

# Auditory signals evolve from hybrid- to eye-centered coordinates in the primate superior colliculus

Jungah Lee and Jennifer M. Groh

*J Neurophysiol* 108:227-242, 2012. First published 18 April 2012;

doi: 10.1152/jn.00706.2011

## You might find this additional info useful...

---

This article cites 69 articles, 30 of which you can access for free at:

<http://jn.physiology.org/content/108/1/227.full#ref-list-1>

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

<http://jn.physiology.org/content/108/1/227.full>

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

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

---

This information is current as of January 9, 2013.

# Auditory signals evolve from hybrid- to eye-centered coordinates in the primate superior colliculus

Jungah Lee and Jennifer M. Groh

Center for Cognitive Neuroscience, Department of Psychology and Neuroscience, Department of Neurobiology, Duke University, Durham, North Carolina

Submitted 28 July 2011; accepted in final form 16 April 2012

**Lee J, Groh JM.** Auditory signals evolve from hybrid- to eye-centered coordinates in the primate superior colliculus. *J Neurophysiol* 108: 227–242, 2012. First published April 18, 2012; doi:10.1152/jn.00706.2011.—Visual and auditory spatial signals initially arise in different reference frames. It has been postulated that auditory signals are translated from a head-centered to an eye-centered frame of reference compatible with the visual spatial maps, but, to date, only various forms of hybrid reference frames for sound have been identified. Here, we show that the auditory representation of space in the superior colliculus involves a hybrid reference frame immediately after the sound onset but evolves to become predominantly eye centered, and more similar to the visual representation, by the time of a saccade to that sound. Specifically, during the first 500 ms after the sound onset, auditory response patterns ( $N = 103$ ) were usually neither head nor eye centered: 64% of neurons showed such a hybrid pattern, whereas 29% were more eye centered and 8% were more head centered. This differed from the pattern observed for visual targets ( $N = 156$ ): 86% were eye centered, <1% were head centered, and only 13% exhibited a hybrid of both reference frames. For auditory-evoked activity observed within 20 ms of the saccade ( $N = 154$ ), the proportion of eye-centered response patterns increased to 69%, whereas the hybrid and head-centered response patterns dropped to 30% and <1%, respectively. This pattern approached, although did not quite reach, that observed for saccade-related activity for visual targets: 89% were eye centered, 11% were hybrid, and <1% were head centered ( $N = 162$ ). The plainly eye-centered visual response patterns and predominantly eye-centered auditory motor response patterns lie in marked contrast to our previous study of the intraparietal cortex, where both visual and auditory sensory and motor-related activity used a predominantly hybrid reference frame (Mullette-Gillman et al. 2005, 2009). Our present findings indicate that auditory signals are ultimately translated into a reference frame roughly similar to that used for vision, but suggest that such signals might emerge only in motor areas responsible for directing gaze to visual and auditory stimuli.

saccade; midbrain; reference frame

VISUAL-AUDITORY INTEGRATION poses a computational challenge for the brain: visual location is determined from the location of active photoreceptors on the retina (an eye-centered reference frame), whereas auditory location is determined from cues that depend on sound location with respect to the head and ears (a head-centered reference frame). Resolving this discrepancy is essential for knowing whether sights and sounds are at the same location in space and for guiding movements independent of modality. Because vision is more spatially accurate than, and can dominate, auditory spatial perception (Knudsen and Knudsen 1985; Recanzone 1998; Woods and Recanzone 2004; Witten and Knudsen 2005; Kopco et al. 2009), it has been

suspected that auditory signals might be translated into an eye-centered reference frame. However, despite extensive efforts to identify a predominantly eye-centered reference frame<sup>1</sup> (Jay and Sparks 1984, 1987a, 1987b; Hartline et al. 1995; Peck et al. 1995; Groh et al. 2001; Zella et al. 2001; Werner-Reiss et al. 2003; Populin et al. 2004; Mullette-Gillman et al. 2005; Porter et al. 2006; Mullette-Gillman et al. 2009), the existing evidence reveals representations in which head- and eye-centered coordinates both appear to contribute to response patterns. Even studies in the parietal cortex, thought to play a role in coordinate transformations (Andersen et al. 1993; Colby and Goldberg 1999), have found only complex representations of visual and auditory signals that are not in a unique coordinate framework (Stricanne et al. 1996; Mullette-Gillman et al. 2005; Schlack et al. 2005; Mullette-Gillman et al. 2009).

