccades at other points in the pursuit trajectory.

Figure 2d-e show single cell examples (red, SCs visual only neuron; blue, SCi visuomotor neuron). The traces were aligned on the time the pursuit stimulus crossed the screen center, which coincided with the time the RF was spatially aligned with the oddball/singleton. Note, the results were qualitatively similar, though not as clean, when aligned on vertical eye position relative to center (see supplementary Fig. S2). For both example neurons (**Fig. 2d, e**), there was a gradual increase then decrease in activation as the RF was gradually brought over the singleton (black thick traces), which was absent when the singleton appeared opposite the RF (black thin traces). This difference was highly significant during the test epoch for both neurons (Z_{19} =

59.01, P = 3.70e-08 for SCs, Z_{18} =34.80 P = 4.82e-09 for SCi, *Wilcoxon Ranksum test, Bonferroni corrected*), which indicates that both neurons were visually activated during smooth pursuit. For the example SCs neuron (**Fig. 2d**), there was also an increase in activation associated with the oddball (red thick trace), which was attenuated when the oddball appeared opposite the RF (red thin trace), and this difference was also significant during the test epoch (Z_{18} = 39.38, P = .00034). For the example SCi neuron (**Fig. 2e**), there was a less noticeable distinction between when the oddball appeared in (blue thick trace) versus opposite (blue thin trace) the RF, though for this neuron the difference was statistically significant (Z_{24} = 6.78, P < .05).

[Fig. 2 here]

Figure 3a, d shows population averaged spike density functions for our sample of 32 SCs neurons and 19 SCi neurons. The tick marks immediately above the x-axis in **Figure 3a, d** indicate the results of a moving statistical test between the oddball/singleton IN versus OPP conditions (P < 0.05, df = 31 and df = 18 for SCs and SCi, respectively, *Wilcoxon signed rank test* in a moving 50ms window at 50ms intervals, *Bonferroni–Holm corrected*). The population averaged results indicate that only SCs neurons reliably encoded the saliency oddball, even though both sets of neurons showed clear visually evoked responses when a unitary singleton appeared in the RF.

[Fig. 3 here]

Figures 3b, c, e, f show a summary and break down across all neurons, averaged within the test epoch (represented by the grey vertical shading in **Fig. 3a, d**). Approximately 69% (22/32) of SCs neurons (**Fig. 3b**, black filled symbols) and 63% (12/19) of SCi neurons (**Fig. 3e**, black filled symbols) showed a greater response when the singleton fell in versus opposite the RF, indicating that these neurons were visually activated during smooth pursuit (P < .05, df = 4 to df

= 43 across the sample of neurons, *Wilcoxon Ranksum test*). The remaining neurons (~31-37%) were visually responsive to an abrupt onset in their RF during fixation in a visual delay task, and this discrepancy is likely due to the fact that SC neurons are more sensitive to transient onsets (Boehnke & Munoz, 2008), than the gradual input induced by smooth pursuit. **Figure 3g** shows that there was no significant difference in the size of this effect for SCs and SCi neurons (Z_{49} = 0.49, P = 0.62, *Wilcoxon Ranksum test*). Critically, in the array condition, approximately 44% (14/32) of SCs neurons showed a significantly greater response when the oddball fell IN versus OPP the RF (**Fig. 3c**, $P < .05$, $df = 4$ to $df = 41$ across the sample of neurons, *Wilcoxon Ranksum test*), indicating that these neurons also encoded the presence of the salient oddball. Although this effect was larger for one animal (monkey U, Fig. 3h, inverted triangles) than the other (monkey I, Fig. 3h, upright triangles), the trend was consistent and statistically significant in both animals. In contrast, 26% (5/19) of individual SCi neurons showed a significantly greater response when the oddball fell IN versus OPP the RF (**Fig. 3f**, $P = .05$, $df = 4$ to $df = 38$ across the sample of neurons, *Wilcoxon Ranksum test*). Importantly, **Figure 3h** shows that the size of the oddball-preference effect for SCs was significantly greater than for SCi ($Z_{49} = 2.42$, $P = 0.01$, *Wilcoxon Ranksum test*). Taken together, these results are qualitatively similar with our previous work (White, Berg, *et al.*, 2017; White, Kan, *et al.*, 2017), and indicate that the SCs also encodes the saliency of extrafoveal stimuli during smooth tracking of a moving object.

SCs and SCi show surround modulation during smooth pursuit

Computational models of visual saliency depend critically on center-surround feature contrasts that operate widely across the visual field (Itti *et al.*, 1998; Itti & Koch, 2001; Borji & Itti, 2013). Similarly, a map that prioritizes spatial locations/objects for the control of attention and gaze also relies on wide-field surround modulation in order to establish a systematic biasing

