Superior colliculus encodes visual saliency before the primary visual cortex

Brian J. White[a,1], Janis Y. Kan[a], Ron Levy[a,b], Laurent Itti[c], and Douglas P. Munoz[a,d,e,f]

[a]Centre for Neuroscience Studies, Queen's University, Kingston, ON K7L 3N6, Canada; [b]Department of Surgery, Queen's University, Kingston, ON K7L 3N6, Canada; [c]Department of Computer Science, University of Southern California, Los Angeles, CA 95120 [d]Department of Biomedical and Molecular Sciences, Queen's University, Kingston, ON K7L 3N6, Canada; [e]Department of Medicine, Queen's University, Kingston, ON K7L 3N6, Canada; and [f]Department of Psychology, Queen's University, Kingston, ON K7L 3N6, Canada

Edited by Marlene Cohen, University of Pittsburgh, Pittsburgh, PA and accepted by Editorial Board Member Marlene Behrmann July 18, 2017 (received for review January 18, 2017)

Models of visual attention postulate the existence of a bottom-up saliency map that is formed early in the visual processing stream. Although studies have reported evidence of a saliency map in various cortical brain areas, determining the contribution of phylogenetically older pathways is crucial to understanding its origin. Here, we compared saliency coding from neurons in two early gateways into the visual system: the primary visual cortex (V1) and the evolutionarily older superior colliculus (SC). We found that, while the response latency to visual stimulus onset was earlier for V1 neurons than superior colliculus superficial visual-layer neurons (SCs), the saliency representation emerged earlier in SCs than in V1. Because the dominant input to the SCs arises from V1, these relative timings are consistent with the hypothesis that SCs neurons pool the inputs from multiple V1 neurons to form a feature-agnostic saliency map, which may then be relayed to other brain areas.

attention | priority | vision | gaze | oculomotor

Most theories and computational models of saliency postulate that visual input is transformed into a topographic representation of visual conspicuity (Fig. 1*A*, red), whereby certain stimuli stand out from others based on low-level features of the input image (Fig. 1*A*, blue) (1–3). The concept of a priority map describes a combined representation of visual saliency and behavioral relevancy (Fig. 1*A*, yellow), which is thought to be the core determinant of attention and gaze (4, 5). To date, most studies have reported evidence of saliency and/or priority maps in a distributed network of cortical brain areas [e.g., primary visual cortex (V1) (6–9), visual area 4 (V4) (10), lateral intraparietal area (LIP) (11–13), and frontal eye fields (14, 15)]. However, there is mounting evidence for a subcortical saliency mechanism in the premammalian optic tectum (16–18) or superior colliculus (SC) in primates (Fig. 1*B*). The primate SC, which has received a lot of attention for its role as an oculomotor hub, might be considered an unlikely candidate for a visual salience map, but there is a rich history of research on visual attention in the SC (a recent review is in ref. 19), which has broadened our perspective of its role in processes previously thought to be the domain of neocortex.

The SC (Fig. 1*B*) is multilayered but is often described as having two dominant functional layers, a superior colliculus superficial layer (SCs) associated exclusively with visual processing and a superior colliculus multisensory–cognitive–motor intermediate layer (SCi) linked to the control of attention and gaze (19–22). Because SCs is interconnected with multiple visual areas (23–25), it is in an ideal location to pool diverse visual inputs to form a feature-agnostic saliency representation. Recently (26), it has been shown that the activity of SCs neurons, with dominant inputs that arise from the retina and V1 (Fig. 1*B*) (24, 25), is well-predicted by a computational saliency model that has been validated on the free viewing behavior of humans (27) and nonhuman primates (28). In addition, a recent lesion study has shown that attention guidance is preserved in the absence of V1 (29), thereby challenging the hypothesis that saliency coding depends critically on V1 (6–9) and instead, implicating a possible pathway through the SC. This poses the question of whether the evolutionarily older SC still plays a significant role in the computation of saliency or whether saliency representations observed there reflect computations done elsewhere.

Here, we explored the hypothesis that SCs, not V1, embodies the role of a saliency map. To test this hypothesis, we recorded from neurons in the SC and V1 (of different animals), while using a task and stimuli (Fig. 1 *C*–*F*) designed to compare the timing of saliency representations, and the long-range spatial interactions that give rise to saliency in complex scenes. Rhesus monkeys performed a visually guided saccade task while presented with salient but goal-irrelevant stimuli that they were required to ignore. The goal-irrelevant stimulus consisted of a wide-field array (210 items spanning 40° to 50° of visual angle) of oriented color bars (~0.4° × 1.2°) with a salient oddball, forming what is traditionally described as perceptual pop out (30). We measured visually evoked responses when the salient oddball appeared within (IN) vs. opposite/contralateral (OPP) to the response field (RF), representing high- vs. low-saliency regions of the display. Because models of visual saliency depend critically on center-surround feature contrasts that extend widely across the visual field (1–3), quantification of the surround suppression characteristics can also provide a strong index of the degree to which each brain area represents saliency. Therefore, a single-item control condition (Fig. 1 *E* and *F*) served as a benchmark of visual responses with no competing surround, which allowed us to differentiate first-order saliency (local luminance change associated with a sudden visual onset) from second-order saliency

Significance

Theories of visual attention postulate the existence of a saliency map that guides attention/gaze toward the most visually conspicuous stimuli in complex scenes. This study compared saliency coding in the two dominant visual gateways: the primary visual cortex (V1) and the evolutionarily older visual system that exists in the midbrain superior colliculus. Our results show that neurons in the superficial visual layers of the superior colliculus (SCs) encoded saliency earlier and more robustly than V1 neurons. This was surprising, because the dominant input to the SCs arises from V1. This result is in line with models that place a feature processing stage (V1) before the feature-agnostic saliency map in SCs.