The superior colliculus (SC) is thought to play a role in multisensory integration, spatial attention, and control of eye movements (e.g. Wurtz and Goldberg 1971; Robinson 1972; Schiller and Stryker 1972; Sparks et al. 1976; Sparks and Mays 1983; Lee et al. 1988; Waitzman et al. 1988; Freedman et al. 1996; Kustov and Robinson 1996; Freedman and Sparks 1997; Carello and Krauzlis 2004; Muller et al. 2005). Several studies have shown evidence that the auditory receptive fields of SC neurons shift when the eyes move, but they do not shift by the exact amount of the eye movement, as would be expected of an eye-centered frame of reference (Jay and Sparks 1984, 1987a, 1987b; Zella et al. 2001; Populin et al. 2004). By implication, these findings suggested that the receptive fields for visual and auditory signals could not maintain perfect alignment across different initial eye positions as the eyes move. However, saccades to visual and auditory stimuli do not show much indication of such a discrepancy (Metzger et al. 2004).

A possible explanation lies in the fact that these studies investigated only sensory activity, i.e., activity occurring time locked to the target onset and well before the movement (Jay and Sparks 1984, 1987b), or they used a spike-counting window that likely included both sensory and activity occurring around the time of the movement (Populin et al. 2004). In the intermediate and deep layers of the SC, neurons exhibit vigorous bursts time locked to saccades (Goldberg and Wurtz 1972) forming a motor map that encodes a gaze shift with a particular direction and amplitude for a given location on the map. This map is traditionally assumed to be a predominantly eye-centered map of saccade vectors (Robinson 1972; Schiller

<sup>1</sup> That is, one in which spatial response functions are stable in position with respect to the eyes. Eye position modulations are allowed under this definition provided that they affect the gain of neural responses but not what locations the neuron responds to.

Address for reprint requests and other correspondence: J. Lee, LSRC, Rm. B245, Box 90999, Duke Univ., Durham, NC 27708 (e-mail: jungah.lee@duke.edu).

and Stryker 1972; Freedman et al. 1996; Freedman and Sparks 1997; Klier et al. 2001) with some modulation of activity patterns or stimulation-evoked eye movements by eye position (Van Opstal et al. 1995, Campos et al. 2006; Groh 2011). Importantly, the reference frame of this motor-related activity has not previously been investigated when the saccade target is auditory. This leaves open the possibility that an eye-centered auditory representation might reside in these activity patterns.

Here, we found that the visual and auditory sensory signals in the primate SC initially arise in different reference frames: the sensory activity is eye centered for visual targets and hybrid for auditory targets consistent with a previous work (Jay and Sparks 1987). However, by the time of the motor readout, this discrepancy has been reduced. The auditory signals reflected in motor-related activity are predominantly eye centered but not quite as strongly so as visual signals. The reduced discrepancy between multisensory reference frames likely facilitates the generation of a reasonably accurate saccade signal by the SC regardless of target modality.

## MATERIALS AND METHODS

### *Animal Preparation*

Two adult rhesus monkeys (*Macaca mulatta*; 1 male and 1 female) served as subjects for this experiment. All animal procedures were conducted in accordance with the principles of laboratory animal care of the National Institutes of Health (National Institutes of Health Pub. No. 86-23, Revised 1985) and were approved by the Institutional Animal Care and Use Committee of Duke University. Surgical procedures were conducted using suitable anesthesia and analgesics. Subjects underwent sterile surgery for the implantation of a head post holder and eye coil (Robinson 1963; Judge et al. 1980). After behavioral training, a recording cylinder was implanted over the left [*monkey W* (male)] or right [*monkey P* (female)] SC using stereotactic techniques. The location of the cylinder over the SC was verified with MRI scans at the Duke Center for Advanced Magnetic Resonance Development.