of one or more locations on that map. Thus, strong surround modulation is a prerequisite for any brain area purported to play a role in the computation of saliency and/or priority. We included the singleton control condition as a benchmark to quantify the effect of the wide-field surround on both SCs and SCi neurons. Based on previous results (White, Kan, *et al.*, 2017), we hypothesized that both SCs and SCi would show significant surround modulation during smooth pursuit. To this end, we compared peak visual responses evoked by the wide-field array (surround; **Fig. 1d**) with a singleton (no surround; **Fig. 1f**) during pursuit. From the averaged population traces (**Fig. 3a, d**), we already observed noticeable response attenuation from the array for both SCs and SCi neurons, comparing the singleton IN (black thick trace) versus oddball IN (red thick trace) conditions. **Figure 4a** and **4b** summarize these differences, by averaging the response over a short epoch (±50ms relative to the zero point) in which the RF was most aligned with the oddball/singleton in order to capture the maximum responses. Approximately 40% of SCs and 50% of SCi neurons showed a significantly attenuated response in the array condition (Oddball IN RF). We calculated the percentage of surround suppression for each neuron type. **Figure 4c** shows that the wide-field array produced significant surround suppression for both SCs neurons (red; 39% response attenuation, Z_{28}=4.46, P = 7.99e-06) and SCi neurons (blue; 32% response attenuation, Z_{18}=2.54, P = .011; *Wilcoxon signed rank test for zero median*). There was no significant difference in the percentage of response attenuation between the SCs and SCi (Z_{49} = 0.22, P = .81, *Wilcoxon Ranksum test*). The magnitude of this response attenuation for SCs and SCi was qualitatively similar to our previous study involving fixation of a stationary stimulus (28-32% for SCs and 15-28% for SCi) (White, Kan, *et al.*, 2017), suggesting that similar mechanisms might operate here during smooth pursuit. These results are certainly in line with the well-established long-range connectivity known to exist

within the SCs and SCi, and have been well documented in previous studies (see for example (Meredith & Ramoa, 1998; Munoz & Istvan, 1998; Phongphanphanee *et al.*, 2014)).

[Fig. 4 here]

DISCUSSION

The saliency map has played a longstanding role in models and theories of visual attention, and it is now well validated in terms of gaze prediction (for an detailed review see (Borji & Itti, 2013)). Moreover, there is now support for something akin to a saliency map in several cortical and subcortical regions of the primate brain (Gottlieb *et al.*, 1998; Li, 2002; Thompson & Bichot, 2004; Li *et al.*, 2006; Burrows & Moore, 2009; Bisley & Goldberg, 2010; Purcell *et al.*, 2012; Zhang *et al.*, 2012; White, Berg, *et al.*, 2017; White, Kan, *et al.*, 2017). In this study, we highlight a novel issue, namely the degree to which the saliency map operates during the foveation of moving objects (i.e., smooth pursuit eye movements). To our knowledge, this is the first study to test this hypothesis directly in a brain area that has been shown to play an important role in early saliency coding (White, Berg, *et al.*, 2017; White, Kan, *et al.*, 2017). We found that about half (14/32) of our sample of SCs neurons encoded the salient extrafoveal oddball while monkeys performed a smooth pursuit task. Although a smaller fraction (5/19) of SCi neurons also encoded the salient oddball, the magnitude of this effect was considerably smaller for SCi than SCs (**Fig. 3h**). Thus, the role of SCi as a saliency map is certainly less evident. This pattern of results is in close agreement with our previous research using similar wide-field stimulus arrays (White, Kan, *et al.*, 2017), and natural dynamic scenes (White, Berg, *et al.*, 2017), to examine saliency coding during periods of fixation. In those studies, SCi also showed a relatively weaker saliency representation than SCs, which led to the proposal that SCi may be more akin to a priority map with strong top down control over goal-irrelevant visual inputs (White, Berg, *et*

al., 2017; White, Kan, *et al.*, 2017) (**Fig. 1a**, blue). In combination with previous research (White, Berg, *et al.*, 2017; White, Kan, *et al.*, 2017), we now have converging evidence across three independent studies that the primate SCs embodies the role of a saliency map, which functions qualitatively similar during foveation of stationary and moving objects.

Although the current study is not a direct test of the role of SCi as a priority map, this idea is not particularly novel or controversial (Fecteau & Munoz, 2006). Earlier studies on the role of SCi in the control of visual attention and target selection point in this direction (Ignashchenkova *et al.*, 2004; Nummela & Krauzlis, 2010; Zénon & Krauzlis, 2012; Krauzlis *et al.*, 2013). With respect to smooth pursuit specifically, one study examined extrafoveal processes in SCi during pursuit (Hafed & Krauzlis, 2008). In that study, the authors reported that SCi neurons initially encode the retinotopic location of extrafoveal stimuli, but during pursuit the active SCi neurons are located at the fovea even though no stimulus is present in the fovea at this point in their task. This is certainly consistent with our interpretation of priority processing in SCi. Future research would benefit from simultaneous recording in the rostral and caudal SCi during a task such as the one used in the current study.

A question that arises is the degree to which divided attention between the pursuit target and the singleton/oddball may have contributed to biasing activation in favor of the latter. There are several reasons why this cannot adequately account for the results, in particular with respect to the critical array condition. First, the animals used in this study were never trained to perform visual selection type tasks using arrays of orientated color stimuli. They were first trained to perform smooth pursuit in the absence of any other stimuli except the pursuit target, and only later in the training phase did we begin to increase the visibility (contrast) of the goal-irrelevant array/singleton. As such, the animals learned early on to disregard the goal-irrelevant stimuli,