Author contributions: B.J.W., J.Y.K., L.I., and D.P.M. designed research; B.J.W. and J.Y.K. performed research; R.L. contributed new reagents/analytic tools; B.J.W. and J.Y.K. analyzed data; R.L. was the primary surgeon in the chronic implantation of a microelectrode array in the primary visual cortex; L.I. contributed significantly to the research idea and hypotheses and cowrote portions of the manuscript; and B.J.W., J.Y.K., and D.P.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. M.C. is a guest editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. Email: brian.white@queensu.ca.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1701003114/-/DCSupplemental.

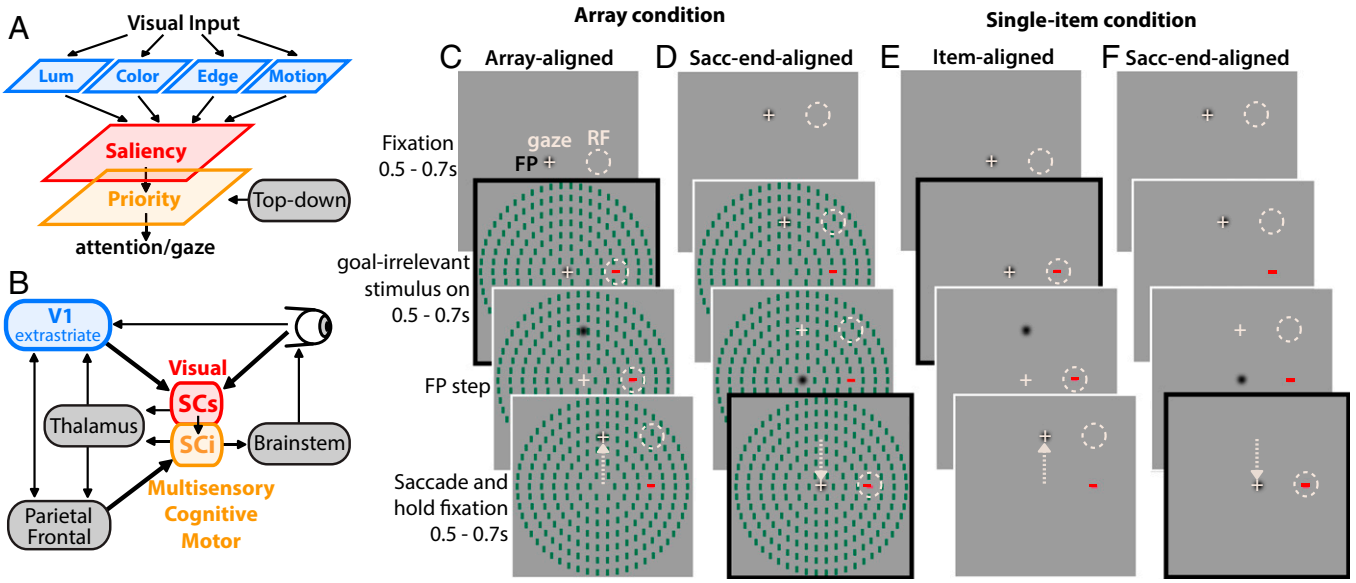

Fig. 1. Comparison of saliency representations in V1, SCs, and SCi. (*A*) Saliency map theory postulates that visual input is decomposed into basic features (blue), which are integrated to form a feature- and behavior-agnostic saliency map (red). A priority map (yellow) is postulated to combine inputs from the saliency map with top-down goal-dependent signals to determine attention and gaze. (*B*) Simplified schematic of the visual attention network, highlighting the dominant inputs and outputs of the primate SC. (*C* and *D*) Rhesus monkeys viewed a wide-field arrangement of goal-irrelevant stimuli with a salient oddball defined by color and orientation, which could appear IN or opposite the RF. Four possible combinations of color and orientation were randomly interleaved, such that the oddball was always the opposite color and orientation as the remaining items. (*E* and *F*) A single-item control condition served as a benchmark with no surround. On a given trial, the animal fixated an FP, which appeared at either the screen center (*C* and *E*) or above or below center (*D* and *F*) at a location orthogonal to the RF at the same eccentricity (i.e., ±90° polar angle relative to the RF). The animals were required to continue fixating on the FP for a 0.5- to 0.7-s random period, after which the goal-irrelevant array/item appeared. The animals were required to continue fixating on the FP for an additional 0.5–0.7 s, after which the FP stepped from center to one of the orthogonal peripheral locations (*C* and *E*) or from the peripheral location to center (*D* and *F*). The frames with the black outlines illustrate the time of alignment for the array/item-aligned (*C* and *E*) and saccade end-aligned (*D* and *F*) conditions. In both cases, visually evoked responses were measured when the eyes were fixated at the screen center, such that the stimulus display was the same.

computations (center-surround contrasts that make a stimulus stand out from its isoluminant neighbors in a complex scene). In addition, the goal-irrelevant oddball/item was centered in the RF in two ways: (*i*) the stimulus appeared abruptly with the oddball/single-item centered IN or OPP the RF; or (*ii*) the oddball/single item was brought IN or OPP the RF via a saccade (11). This allowed us to examine possible differences in saliency coding associated with passive viewing of an abrupt visual onset vs. active gaze-dependent shifting of stimuli over different parts of the visual field via saccades (11), which is characteristic of real-world viewing.