### *Experimental Setup*

All experimental and behavioral training sessions were conducted in a dimly illuminated sound-attenuation room (IAC, single walled) lined on the inside with sound-absorbing foam (Sonex PaintedOne). Dim illumination helped prevent nystagmus and thus improved saccade performance generally but did not provide any visual cues that were useful for performance of the task because there was no association between the visual scene and which sound target was used on any given trial. Behavioral performance on both visual and auditory trials was comparable with our previous studies using similar methods in complete darkness (Metzger et al. 2004; Mullette-Gillman et al. 2005, 2009).

Stimulus presentation and data collection were controlled through GramakIn 2.0 software. Eye position was sampled at 500 Hz. Velocity criteria to detect saccades were 20°/s for both saccade onset and offset and were implemented using the EyeMove software program (written by Kathy Pearson). All subsequent analysis was performed in Matlab 7.1 (Mathworks Software).

Sensory targets were presented from a stimulus array that was 58 in. in front of the monkey. The array contained nine speakers (model TXO25V2, Audax) with a light-emitting diode (LED) attached to each speaker's face (Fig. 1A). The speakers were placed from 24° left to 24° right of the monkey in 6° increments at an elevation of 0°. Three visual fixation points were spaced 12° apart horizontally and were presented at a suitable elevation above or below the horizontal row of targets (Fig. 1A; see more details below in *Recording Procedure and Strategy*). Auditory targets were bandpass white noise bursts (500 Hz

to 18 kHz, rise time of 10 ms) at  $50 \pm 2$  dB sound pressure level. The luminance of each LED was  $\sim 26.4$  cd/m<sup>2</sup>.

### *Recording Procedure and Strategy*

At the start of each recording sessions, a stainless steel guide tube was manually advanced through the dura. Monkeys performed the overlap saccade task (described below in *Behavioral Task*) while a tungsten electrode (FHC, impedance between 2 and 4.5 M $\Omega$  at 1 kHz) was extended further into the brain with an oil hydraulic pulse micropositioner (Narishige-group.com). Extracellular neural signals were amplified, and action potentials from single neurons were isolated using a PLEXON system (Sort Client software, PLEXON). The time of occurrence of each action potential was stored for offline analysis.

We focused on recording SC neurons in the intermediate or deep layers. About 1 mm after encountering cells responsive to visual stimuli (the superficial layers), we reached cells showing visual or/and auditory responses and/or motor activity related to eye movements (the intermediate and/or deep layers). We selected any such neuron for further study. Once a neuron was isolated, we first qualitatively determined the elevation of the receptive or movement field while subjects performed the overlap task. The vertical position of the fixation LEDs was then chosen to position the (fixed) row of visual and auditory saccade targets at or near that preferred elevation of the neuron under study for the remainder of the experiment. We used three horizontal fixations ( $-12^\circ$ ,  $0^\circ$ , and  $12^\circ$  from center) and the full set of target locations as shown in Fig. 1A. [We note that choosing the fixation elevation on a neuron-by-neuron basis could, in principle, introduce a selection bias for eye-centered response patterns, because the sampled range of eye-centered space (9 head-fixed targets with respect to 3 fixation positions with tailored elevation) was larger than the sampled range of head-centered space (9 head-fixed targets only).]