and to facilitate this, trials were instantly aborted if gaze slipped from the invisible computer controlled window surrounding the pursuit target (see Materials and Methods). Second, it was largely SCs neurons that showed a saliency preference, yet it is SCi neurons that are most commonly associated with attention modulation (Ignashchenkova *et al.*, 2004; Krauzlis *et al.*, 2013). Third, it has been shown that during smooth pursuit, visual attention is heavily focused on the pursuit target (Lovejoy *et al.*, 2009; Khan *et al.*, 2010; Chen *et al.*, 2017), and the maintenance of smooth pursuit gain requires strong attention allocation on the tracked stimulus (Khijrana & Kowler, 1987; Kerzel *et al.*, 2008). The fact that SCs neurons in particular continued to signal the presence of the salient oddball under these conditions further supports its proposed function. Because the dominant input to SCs arises from early visual cortical areas, and the dominant output is SCi, the structure and function of SCs is ideally suited for a saliency map that monitors peripheral inputs irrespective of the goals of the animal. The strongest case for divided attention might be the singleton condition because there was only one other highly salient stimulus in the visual field besides the pursuit target, and this condition always showed the greatest visually evoked response during pursuit (**Fig. 2, 3**). However, there was no difference in steady-state pursuit gain (an index of divided attention (Khijrana & Kowler, 1987; Kerzel *et al.*, 2008)) between the singleton and array conditions (**Fig. 2c**), indicating that biased attention cannot adequately account the difference between those conditions in which it would have been most likely. Although this does not rule out the possibility that attention played a role in biasing activation in both conditions, it certainly does not explain the difference between conditions, nor the greater saliency response for SCs neurons over SCi neurons. Most importantly, this pattern of results qualitatively mirrors previously published work (White, Berg, *et al.*, 2017; White, Kan, *et al.*, 2017), and therefore likely represents a similar process. For these

reasons, we think these results are consistent with a genuine saliency response to the orientation and color feature contrasts between the oddball and the remaining homogeneous items, as described by several notable computational saliency models (Itti *et al.*, 1998; Itti & Koch, 2001; Borji & Itti, 2013).

Although the current study is in agreement with our previously work (White, Berg, *et al.*, 2017; White, Kan, *et al.*, 2017), the results raise novel questions that are worth pursuing in future research. For example, there are important differences here from our earlier work that are worthy of mention. In particular, because the stationary visual stimuli sweep across the visual field during the trial in the current design, we are unable to make a direct comparison to the visual conditions associated with fixation in our previous work. As a result, although the data support the general point that SCs neurons encode salient stimuli during pursuit, the results fall short of showing that the saliency map within the SC is the same during smooth pursuit and active fixation. One way to address this would be to have the entire array move synchronous with the fovea during pursuit, such that the oddball remains centered in the RF across the duration of the movement. Another useful control condition would be to have gaze fixed at center while the array/singleton sweeps across the visual field at the same speed associated with the pursuit movement (resulting in the same visual input as the current study without invoking the pursuit system at all). This approach would allow a more direct comparison to the state of the saliency map during active fixation (White, Kan, *et al.*, 2017) and smooth pursuit (as in the current study). Additionally, this would allow us to examine the role of spatial updating of the saliency map during smooth pursuit (Dash *et al.*, 2015, 2016).

In conclusion, the primate SC, which has played a longstanding role in gaze control (Gandhi & Katnani, 2011; White & Munoz, 2011a) and visual attention (Goldberg & Wurtz, 1972;

418 Krauzlis *et al.*, 2013), and more recently visual saliency processing (White, Berg, *et al.*, 2017;
419 White, Kan, *et al.*, 2017), functions qualitatively similar during foveation of stationary or moving
420 stimuli. The results of this study should provide another layer of validation for models and
421 theories of the saliency map, and the particularly important role that the superior colliculus
422 appears to play in this respect.

ACKNOWLEDGEMENTS

The authors thank Ann Lablans, Donald Brien, Sean Hickman and Mike Lewis for outstanding technical assistance. This project was funded by the Canadian Institutes of Health Research (Grant number MOP-FDN-148418). DPM was supported by the Canada Research Chair Program.

COMPETING INTERESTS

The authors have no competing interests to declare.

DATA ACCESSIBILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

AUTHOR CONTRIBUTIONS

B.J.W. and D.P.M. conceptualized and designed the experiment. B.J.W wrote the REX code to run the experiment, and wrote the C code to present and control the visual display. B.J.W. collected and analyzed the data using custom scripts in Matlab (*Mathworks, Inc.*) and wrote the manuscript. D.P.M and L.I provided guidance in analyses, interpretation of results, and manuscript preparation/revision.

REFERENCES

Berg, D.J., Boehnke, S.E., Marino, R.A., Munoz, D.P., & Itti, L. (2009) Free viewing of dynamic stimuli by humans and monkeys. *J. Vis.*, **9**, 1–15.

Bisley, J.W. & Goldberg, M.E. (2010) Attention, intention, and priority in the parietal lobe. *Annu. Rev. Neurosci.*, **33**, 1–21.

Boehnke, S.E. & Munoz, D.P. (2008) On the importance of the transient visual response in the superior colliculus. *Curr. Opin. Neurobiol.*, **18**, 544–551.

Borji, A. & Itti, L. (2013) State-of-the-art in visual attention modeling. *IEEE Trans. Pattern Anal. Mach. Intell.*, **35**, 185–207.

Braun, D.I., Schütz, A.C., & Gegenfurtner, K.R. (2017) Visual sensitivity for luminance and chromatic stimuli during the execution of smooth pursuit and saccadic eye movements. *Vision Res.*, **136**, 57–69.

Burrows, B.E. & Moore, T. (2009) Influence and limitations of popout in the selection of salient visual stimuli by area V4 neurons. *J. Neurosci.*, **29**, 15169–15177.

Cerkevich, C.M., Lyon, D.C., Balaram, P., & Kaas, J.H. (2014) Distribution of cortical neurons projecting to the superior colliculus in macaque monkeys. *Eye Brain*, **2014**, 121–137.