Results

Saliency Coding Emerges Earlier in SCs than V1. We compared visually evoked responses when a salient but goal-irrelevant oddball appeared IN vs. OPP the RF, representing high- vs. low-saliency

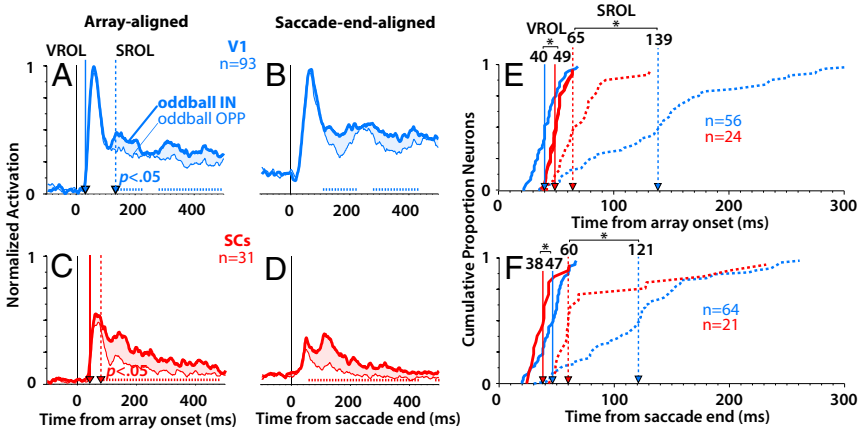

Fig. 2. The time course of saliency representations in V1 and SC. Normalized population average responses for V1 (blue; $n = 93$; monkeys D and Y) and SCs (red; $n = 31$; monkeys I and U) when the salient oddball appeared abruptly in the RF (*A*, *C*, and *E*; array-aligned condition), or was brought there via a saccade (*B*, *D*, and *F*; saccade end-aligned condition). The tick marks immediately above the *x* axis in *A*–*D* indicate the points where the oddball IN response was significantly greater than the oddball OPP response ($P < 0.05$, Wilcoxon signed rank test at 10-ms intervals, Bonferroni–Holm corrected). (*E* and *F*) Cumulative distributions of VROL (solid traces) and SROL (dotted traces), with the relative median latencies indicated by the vertical lines with arrows. *$P < 0.05$, rank sum test.

regions of the display, the difference of which defined the saliency index (*SI Materials and Methods*). Fig. 2 *A–D* shows normalized population-averaged responses for V1 (Fig. 2, blue) and SCs (Fig. 2, red) neurons (note that these results were not affected by the normalization procedure as indicated by the identical statistical outcome using absolute firing rates) (Fig. S1). For a given neuron, we normalized to the averaged peak response of the single-item condition within the window from 0 to 500 ms (*SI Materials and Methods*). Because the single item always produced the greatest visual response, this ensured that the relative differences between conditions for a given neuron were retained. Although V1 and to a lesser degree, SCs (31) neurons showed some feature preference (Fig. S2), the response curves depicted in Fig. 2 represent the data collapsed across the feature combinations, effectively canceling out feature-specific differences. It should be noted that the V1 response curves represent combined single units plus multiunit sites (*Materials and Methods*), the results of which were qualitatively similar but less robust with the single units alone (Fig. S3). At the population level, V1 and SCs neurons showed a significant preference for the salient oddball as indicated by the difference between the oddball IN (Fig. 2, thick traces) and oddball OPP (Fig. 2, thin traces) conditions ($P < 0.05$, Wilcoxon signed rank test at 10-ms intervals, Bonferroni–Holm corrected). In addition, this oddball selectivity was qualitatively similar whether the oddball appeared abruptly in the RF (Fig. 2 *A* and *C*, array-aligned condition) or was brought into the RF via a saccade (Fig. 2 *B* and *D*, saccade end-aligned condition).

To explore these differences in detail, we derived two quantities from the response profiles of each neuron, the visual response onset latency (VROL; the earliest emergence of a visual response) and the saliency response onset latency (SROL; the earliest point at which the neuron signaled a preference for the oddball) (*Materials and Methods*). Fig. 2 *E* and *F* shows the cumulative distributions of VROLs (Fig. 2, solid traces) and SROLs (Fig. 2, dotted traces) for V1 and SCs neurons. Only the subset of neurons which showed significant preference for the oddball was included in the cumulative distribution of SROLs. For the array-aligned condition (Fig. 2*E*), VROL was significantly earlier for V1 (median = 40 ms, $n = 56$) (Fig. 2, vertical blue line) than SCs (median = 49 ms, $n = 24$; $P = 0.0008$, rank sum test) (Fig. 2, vertical red line), whereas SROL was significantly earlier for SCs (median = 65 ms, $n = 24$) than V1 (median = 139 ms, $n = 56$; $P = 0.000021$, rank sum test). These results were qualitatively similar in the saccade end-aligned condition (Fig. 2*F*). There, VROL was significantly longer for V1 (median = 47 ms, $n = 64$) than SCs (median = 38 ms, $n = 21$; $P = 0.013$, rank sum test). More importantly, SROL was significantly earlier for SCs (median = 60 ms, $n = 21$) than V1 (median = 121 ms, $n = 64$; $P = 0.0041$, rank sum test). The results were qualitatively similar whether RFs were overlapping or nonoverlapping between brain areas (Fig. S4).

These results indicate that, although V1 signaled the earliest visual arrival times in the stimulus-aligned condition (although not in the saccade end-aligned condition), on average, the saliency representation in V1 did not emerge until 60–75 ms after it had already appeared in the SCs. This delay was relatively shorter for V1 single units (14–37 ms) (Fig. S3 *E* and *F*) than V1 single units plus multiunits (Fig. 2 *E* and *F*), but in either case, the results were generally the same. Using the larger dataset that included single-unit plus multiunit V1 recordings led to more reliable and statistically powerful V1 results (compare, for example, the difference between the population-averaged spike density traces in Fig. 2 *A* and *B* and Fig. S3 *A* and *B*).