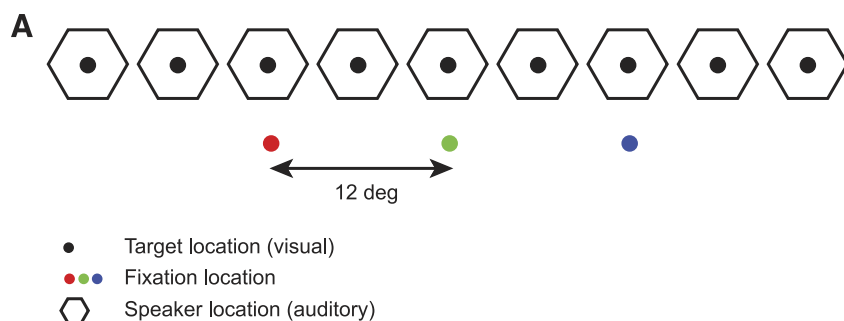
We examined each SC neuron's sensitivity to the location of auditory and visual targets and the reference frame of sensory and motor responses. The location of horizontal fixation, the location of the target, and the target modality varied randomly on a trial-by-trial basis. Data were collected as long as the neuron was well isolated and the monkey performed the task. On average, we collected  $11.2 \pm 5.4$  (mean  $\pm$  SD) successful trials per task condition (fixation location  $\times$  target location  $\times$  target modality). In our final data set of 180 cells, there were on average 604 ( $\pm 164$ ) successful trials for each neuron.

As we have previously shown that pinna movements in monkeys are small and uncorrelated with eye movements (Groh et al. 2001; Werner-Reiss et al. 2003), we did not monitor pinna movements in the present study.

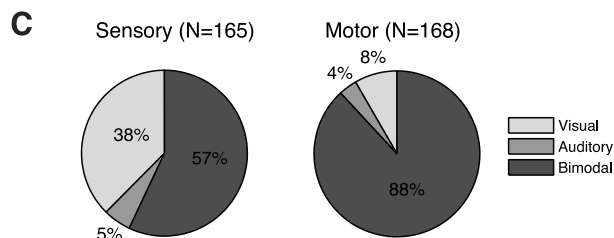
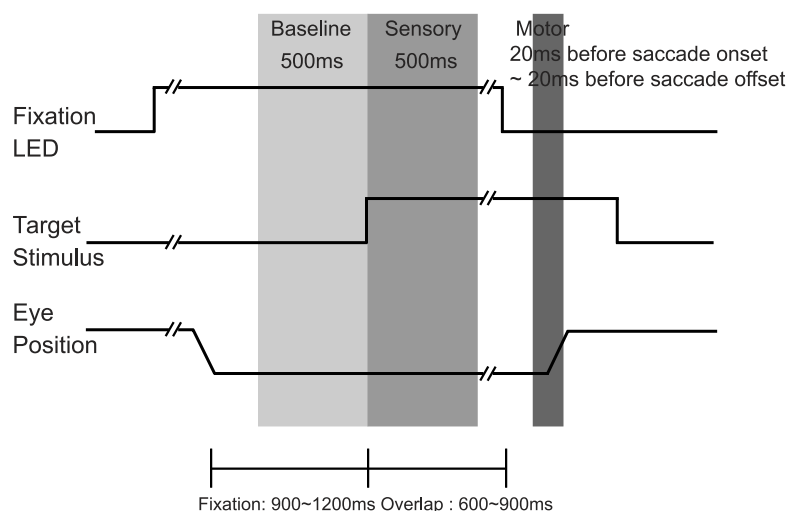
### *Behavioral Task*

Two monkeys performed an overlap saccade task (Fig. 1B) to auditory and visual targets in which all conditions were randomly interleaved. During the task, 900–1200 ms after a visual stimulus was fixated, a sensory target (either auditory or visual) was presented. Both fixation and target stayed on for 600–900 ms (the overlap period), after which the fixation light was extinguished and the monkey had 500 ms to shift its gaze to within 6° of the location of the target. This allowed us to dissociate sensory-related activity from motor-related activity. Animals received a juice or water reward for a successful trial. Incorrect trials were excluded from the analysis except where specifically noted (see Fig. 7).