Chen, J., Valsecchi, M., & Gegenfurtner, K.R. (2017) Attention is allocated closely ahead of the target during smooth pursuit eye movements: Evidence from EEG frequency tagging. *Neuropsychologia*, **102**, 206–216.

Chukoskie, L. & Movshon, J.A. (2009) Modulation of Visual Signals in Macaque MT and MST Neurons During Pursuit Eye Movement. *J. Neurophysiol.*, **102**, 3225–3233.

Dash, S., Nazari, S.A., Yan, X., Wang, H., & Crawford, J.D. (2016) Superior Colliculus Responses to Attended, Unattended, and Remembered Saccade Targets during Smooth Pursuit Eye Movements. *Front. Syst. Neurosci.*, **10**, 1–12.

Dash, S., Yan, X., Wang, H., & Crawford, J.D. (2015) Continuous updating of visuospatial memory in superior colliculus during slow eye movements. *Curr. Biol.*, **25**, 267–274.

Derrington, A., Krauskopf, J., & Lennie, P. (1984) Chromatic mechanisms in lateral geniculate-nucleus of macaque. *J. Physiol.*, **357**, 241–265.

Erickson, R.G. & Thier, P. (1991) A neuronal correlate of spatial stability during periods of self-induced visual motion. *Exp. Brain Res.*, **86**, 608 616.

Fecteau, J.H. & Munoz, D.P. (2006) Salience, relevance, and firing: a priority map for target selection. *Trends Cogn. Sci.*, **10**, 382–390.

Gandhi, N.J.N. & Katnani, H.A.H. (2011) Motor functions of the superior colliculus. *Annu. Rev. Neurosci.*, **34**, 205–231.

Goldberg, M.E. & Wurtz, R.H. (1972) Activity of superior colliculus in behaving monkey. II. Effect of attention on neuronal responses. *J Neurophysiol*, **35**, 560–574.

Gottlieb, J.P., Kusunoki, M., & Goldberg, M.E. (1998) The representation of visual salience in monkey parietal cortex. *Nature*, **391**, 481–484.

Hafed, Z.M. & Krauzlis, R.J. (2008) Goal representations dominate superior colliculus activity during extrafoveal tracking. *J. Neurosci.*, **28**, 9426–9439.

Hays, A.V.J., Richmond, B.J., & Optican, L.M. (1982) Unix-based multiple-process system, for real-time data acquistion and control. In *WESCON Conf. Proc.* pp. 1–10.

Ignashchenkova, A., Dicke, P.W., Haarmeier, T., & Thier, P. (2004) Neuron-specific contribution of the superior colliculus to overt and covert shifts of attention. *Nat. Neurosci.*, **7**, 56–64.

Ilg, U.J. (1996) Inability of rhesus monkey area V1 to discriminate between self-induced and externally induced retinal image slip. *Eur. J. Neurosci.*, **8**, 1156–1166.

Inaba, N., Shinomoto, S., Yamane, S., Takemura, A., & Kawano, K. (2007) MST Neurons Code for Visual Motion in Space Independent of Pursuit Eye Movements. *J. Neurophysiol.*, **97**, 3473–3483,.

Itti, L. (2005) Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. *Vis. cogn.*, **12**, 1093–1123.

494 Itti, L. & Koch, C. (2001) Computational modeling of visual attention. *Nat. Rev. Neurosci.*, **2**,
495 194–203.

496 Itti, L., Koch, C., & Niebur, E. (1998) A model of saliency-based visual attention for rapid scene
497 analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, **20**, 1254–1259.

498 Kerzel, D., Souto, D., & Ziegler, N.E. (2008) Effects of attention shifts to stationary objects
499 during steady-state smooth pursuit eye movements. *Vision Res.*, **48**, 958–969.

500 Khan, A.Z., Lefèvre, P., Heinen, S.J., & Blohm, G. (2010) The default allocation of attention is
501 broadly ahead of smooth pursuit. *J. Vis.*, **10**.

502 Khijrana, B. & Kowler, E. (1987) Shared attentional control of smooth eye movement and
503 perception. *Vision Res.*, **27**, 1603–1618.

504 Krauzlis, R.J. (2003) Recasting the Smooth Pursuit Eye Movement System. *J. Neurophysiol.*, **91**,
505 591–603.

506 Krauzlis, R.J., Lovejoy, L.P., & Zénon, A. (2013) Superior colliculus and visual spatial attention.
507 *Annu. Rev. Neurosci.*, **36**, 165–182.

508 Kreyenmeier, P., Fooken, J., & Spering, M. (2017) Context effects on smooth pursuit and
509 manual interception of a disappearing target. *J. Neurophysiol.*, **118**, 404–415.

510 Li, W., Piech, V., & Gilbert, C.D. (2006) Contour Saliency in Primary Visual Cortex. *Neuron*,
511 **50**, 951–962.

512 Li, Z. (2002) A saliency map in primary visual cortex. *Trends Cogn. Sci.*,.

513 Lock, T.M., Baizer, J.S., & Bender, D.B. (2003) Distribution of corticotectal cells in macaque.
514 *Exp. Brain Res.*, **151**, 455–470.

515 Lovejoy, L.P., Fowler, G.A., & Krauzlis, R.J. (2009) Spatial Allocation of Attention During
516 Smooth Pursuit Eye Movements. *Vision Res.*, **49**, 1275–1285.

517 Marino, R.A., Levy, R., Boehnke, S., White, B.J., Itti, L., & Munoz, D.P. (2012) Linking visual
518 response properties in the superior colliculus to saccade behavior. *Eur. J. Neurosci.*, **35**,

519 1738–1752.