In addition, in V1, there was an average delay of 45–99 ms from visual onset to the emergence of the saliency representation (comparing the blue solid traces with the blue dotted traces in both Fig. 2 *E* and *F* and Fig. S3 *E* and *F*). This delay is exemplified by the fact that the saliency representation in V1 did not emerge until after the initial volley of visual activity, as indicated by the blue vertical dotted line in Fig. 2*A* (Fig. S3*A*). In contrast, while SCs neurons showed overall later VROLs than V1 in the stimulus-aligned condition, the saliency representation emerged within the initial volley of visual activity around the earliest part of the visual response (~15–20 ms after visual response onset) as indicated by the red vertical dotted line in Fig. 2*C*. This relatively slow signaling of saliency in V1 suggests either a processing delay within V1 itself to generate a saliency representation or that the saliency representation in V1 emerged via feedback from other brain areas. Interestingly, VROL was the same or shorter for SCs than V1 in the saccade end-aligned case (Fig. 2*F* and Fig. S3*F*). This might indicate that, under real-world viewing conditions, where stimuli are more often brought into a neuron's RF as a result of saccades, VROL may be earlier in SCs than V1. Most importantly, the saliency representation emerged reliably earlier in SCs than V1.

Surround Modulation Emerges Earlier and Stronger in SCs than V1. Most models of visual saliency depend critically on center-surround feature contrasts that extend widely across the visual field (1–3), and therefore, comparison of the surround suppression characteristics can provide a strong index of the degree to which each brain area represents saliency. To quantify surround suppression, we compared visual responses evoked by the wide-field array (surround) (Fig. 1*C*) with a single item (no surround) (Fig. 1*E*). Fig. 3 *A–D* shows population responses in the single-item IN condition (no surround) (Fig. 3, black trace) vs. the oddball IN condition (surround) (Fig. 3, color trace). From these averaged population traces, one can see noticeable differences in the timing of surround suppression between brain areas, with SCs (Fig. 3 *C* and *D*) showing markedly earlier suppression than V1 (Fig. 3 *A* and *B*) [$P < 0.05$, Wilcoxon signed rank test at 10-ms intervals, Bonferroni–Holm corrected; note that these results were not affected by the normalization procedure as indicated by the identical statistical outcome using absolute firing rates (Fig. S5)]. In particular, the suppression in V1 did not emerge until well after the initial volley of visual activity as exemplified by the blue vertical dotted line in Fig. 3*A* (Fig. S6*A*). This is in contrast to SCs, in which the suppression emerged within or before the peak of the initial volley of visual activity as exemplified by the red vertical dotted line in Fig. 3*C*. To quantify the magnitude of the differences, we compared the percentage of suppression within two time periods: (*i*) during the peak of the visual response and (*ii*) during the sustained portion of the response (details are in *SI Materials and Methods*). Percentage suppression (Fig. 3 *E–H*) was significantly greater for SCs neurons (Fig. 3, red) than for V1 neurons (Fig. 3, blue) during both the peak (Fig. 3 *E*, array-aligned; $P = 9.3e\text{-}07$ and *G*, saccade end-aligned; $P = 1.2e\text{-}05$) and sustained (Fig. 3 *F*, array-aligned; $P = 0.0089$ and *H*, saccade end-aligned; $P = 2.8e\text{-}04$) portions of the response (rank sum test).

As with SROL above, we computed the surround suppression onset latency (SSOL) for each neuron that showed significant surround suppression, defined as the earliest time in which a neuron signaled a significantly greater response for the single-item IN condition (no surround) vs. the oddball IN condition (surround) (*SI Materials and Methods*). Fig. 3*I* shows that, for V1 neurons in the stimulus-aligned condition, SSOL (median = 73 ms, $n = 69$) was significantly delayed relative to VROL (median = 40 ms, $n = 69$; $P = 6.8e\text{-}16$). In contrast, for SCs neurons in the stimulus-aligned condition, SSOL (median = 51 ms, $n = 25$) was not significantly delayed relative to VROL (median VROL = 49 ms, $n = 25$; $P = 0.18$). Fig. 3*J* shows that, for V1 neurons in the saccade end-aligned condition, SSOL (median = 95 ms, $n = 52$) was again significantly delayed relative to VROL (median = 46 ms, $n = 52$; $P = 2.2e\text{-}12$). For SCs neurons in the saccade end-aligned condition, SSOL (median = 51 ms, $n = 22$) was moderately delayed relative to VROL (median VROL = 39 ms, $n = 22$; $P = 0.0004$). Importantly, SSOL emerged in V1, on average, 22–44 ms later