Animals were initially trained with visual targets, and trials with auditory targets were then introduced. Monkeys generally make saccades to sounds readily, especially when the sounds are novel or unexpected, as they are when first presented to a monkey accustomed to a silent recording booth (Metzger et al. 2004). On the first several days of auditory training, visual targets were presented at the location of the sounds if the monkey did not make a saccade to the sound's location within several hundred milliseconds. These visual reinforcers



**B** Overlap saccade task



were only presented at the most peripheral target locations ( $\pm 24^\circ$ ) and only for the first several days of auditory training, after which they were no longer necessary.

**Data Analysis**

To analyze activity, action potentials were counted during several different time periods. The baseline period was the 500-ms period before the target onset, and the sensory-related target period was the 500-ms period after the target onset (i.e., ending a minimum of 100 ms before the offset of the fixation light for the shortest overlap periods; Fig. 1B). The neural responses during the 500 ms after the visual target onset are henceforth referred to as “visual sensory” and those after auditory target onset are referred to as “auditory sensory” (Fig. 1B).

The motor-related activity period was synchronized with the saccade, beginning 20 ms before the saccade onset and ending 20 ms before the saccade offset, and is referred to as “visual motor” or “auditory motor” as appropriate. This time period was chosen based on the minimum latency of stimulation-evoked movements in the SC (Robinson 1972; Guitton et al. 1980; Schlag-Rey et al. 1989; Kustov and Robinson 1995) and captures the peak of the saccade-related

Fig. 1. Experimental design. *A*: spatial layout of the targets. Initial eye fixation targets [red light-emitting diode (LED)] were located  $12^\circ$  left (red),  $0^\circ$  (green), and  $12^\circ$  right (blue) at an elevation chosen based on the spatial sensitivity of the neuron under study. Targets were either auditory (white noise burst) or visual (green LED) and were presented from a stimulus array of nine speakers each with an LED attached to its face. The speakers were spaced from  $24^\circ$  left to  $24^\circ$  right with  $6^\circ$  intervals at an elevation of  $0^\circ$  with respect to the animal’s head. *B*: events and spike counting time periods of the behavioral task. The baseline period was 500 ms before the target onset, the target period was 0–500 ms after the target onset, and the saccade-related period began 20 ms before the saccade onset and ended 20 ms before the saccade offset. *C*: pie chart showing the percentage of neurons showing responses for each modality and time period. Data are from Table 1.

burst. Since saccade duration is not a constant but depends on saccade amplitude (Bahill et al. 1975) and target modality (Jay and Sparks 1990; Groh and Sparks 1996; Yao and Peck 1997; John and Van Lieshout 2000) the motor activity period varied with these factors as well. The average saccade duration was  $48.5 (\pm 17.7 \text{ SD})$  ms for visual saccades and  $53.8 (\pm 19.8 \text{ SD})$  ms for auditory saccades, which are a little slower than for visual saccades. Unless otherwise noted, all data analyses used the average firing rate during the activity period so as to control for the variation in duration.

**Responsiveness and Spatial Tuning**

Neurons were included for the relevant population analysis if 1) they responded significantly [one-tailed *t*-test comparing baseline activity and sensory or motor-related activity on visual or auditory trials ( $P < 0.05$ )] and 2) their firing rate was modulated by target location as assessed by ANOVA [main effect for target location, defined either with respect to the head or eyes ( $P < 0.05$ ) or an interaction effect ( $P < 0.05$ ) between the target location and fixation position]. To ensure that these ANOVAs were not biased in favor of one reference frame, only the target locations that existed in both

reference frames were included (i.e.,  $\pm 12$ , 6, and  $0^\circ$  with respect to the head or eyes; see Mullette-Gillman et al. 2005 for details).<sup>2</sup>

### Quantitative Analyses of Reference Frame

**Eye-centered versus head-centered correlation coefficient.** The reference frame was assessed by determining the correlation coefficient between the mean responses evoked by this set of target locations across different fixation positions (Mullette-Gillman et al. 2005; Porter et al. 2006; Mullette-Gillman et al. 2009). For most analyses (except for those shown in Figs. 7, D–G, and 10 and portions of the data shown in Figs. 6 and 9), we used saccade amplitude and saccade end point rather than target location per se because these measures incorporate the animal's estimate of sound location. Saccade amplitudes and end points were binned in  $6^\circ$  increments matching the target locations.