520 Marino, R.A., Rodgers, C.K., Levy, R., & Munoz, D.P. (2008) Spatial relationships of
521 visuomotor transformations in the superior colliculus map. *J. Neurophysiol.*, **100**, 2564–
522 2576.

523 McPeek, R.M. & Keller, E.L. (2002) Saccade target selection in the superior colliculus during a
524 visual search task. *J. Neurophysiol.*, **88**, 2019–2034.

525 Meredith, M. a & Ramoa, a S. (1998) Intrinsic circuitry of the superior colliculus:
526 pharmacophysiological identification of horizontally oriented inhibitory interneurons. *J.*
527 *Neurophysiol.*, **79**, 1597–1602.

528 Munoz, D.P. & Istvan, P.J. (1998) Lateral inhibitory interactions in the intermediate layers of the
529 monkey superior colliculus. *J. Neurophysiol.*, **79**, 1193–1209.

530 Nummela, S.U. & Krauzlis, R.J. (2010) Inactivation of primate superior colliculus biases target
531 choice for smooth pursuit, saccades, and button press responses. *J. Neurophysiol.*, **104**,
532 1538–1548.

533 Phongphanphanee, P., Marino, R.A., Kaneda, K., Yanagawa, Y., Munoz, D.P., & Isa, T. (2014)
534 Distinct local circuit properties of the superficial and intermediate layers of the rodent
535 superior colliculus. *Eur. J. Neurosci.*, **40**, 2329–2343.

536 Purcell, B.A., Schall, J.D., Logan, G.D., & Palmeri, T.J. (2012) From Salience to Saccades:
537 Multiple-Alternative Gated Stochastic Accumulator Model of Visual Search. *J. Neurosci.*,
538 **32**, 3433–3446.

539 Rashbass, C. (1961) The relationship between saccadic and smooth tracking eye movements. *J.*
540 *Physiol.*, **159**, 326–338.

541 Schütz, A.C., Braun, D.I., Kerzel, D., & Gegenfurtner, K.R. (2008) Improved visual sensitivity
542 during smooth pursuit eye movements. *Nat. Neurosci.*, **11**, 1211.

543 Serences, J.T. & Yantis, S. (2006) Selective visual attention and perceptual coherence. *Trends*
544 *Cogn. Sci.*, **10**, 38–45.

Shen, K. & Paré, M. (2014) Predictive saccade target selection in superior colliculus during visual search. *J. Neurosci.*, **34**, 5640–5648.

Spering, M. & Gegenfurtner, K.R. (2007) Contextual effects on smooth-pursuit eye movements. *J. Neurophysiol.*, **97**, 1353–1367.

Thompson, K.G. & Bichot, N.P. (2004) A visual salience map in the primate frontal eye field. *Prog. Brain Res.*, **147**, 251–262.

Veale, R., Hafed, Z.M., & Yoshida, M. (2017) How is visual salience computed in the brain? Insights from behaviour, neurobiology and modelling. *Philos. Trans. R. Soc. B Biol. Sci.*, **372**, 20160113.

White, B.J., Berg, D.J., Kan, J.Y., Marino, R.A., Itti, L., & Munoz, D.P. (2017) Superior colliculus neurons encode a visual saliency map during free viewing of natural dynamic video. *Nat. Commun.*, **8**, 1–9.

White, B.J., Kan, J.Y., Levy, R., Itti, L., & Munoz, D.P. (2017) Superior colliculus encodes visual saliency before the primary visual cortex. *Proc. Natl. Acad. Sci. U. S. A.*, **114**, 9451–9456.

White, B.J. & Munoz, D.P. (2011a) The superior colliculus. In *The Oxford Handbook of Eye Movements*. Oxford University Press, pp. 195–213.

White, B.J. & Munoz, D.P. (2011b) Separate visual signals for saccade initiation during target selection in the primate superior colliculus. *J. Neurosci.*, **31**, 1570–1578.

Yan, Y., Zhaoping, L., & Li, W. (2018) Bottom-up saliency and top-down learning in the primary visual cortex of monkeys. *Proc. Natl. Acad. Sci.*, **115**, 201803854.

Yoshida, M., Itti, L., Berg, D.J., Ikeda, T., Kato, R., Takaura, K., White, B.J., Munoz, D.P., & Isa, T. (2012) Residual attention guidance in blindsight monkeys watching complex natural scenes. *Curr. Biol.*, **22**, 1429–1434.

Zénon, A. & Krauzlis, R.J. (2012) Attention deficits without cortical neuronal deficits. *Nature*, **489**, 434–437.

Zhang, X., Zhaoping, L., Zhou, T., & Fang, F. (2012) Neural Activities in V1 Create a Bottom-Up Saliency Map. *Neuron*, **73**, 183–192.