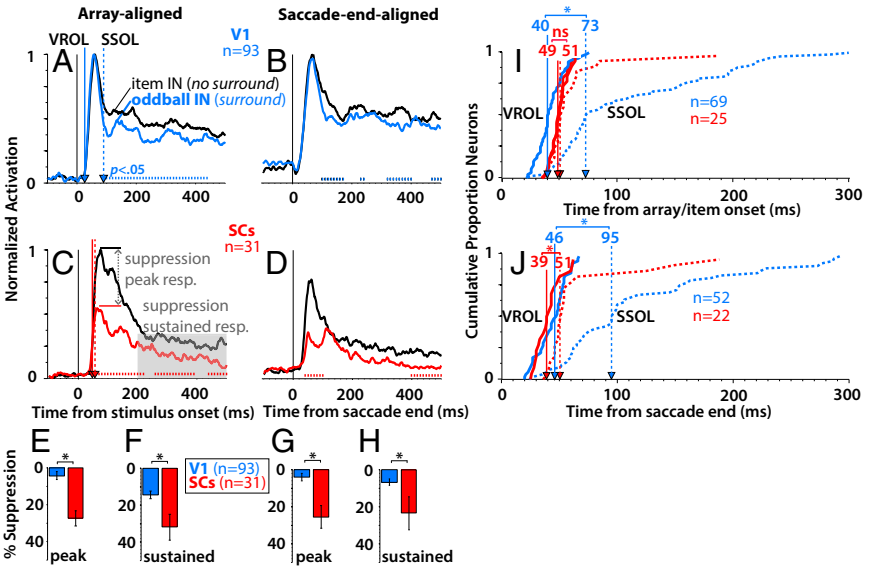

Fig. 3. Surround suppression in V1 and SC. (*A–D*) Normalized population average responses for V1 (blue; $n = 93$; monkeys D and Y) and SCs (red; $n = 31$; monkeys I and U) comparing the single-item IN condition (black traces; no surround) with the oddball IN condition (colored traces; surround) aligned on array/item onset (*A* and *C*) and saccade end (*B* and *D*). The tick marks immediately above the *x* axis in *A–D* indicate the points where the single-item IN response was significantly greater than the oddball IN response ($P < 0.05$, Wilcoxon signed rank test at 10-ms intervals, Bonferroni–Holm corrected). (*E–H*) Averaged percentage suppression across V1 ($n = 93$) and SCs ($n = 31$) recordings. Percentage suppression was computed from the ratio of response evoked by the single item (no surround) vs. the array (surround) with the oddball IN the RF and was computed during the peak and during a sustained period of the response (*Materials and Methods*). (*I* and *J*) Cumulative distributions showing a comparison of the SSOL (dotted traces) relative to VROL (solid traces). ns, Not significant. *$P < 0.05$, rank sum test.

than in SCs (note the difference between the red and blue dotted curves in Fig. 3 *I* and *J*) ($P = 2.8e\text{-}05$ for stimulus-aligned data; $P = 6.1e\text{-}05$ for saccade end-aligned data, rank sum test). These results were qualitatively similar using V1 single units alone (Fig. S6 *I* and *J*). The results were also qualitatively similar whether RFs were overlapping or nonoverlapping between brain areas (Fig. S7). This indicates that, similar to the SROL results described above, SSOL also showed relatively slow dynamics in V1. This suggests either a processing delay to generate surround suppression within V1 itself or that surround modulation in V1 arises, in part, via feedback from other brain areas (32–34). In either case, these results are in agreement with the hypothesis that the SCs, but not V1, has center-surround characteristics suited for computing saliency from wide-field visual inputs (35).

Goal-Irrelevant Saliency Is Weakly Represented in SCi. The input–output structure and response characteristics of SCi neurons differ in many important ways from SCs neurons (23–25). In particular, SCi neurons receive a diverse set of inputs predominantly from frontal and parietal cortices and the basal ganglia (Fig. 1*B*) and project directly to the brainstem saccade generator (23–25). Functionally, these neurons have multisensory and oculomotor responses and have been shown to play a central role in spatial attention (19, 20). This has led to the hypothesis that SCi is best described as a priority map for the control of attention and gaze (4) (Fig. 1 *A* and *B*). For these reasons, we predicted that SCi neurons would show important differences from SCs neurons on our task designed to measure goal-irrelevant saliency.

Fig. 4 shows the results of our sample of $n = 25$ SCi neurons. In contrast to SCs neurons (Fig. 2), SCi neurons only weakly signaled the presence of the salient oddball, showing a small difference over a shorter period than observed in SCs (Fig. 4 *A* and *B*) ($P < 0.05$, moving Wilcoxon signed rank test, Bonferroni–Holm corrected). Moreover, the response of SCi neurons was especially attenuated in the saccade end-aligned condition (Fig. 4*B*). Recall that these response profiles were normalized to the response of a single unitary item (*Materials and Methods* and Fig. 1*E*), which indicates that, while SCi neurons respond well to a visual stimulus (Fig. 4*C*, black trace), presentation of the entire array resulted in a dramatic response attenuation (Fig. 4 *A–D*, yellow traces). In addition, SCi neurons did show significant surround suppression during the peak response in the stimulus-aligned condition (Fig. 4*E*) ($P = 0.0005$, Wilcoxon signed rank test against zero median) but not during the sustained portion of the response (Fig. 4*F*) ($P = 0.072$) or in the saccade end-aligned condition (Fig. 4 *G* and *H*) ($P > 0.49$ in both cases). SSOL emerged around the same time or just after VROL, and the difference was not significant (Fig. 4*I*) (median VROL = 48 ms, median SROL = 52 ms, $n = 14$; $P = 0.24$, rank sum test).

These results suggest that SCi neurons are not a good candidate for the representation of a saliency map, a hypothesis that is consistent with the results of an earlier study (26). Moreover, the strong attenuation of the overall response in the saccade end-aligned case implies that locations on the SCi map not associated with the saccade goal were particularly suppressed around the time of the saccade, which continued momentarily into the subsequent fixation when the stimuli were brought into the RF (Fig. 4*D*). This was not the result of a misalignment between the stimulus and RF (because of less than perfect saccade accuracy), because V1 neurons showed little attenuation under the same conditions (Fig. 3*B*), and stimulus–RF misalignment would have been more likely in V1 given the characteristically smaller RFs (36, 37). These results suggests a mechanism that restricts goal-irrelevant saliency signals in SCi from passing through to downstream gaze control circuitry.