The equation for this calculation was as follows:

Correlation coefficient =

$$\frac{1}{2} \left( \frac{(\vec{R}_{l,i} - \vec{R}) \times (\vec{R}_{c,i} - \vec{R})}{\|\vec{R}_{l,i} - \vec{R}\| \|\vec{R}_{c,i} - \vec{R}\|} + \frac{(\vec{R}_{r,i} - \vec{R}) \times (\vec{R}_{c,i} - \vec{R})}{\|\vec{R}_{r,i} - \vec{R}\| \|\vec{R}_{c,i} - \vec{R}\|} \right) \quad (1)$$

where  $\vec{R}_{l,i}$ ,  $\vec{R}_{r,i}$  and  $\vec{R}_{c,i}$  are the vectors of average responses of the neuron for saccade amplitudes or end points at location  $i$  when the monkey's eyes were fixated at the left (l), right (r), or center (c).  $\vec{R}$  is the mean response across all tested target locations and eye positions.

The correlation coefficient was obtained in the two response periods (sensory and motor period) with two reference frames (eye centered and head centered) and for two modalities (visual and auditory). To be considered eye centered, the correlation coefficient had to be significantly higher across fixation positions as a function of the saccade amplitude or eye-centered target location than as a function of the saccade end point or head-centered location of the target. The opposite pattern was considered head centered.

The correlation coefficient could range from  $-1$  to  $1$ , with a value of  $1$  indicating that the response function was perfectly aligned in either an eye or a head-centered reference frame. A value of  $0$  means that the response functions were not correlated each other, and a value of  $-1$  means that they were negatively correlated. We calculated the variance of the correlation coefficient using a bootstrap analysis (100 iterations of 80% of data for each target location/eye position combination). This allowed us to estimate the 95% confidence level ( $\pm 1.96$  SD) around the correlation coefficient.

<sup>2</sup> Targets were presented between  $+24$  and  $-24^\circ$  with respect to the room and head. So, the head-centered space was sampled between  $+24$  and  $-24^\circ$ . The fixation positions ranged between  $+12$  and  $-12^\circ$ . So, eye-centered space was sampled from  $-36$  to  $12^\circ$  from the right eye position, from  $-24$  to  $24^\circ$  from the center, and from  $-12$  to  $36^\circ$  from the left eye position. The intersection of these four ranges of locations was  $-12$  to  $12^\circ$ .

Since the monkeys had their heads restrained in the present study, we could not dissociate a head-centered reference frame from a body- or world-centered reference frame. Note also that this correlation analysis is likely to be relatively invariant to changes in gain, such as the eye position modulations that have been observed with recording and stimulation studies of the SC (Van Opstal et al. 1995; Campos et al. 2006; Groh 2011); thus, our correlation analysis does not speak to the presence or role of such eye position modulations.

**Displacement index.** We also performed a cross-correlation analysis to estimate the degree of shift, if any, of the response function. This analysis was based on the method described by d'Avossa et al. (2007), which has been used to study the reference frames of visual or tactile receptive fields (Avillac et al. 2005; Gardner et al. 2008). Briefly, we first interpolated response functions linearly in  $0.01^\circ$  steps. We then calculated correlation coefficients between any two response functions involving different fixation positions while systematically shifting one function relative to the other in  $0.01^\circ$  steps. The shift that produced the highest correlation coefficient was normalized with respect to the separation in eye position, yielding a displacement index (DI). Like the eye- versus head-centered correlation coefficient analysis, the DI is insensitive to variation in the level of activity associated with any gain field effect by eye position but differs in that it produces an estimate of the shift that yields the best alignment between response functions. The DI equals  $0$  when the response functions do not move with the eyes: a head-centered reference frame. A DI value of  $1$  corresponds to response functions that shift by the same amount as the eyes move: an eye-centered reference frame.