FIGURE CAPTIONS

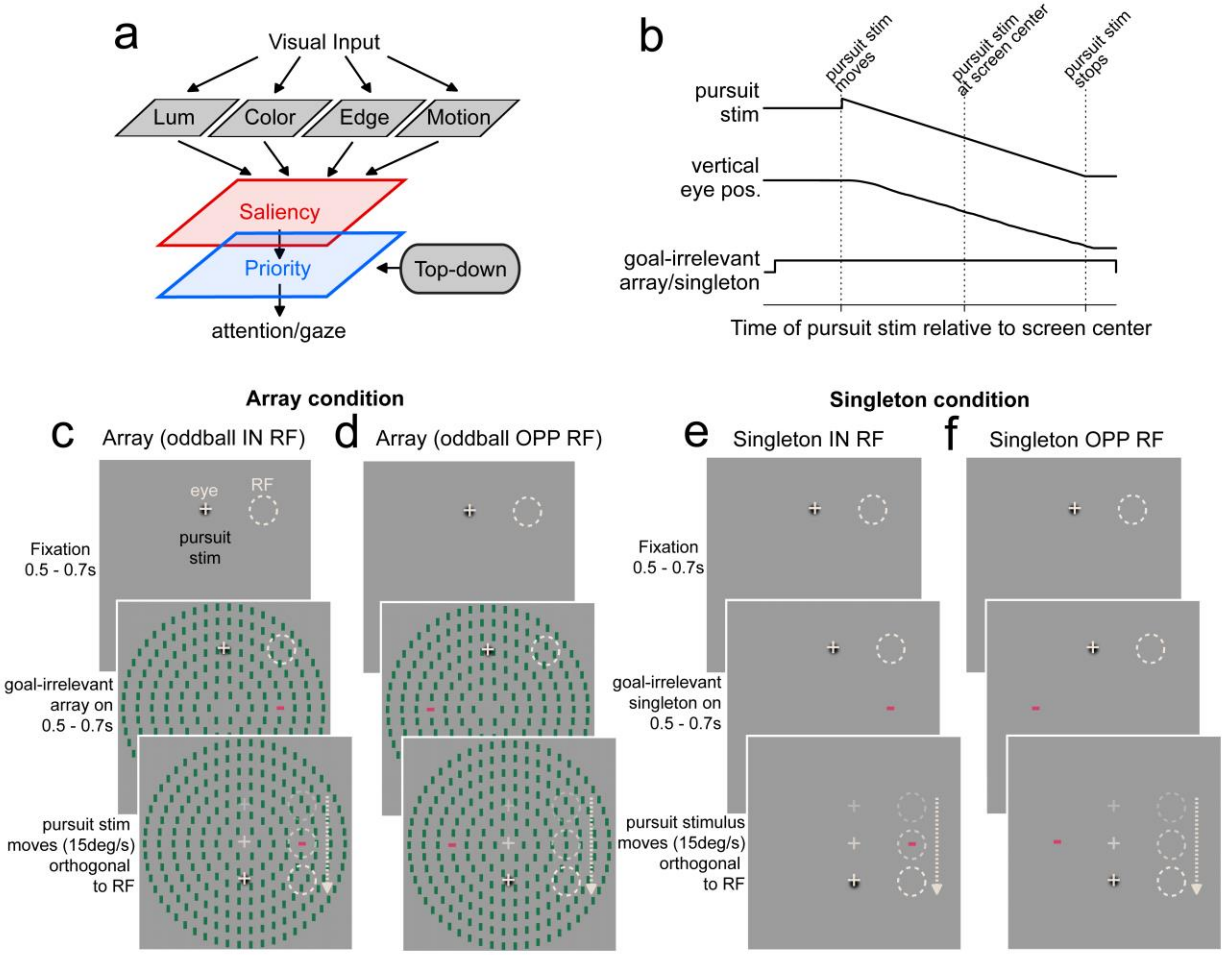

Figure 1. Saliency coding during smooth pursuit eye movements. **(a)** Conceptual model of the saliency/priority map. Visual input is transformed into a topographic map of conspicuity whereby certain stimuli stand out from others based on low-level visual features (saliency map, red). The priority map (blue) combines inputs from the saliency map with top-down goal-dependent signals to determine attention and gaze. **(b)** Temporal illustration of the smooth pursuit task. Rhesus monkeys were trained to smoothly pursue a single black Gaussian-windowed stimulus (~1.5º in diameter, SD=0.3º) which moved at 15º/s. Upper trace shows downward moving pursuit stimulus using the Rashbass step-ramp procedure to reduce saccades during pursuit initiation(Rashbass, 1961) (see Materials and Methods). Middle trace shows example vertical eye position. Lower trace indicates the relative timing of the goal-irrelevant array/singleton. **(c-f)** Spatial illustration of the stimuli and task. On a given trial, the animal fixated the pursuit stimulus for 0.5s-0.7s, after which the array (**c, d**) or singleton (**e, f**) appeared. After an additional 0.5s-0.7s the pursuit stimulus moved towards then past center screen along a trajectory that was 90º relative to the radial direction of the RF, such that the RF was drawn over the salient oddball/singleton (in RF conditions, **c, e**). On half the trials the oddball/singleton appeared opposite the RF (**d, f**). Pursuit occurred in both upward and downward directions, and all conditions were randomly interleaved. Note, the array/singleton remained stationary and did not move with the pursuit target. Also, because RFs were not strictly along the horizontal meridian (**c**), pursuit directions were not strictly vertical as in this illustration. Visually evoked responses were aligned on the time in which the pursuit stimulus crossed screen center, which corresponded to the time the RF was aligned with the salient oddball/singleton.