Discussion

This study compared saliency coding in two early visual areas of the primate visual system, V1, and the hub of the evolutionarily older visual system, SCs. We found that, while both V1 and SCs neurons encoded the presence of a salient but goal-irrelevant oddball, the saliency representation in V1 did not emerge until after it had already emerged in SCs (Fig. 2 *E* and *F*). This indicates that, although V1 is the dominant input to SCs (24, 25),

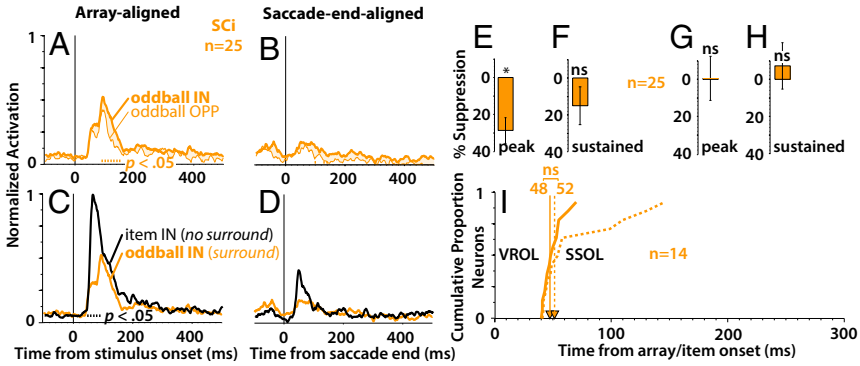

Fig. 4. Saliency representation in SCi. (*A* and *B*) Normalized population average responses for SCi ($n = 25$; monkeys I and U) in the array-aligned condition (*A*) and the saccade end-aligned condition (*B*). (*C* and *D*) Normalized population average responses comparing the single-item IN condition (black traces; no surround) with the oddball IN condition (yellow traces; surround) aligned on array/item onset (*C*), and saccade end (*D*). The tick marks immediately above the *x* axis in *A*–*D* indicate the significant differences between the response curves ($P < 0.05$, Wilcoxon signed rank test at 10-ms intervals, Bonferroni–Holm corrected). (*E*–*H*) Averaged percentage suppression across SCi recordings ($n = 25$). Percentage suppression associated with the peak and sustained portions of the response (*Materials and Methods*). (*I*) Cumulative distribution of SCi neurons comparing SSOL (dotted traces) relative to VROL (solid traces). ns, Not significant. *$P < 0.05$, signed rank test against zero median.

the saliency representation in SCs was not simply a readout of computations done upstream in V1. This is interesting, because V1 neurons registered the arrival of visual signals reliably earlier than SCs in the stimulus-aligned condition (although not the saccade end-aligned condition), but V1 was slower to signal both saliency (Fig. 2 *E* and *F*) and surround modulation (Fig. 3), an essential component for computing saliency across the visual field (1–3). Moreover, the saliency representation was generally limited to the superficial visual SC layers. SCi visuomotor neurons, with dominant inputs that arise from frontal–parietal cortices and the basal ganglia (23–25), did not show a significant preference for the salient oddball (Fig. 4 *A* and *B*). This finding is consistent with an earlier study that reported a weak saliency representation in SCi during free viewing of natural scenes (26). Taken together, the results provide evidence that the evolutionarily older SCs, not V1, embodies the role of a saliency map.

The relative timing of these visual representations offers insight into the flow of information processing between V1 and SC. Using the traditional index of visual latency timed from stimulus onset (Fig. 2 *A* and *C*), V1 was the earliest to receive visual inputs (Fig. 2), which is consistent with the role of an early feature extraction stage (Fig. 1 *A*, blue and *B*, blue). The evolutionarily older SCs then signaled the arrival of visual inputs, which may, in part, be inherited from pooled inputs of multiple V1 neurons (24, 25). This would effectively cancel out most feature-specific preferences, consistent with the broadly tuned characteristics of SCs neurons (31, 38, 39). Interestingly, VROL was the same or shorter for SCs than V1 in the saccade end-aligned case (Fig. 2*F* and Figs. S1, S3–S5, and S7). This might indicate that, under real-world viewing conditions, where stimuli are more often brought into a neuron's RF as a result of saccades, retinotectal inputs might dominate the earliest part of the SCs visual response. In either case, SCs signaled the presence of the salient oddball before V1. Interestingly, the saliency representation in SCs was immediately preceded by robust surround suppression, which was markedly stronger and earlier than in V1. This is strong evidence that SCs has center-surround characteristics better suited for computing saliency from wide-field visual inputs (35). For SCi, visual inputs not associated with the saccade goal were strongly attenuated (Fig. 4*D*). This result suggests a mechanism that suppresses goal-irrelevant visual inputs in SCi to prioritize which salient signals gain access to the downstream gaze control system. Taken together, these results support a simple, intuitive framework for the formation of the saliency map in the primate midbrain, echoing earlier studies in the premammalian optic tectum (16–18).

The timing of saliency representations in other cortical brain areas also suggests that they are not the source of the saliency representation observed here in SCs. For example, V4 neurons signal the presence of an oddball too late (~85–100 ms poststimulus) to be the source of the saliency representation observed in SCs (~65 ms poststimulus) (10). Similarly, estimates of a pure saliency response in LIP emerge too late (~70–75 ms poststimulus onset) (13) to be the source of the saliency representation in SCs, other than the fact that LIP does not project to SCs (24, 25). It is also worth mentioning that the timing of the V1 saliency representation in our study is very similar to the timing of V1 figure-ground modulation (40, 41), both of which indicate a modulation that only occurs well after the initial transient visual response. In sum, the timing of these cortical saliency representations appears too slow to be the source of the saliency representation observed in SCs. Whether these cortical saliency representations arise from ascending tectothalamocortical pathways [for example, via pulvinar (42, 43)] is unknown.

Although there have been viable arguments for the existence of a feature-agnostic saliency map in V1 (6, 7), these arguments are weakened by the lack of direct neurophysiological evidence (a detailed discussion is in ref. 44). The results of this study further challenge the hypothesis that V1 generates the visual saliency map by showing that SCs neurons, with dominant inputs that arise from V1, encoded the presence of a salient oddball sooner and more robustly than V1. This is in close agreement with the results of a recent V1 lesion study by Yoshida et al. (29), which used a computational saliency model to weigh the contribution of each visual feature on free viewing gaze behavior after the surgical removal of V1. In that study, the authors showed that the removal of V1 abolished the contribution of oriented features, whereas the remaining features (luminance, color, motion) remained relatively unaffected. Taken together, this is strong evidence that V1 is not essential for the computation of saliency, whereas SCs remains a strong candidate for the early coding of saliency.

While it has been argued that the saliency map has evolved from the optic tectum (SC) in lower vertebrates to the V1 in primates (7), the results of this study provide a serious challenge for this hypothesis. Certainly, it would seem advantageous for primitive species to have a mechanism that allows one to rapidly engage the orienting system toward stimuli defined by conspicuous features, and the evolutionarily old visual system in the

midbrain seems to have solved this biological problem long before the elaboration of visual cortex (16–18). A potential functional advantage of a saliency mechanism in SC is that it would reside close to the gaze-orienting circuitry. The sensory and motor maps of the SCs and SCi are closely aligned (45) and ideally situated to integrate diverse inputs from cortex and engage the orienting system to act on such inputs. Thus, we postulate that, through evolution, V1 was integrated into an already existing saliency system, providing greater feature specificity and fine-tuned control over visually guided behavior.

Materials and Methods

Data were collected from four male Rhesus monkeys (*Macaca mulatta*): two for the SC recordings and two for the V1 recordings (details are in *SI Materials and Methods*). All procedures were approved by the Queen's University Animal Care Committee in accordance with the guidelines of the Canadian Council on Animal Care.

The stimuli consisted of a radial arrangement of equally spaced color bars (210 items), with the diameter of the entire display spanning 40° to 45° visual angle (Fig. 1 *C* and *D*). The oddball, which could appear IN or OPP the RF, was always the opposite color and orientation as the remaining items, but all stimulus combinations were run interleaved. A single-item control condition served as a benchmark with no surround stimuli (Fig. 1 *E* and *F*).

On a given trial (Fig. 1), the animal fixated a fixation point (FP), which appeared at the screen center (stimulus-aligned conditions) (Fig. 1 *C* and *E*) or above or below center (saccade end-aligned conditions) (Fig. 1 *D* and *F*) orthogonal to the RF at the same eccentricity. The animals were required to fixate the FP for 0.5–0.7 s, after which the goal-irrelevant stimulus/array appeared. The animals continued to fixate the FP for an additional 0.5–0.7 s, after which the FP stepped from center to one of the specified peripheral locations (array/item-aligned conditions) (Fig. 1 *C* and *E*) or from the peripheral location to center (saccade end-aligned conditions) (Fig. 1 *D* and *F*). The animals were then required to launch a saccade to the new FP location and hold fixation for 0.5–0.7 s within a ~3° × 3° computer-controlled window, after which a liquid reward was issued. Visually evoked responses were measured during the stimulus-aligned vs. saccade end-aligned time points, highlighted by the black outlines in key frames of Fig. 1 *C*–*F*.

SI Materials and Methods has extended details about stimuli, equipment, procedure, and data analyses. The data, materials, and code that support the findings of this study are available from B.J.W. on reasonable request.

ACKNOWLEDGMENTS. We thank Ann Lablans, Donald Brien, Sean Hickman, and Mike Lewis for technical assistance. This project was funded by Canadian Institutes of Health Research Grant FDN-148418 and National Science Foundation Grants BCS-0827764 and CCF-1317433. D.P.M. was supported by the Canada Research Chair Program.