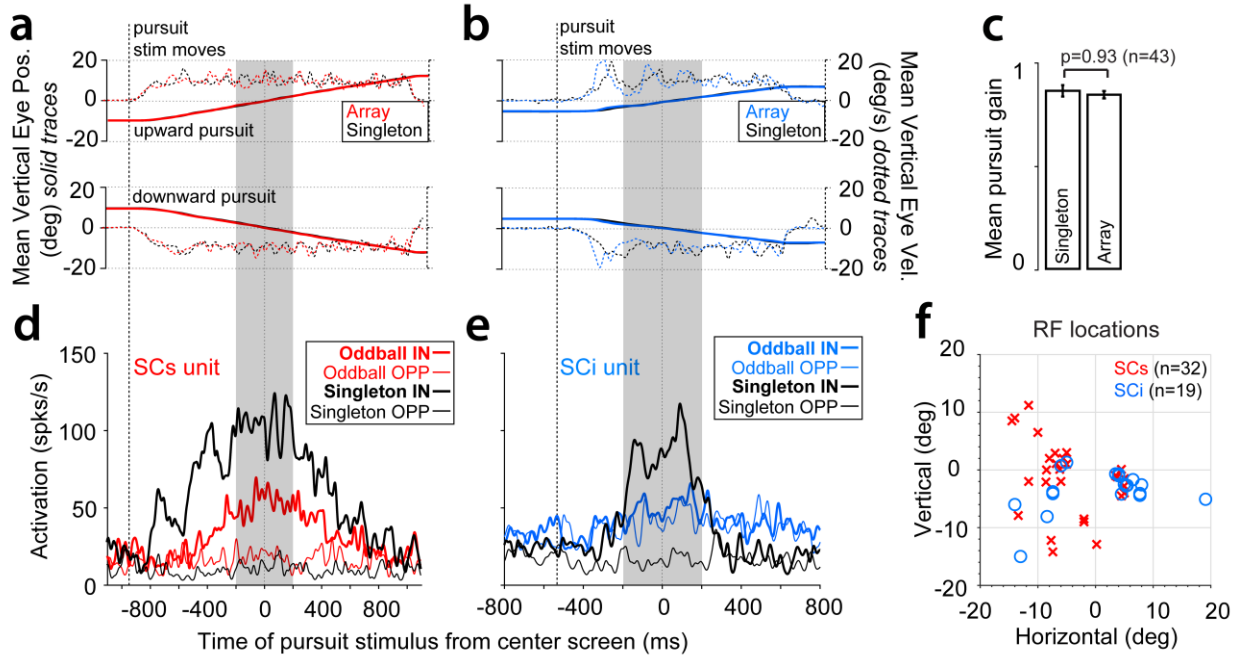

Figure 2. SCs signals saliency during smooth pursuit eye movements. (**a, b**) Mean vertical eye position from two representative recording sessions. Colored traces indicate the array condition, and black traces indicate the singleton condition. Note the duration of the pursuit stimulus was slightly different between these two examples, based on the fact that it is dependent upon RF eccentricity (see Materials and Methods). (**c**) Comparison of mean smooth pursuit gain between the singleton and array conditions. (**d, e**) Spike density functions for the main conditions for a single SCs visual neuron (red) and a single SCi visuomotor neuron (blue). Black traces indicate the singleton control condition for each neuron type. Differences between the IN versus OPP conditions were tested statistically by averaging across the epoch illustrated by the grey vertical shading. **(f)** Distribution of RF locations across n=51 recordings. Errorbars in **c** refer to ±1 SEM between sessions (n=51).

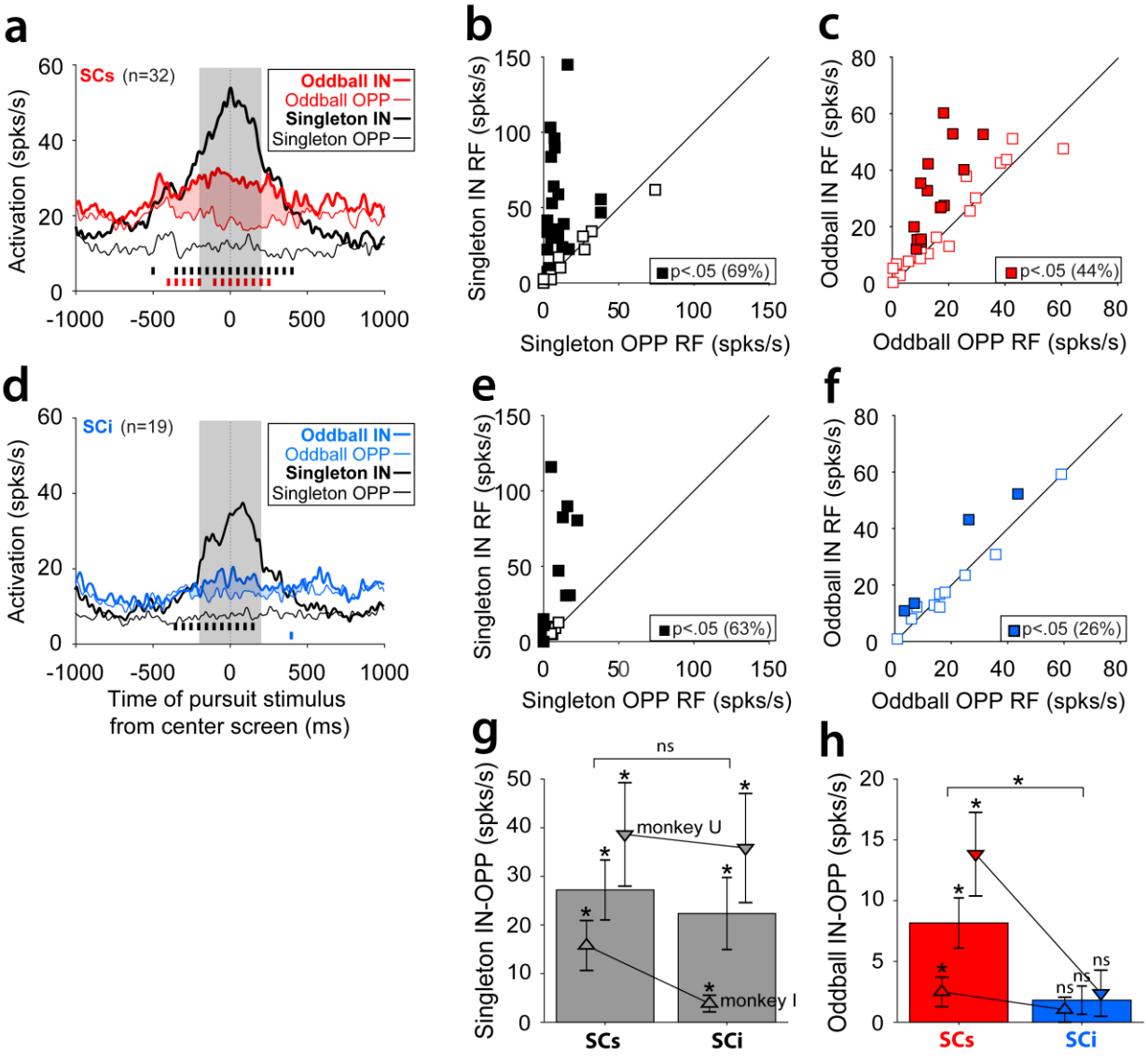

Figure 3. Population spike density functions across the critical conditions. Average spike density functions for the main conditions for a sample of (**a-c**) n=32 SCs visual neurons and (**d-f**) n=19 SCi visuomotor neurons. Tick marks along the x-axis in **a** and **d** indicate a significant difference between the oddball/singleton IN versus oddball/singleton OPP conditions ($P < 0.05$, *Wilcoxon signed rank test* in a 50ms moving window at 50ms intervals from -500ms to +500ms, Bonferroni–Holm corrected). (**b, c, e, f**) Mean responses of each cell for the oddball/singleton IN versus OPP conditions for the sample of 32 SCs visual neurons, and 19 SCi visuomotor neurons, averaged across the test epoch represented by the vertical grey shading in **a** and **d** (±200ms around time the pursuit stimulus crossed center screen). The filled symbols indicate neurons that showed a significantly greater response for the oddball IN versus OPP conditions, $P < .05$, $df = 4$ to $df = 43$ across all neurons, *Wilcoxon Ranksum test*. (**g-h**) The mean difference in firing rate when the singleton/oddball appeared IN versus OPP the RF, comparing SCs and SCi neurons, and comparing monkey U (n=27) and monkey I (n=24). Note, the asterisk above each bar refers to a significant within-neuron *Wilcoxon signed rank test against zero median*, whereas the asterisk above the horizontal bars refers to a significant between-neuron difference using a *Wilcoxon Ranksum test*. Errorbars in **g** and **h** denote ±1 SEM.

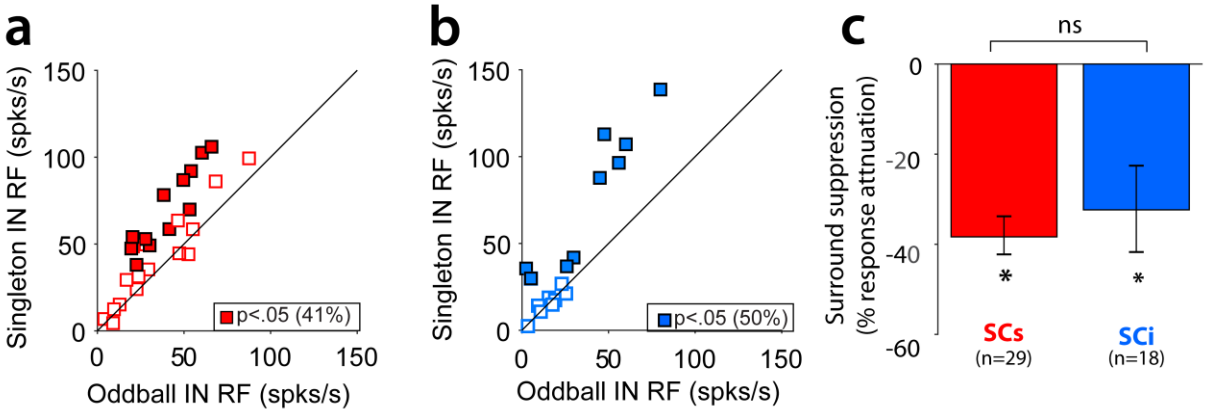

Figure 4. SCs and SCi surround suppression. (**a-b**) Comparison between activation evoked by singleton IN RF (y-axis) versus array with oddball IN RF (x-axis) for sample of n=29 SCs neurons (red) and n=18 SCi neurons (blue). The filled symbols indicate neurons that showed a significantly greater response for the singleton IN (no surround) versus oddball IN (surround) conditions, *Wilcoxon Ranksum test, P < .05*. (**c**) Mean percent surround suppression was defined as the percentage of response attenuation associated with the array condition (Oddball IN RF) relative to the singleton IN RF condition, which did not contain surrounding stimulation. Note, 3 SCs neurons and 1 SCi neuron were excluded due to highly inflated surround suppression scores caused by small differences at low firing rates. *$P < .05$, *Wilcoxon signed rank test against zero median*. ns=not significant. Errorbars refer to ±1 SEM between neurons.