1. Itti L, Koch C, Niebur E (1998) A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans Pattern Anal Mach Intell* 20:1254–1259.
2. Itti L, Koch C (2001) Computational modelling of visual attention. *Nat Rev Neurosci* 2: 194–203.
3. Borji A, Itti L (2013) State-of-the-art in visual attention modeling. *IEEE Trans Pattern Anal Mach Intell* 35:185–207.
4. Fecteau JH, Munoz DP (2006) Salience, relevance, and firing: A priority map for target selection. *Trends Cogn Sci* 10:382–390.
5. Serences JT, Yantis S (2006) Selective visual attention and perceptual coherence. *Trends Cogn Sci* 10:38–45.
6. Li Z (2002) A saliency map in primary visual cortex. *Trends Cogn Sci* 6:9–16.
7. Zhaoping L (2016) From the optic tectum to the primary visual cortex: Migration through evolution of the saliency map for exogenous attentional guidance. *Curr Opin Neurobiol* 40:94–102.
8. Zhang X, Zhaoping L, Zhou T, Fang F (2012) Neural activities in v1 create a bottom-up saliency map. *Neuron* 73:183–192.
9. Li W, Piëch V, Gilbert CD (2006) Contour saliency in primary visual cortex. *Neuron* 50: 951–962.
10. Burrows BE, Moore T (2009) Influence and limitations of popout in the selection of salient visual stimuli by area V4 neurons. *J Neurosci* 29:15169–15177.
11. Gottlieb JP, Kusunoki M, Goldberg ME (1998) The representation of visual salience in monkey parietal cortex. *Nature* 391:481–484.
12. Bisley JW, Goldberg ME (2010) Attention, intention, and priority in the parietal lobe. *Annu Rev Neurosci* 33:1–21.
13. Arcizet F, Mirpour K, Bisley JW (2011) A pure salience response in posterior parietal cortex. *Cereb Cortex* 21:2498–2506.
14. Thompson KG, Bichot NP (2005) A visual salience map in the primate frontal eye field. *Prog Brain Res* 147:251–262.
15. Purcell BA, Schall JD, Logan GD, Palmeri TJ (2012) From salience to saccades: Multiple-alternative gated stochastic accumulator model of visual search. *J Neurosci* 32:3433–3446.
16. Mysore SP, Asadollahi A, Knudsen EI (2011) Signaling of the strongest stimulus in the owl optic tectum. *J Neurosci* 31:5186–5196.
17. Knudsen EI (2011) Control from below: The role of a midbrain network in spatial attention. *Eur J Neurosci* 33:1961–1972.
18. Asadollahi A, Mysore SP, Knudsen EI (2010) Stimulus-driven competition in a cholinergic midbrain nucleus. *Nat Neurosci* 13:889–895.
19. Krauzlis RJRRJ, Lovejoy LP, Zénon A (2013) Superior colliculus and visual spatial attention. *Annu Rev Neurosci* 36:165–182.
20. Ignashchenkova A, Dicke PW, Haarmeier T, Thier P (2004) Neuron-specific contribution of the superior colliculus to overt and covert shifts of attention. *Nat Neurosci* 7: 56–64.
21. Marino RA, et al. (2012) Linking visual response properties in the superior colliculus to saccade behavior. *Eur J Neurosci* 35:1738–1752.
22. White BJ, Munoz DP (2011) Separate visual signals for saccade initiation during target selection in the primate superior colliculus. *J Neurosci* 31:1570–1578.
23. White BJ, Munoz DP (2011) The superior colliculus. *Oxford Handbook of Eye Movements* (Oxford University Press, Oxford, UK), pp 195–213.
24. Cerkevich CM, Lyon DC, Balaram P, Kaas JH (2014) Distribution of cortical neurons projecting to the superior colliculus in macaque monkeys. *Eye Brain* 2014:121–137.
25. Lock TM, Baizer JS, Bender DB (2003) Distribution of corticotectal cells in macaque. *Exp Brain Res* 151:455–470.
26. White BJ, et al. (2017) Superior colliculus neurons encode a visual saliency map during free viewing of natural dynamic video. *Nat Commun* 8:14263.
27. Itti L (2005) Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. *Vis cogn* 12:1093–1123.
28. Berg DJD, Boehnke SES, Marino RAR, Munoz DP, Itti L (2009) Free viewing of dynamic stimuli by humans and monkeys. *J Vis* 9:1–15.
29. Yoshida M, et al. (2012) Residual attention guidance in blindsight monkeys watching complex natural scenes. *Curr Biol* 22:1429–1434.
30. Treisman AM, Gelade G (1980) A feature-integration theory of attention. *Cognit Psychol* 12:97–136.
31. White BJ, Boehnke SE, Marino RA, Itti L, Munoz DP (2009) Color-related signals in the primate superior colliculus. *J Neurosci* 29:12159–12166.
32. Nurminen L, Angelucci A (2014) Multiple components of surround modulation in primary visual cortex: Multiple neural circuits with multiple functions? *Vision Res* 104: 47–56.
33. Cavanaugh JR, Bair W, Movshon JA (2002) Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J Neurophysiol* 88: 2530–2546.
34. Nassi JJ, Lomber SG, Born RT (2013) Corticocortical feedback contributes to surround suppression in V1 of the alert primate. *J Neurosci* 33:8504–8517.
35. Phongphanphanee P, et al. (2014) Distinct local circuit properties of the superficial and intermediate layers of the rodent superior colliculus. *Eur J Neurosci* 40: 2329–2343.
36. Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 195:215–243.
37. Goldberg ME, Wurtz RH (1972) Activity of superior colliculus in behaving monkey. I. Visual receptive fields of single neurons. *J Neurophysiol* 35:542–559.
38. Davidson RM, Bender DB (1991) Selectivity for relative motion in the monkey superior colliculus. *J Neurophysiol* 65:1115–1133.
39. Marrocco RT, Li RH (1977) Monkey superior colliculus: Properties of single cells and their afferent inputs. *J Neurophysiol* 40:844–860.
40. Lamme VAF (1995) The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* 15:1605–1615.
41. Self MW, van Kerkoerle T, Supèr H, Roelfsema PR (2013) Distinct roles of the cortical layers of area V1 in figure-ground segregation. *Curr Biol* 23:2121–2129.
42. Berman RA, Wurtz RH (2010) Functional identification of a pulvinar path from superior colliculus to cortical area MT. *J Neurosci* 30:6342–6354.
43. Berman RA, Wurtz RH (2011) Signals conveyed in the pulvinar pathway from superior colliculus to cortical area MT. *J Neurosci* 31:373–384.
44. Veale R, Hafed ZM, Yoshida M (2017) How is visual salience computed in the brain? Insights from behaviour, neurobiology and modelling. *Philos Trans R Soc Lond B Biol Sci* 372:20160113.
45. Marino RA, Rodgers CK, Levy R, Munoz DP (2008) Spatial relationships of visuomotor transformations in the superior colliculus map. *J Neurophysiol* 100:2564–2576.
46. Fellows M, Suner S (2009) *Blackrock Microsystems Array Surgical Implant Procedure Training Manual Purpose* (Blackrock Microsystems LLC, Salt Lake City), Rev 5.0, LB-0220.
47. Derrington AM, Krauskopf J, Lennie P (1984) Chromatic mechanisms in lateral geniculate nucleus of macaque. *J Physiol* 357:241–265.
48. Hays AV, Richmond BJ, Optican LM (1982) A UNIX-based multiple-process system for real-time data acquisition and control. *WESCON Conference Proceedings* (Electron Conventions, El Segundo, CA), Vol 2, pp 1–10.
49. 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.
50. Cohen JY, et al. (2010) Cooperation and competition among frontal eye field neurons during visual target selection. *J Neurosci* 30:3227–3238.
51. Thomas NWD, Paré M (2007) Temporal processing of saccade targets in parietal cortex area LIP during visual search. *J Neurophysiol* 97:942–947.