## RESULTS

We recorded 180 neurons (100 neurons from *monkey W*, left SC; and 80 neurons from *monkey P*, right SC). Most neurons exhibited motor-related bursts for at least one target modality (168 of 180 neurons, 93.3%; Table 1), and most of these were active for both target modalities ( $N = 148$  of 180 neurons, 82.2%, or 88% of the 168 neurons; Fig. 1C, right). Sensory-related responses were also quite prevalent ( $N = 165$  of 180 neurons, 92%), but a smaller proportion of the neurons were bimodal for this response period (94 of 180 neurons, 52%, or 57% of the 165 neurons; Fig. 1C, left). Among neurons showing unimodal sensory responses, the majority of these were visual (38% visual vs. 5% auditory; Fig. 1C, left). Nearly all neurons with auditory sensory activity were also active for the other three categories, for a total of 51% ( $N = 91$  of 180 neurons; Table 1) of SC neurons showing spatially modulated activity to both visual and auditory stimuli during both sensory and motor periods.

### Temporal Response Profile

The combination of sensory and motor-related activity seen in many of these neurons includes response patterns previously

Table 1. Statistical analysis of sensitivity to target location during the sensory and motor periods

	Visual		Auditory		Both	
	N	Total, %	N	Total, %	N	Total, %
<b>Sensory period</b>						
A. ANOVA: head-centered target $\times$ eye position, main or interaction for target	151	83.9	101	56.1	93	51.7
B. ANOVA: eye-centered target $\times$ eye position, main or interaction for target	150	83.3	80	44.4	69	38.3
C. A or B	156	86.7	103	57.2	94	52.2
<b>Motor period</b>						
D. ANOVA: head-centered target $\times$ eye position, main or interaction for target	161	89.4	154	85.6	145	80.6
E. ANOVA: eye-centered target $\times$ eye position, main or interaction for target	160	88.9	144	80.0	139	72.2
F. D or E	162	90.0	154	85.6	148	82.2
G. C and F	152	84.4	96	53.3	91	50.6

N, number of neurons ( $N = 180$  total).

described as visual-motor (Wurtz and Albano 1980), prelude bursting (Glimcher and Sparks 1992), or buildup neurons (Munoz and Wurtz 1995). An individual example neuron that illustrates activity beginning at the target onset and continuing through the movement for both visual and auditory targets is shown in Fig. 2. For visual targets (Fig. 2, *A* and *B*), the neuron began responding  $\sim 50$  ms after the target onset and reached a peak response at  $\sim 90$  ms (Fig. 2*A*). This neuron also had distinct motor-related activity, reaching its second peak  $\sim 20$  ms before the saccade onset (Fig. 2*B*). For auditory targets (Fig. 2, *C* and *D*), there was no sharp initial sensory-related peak but rather a slowly increasing activity level beginning soon after the target onset (Fig. 2*C*). The motor-related burst for auditory targets was similar to that observed for visual targets (Fig. 2*D*). The response amplitude was the largest when the targets were  $12\text{--}24^\circ$  right of the midline, suggesting the location of the receptive field and movement field of this neuron.

Figure 3 shows the temporal profile of the discharge patterns of the population of neurons. Figure 3 shows a population perievent time histogram (PETH) that is aligned at the target onset (*left* portion of the *x*-axis) and saccade onset (*right* portion of the *x*-axis). The PETHs were constructed from the activity of neurons that had significant sensory and motor-related activity (by one-tailed *t*-test and ANOVA; see Table 1 and MATERIALS AND METHODS) to either visual ( $n = 152$ ; Fig. 3*A*) or auditory targets ( $n = 96$ ; Fig. 3*B*).

#### Reference Frames of Example Neurons

To test which frame of reference best describes the response patterns of SC neurons, we evaluated the alignment of the responses as a function of saccade end point or saccade amplitude (i.e., head- vs. eye-centered coordinates) across

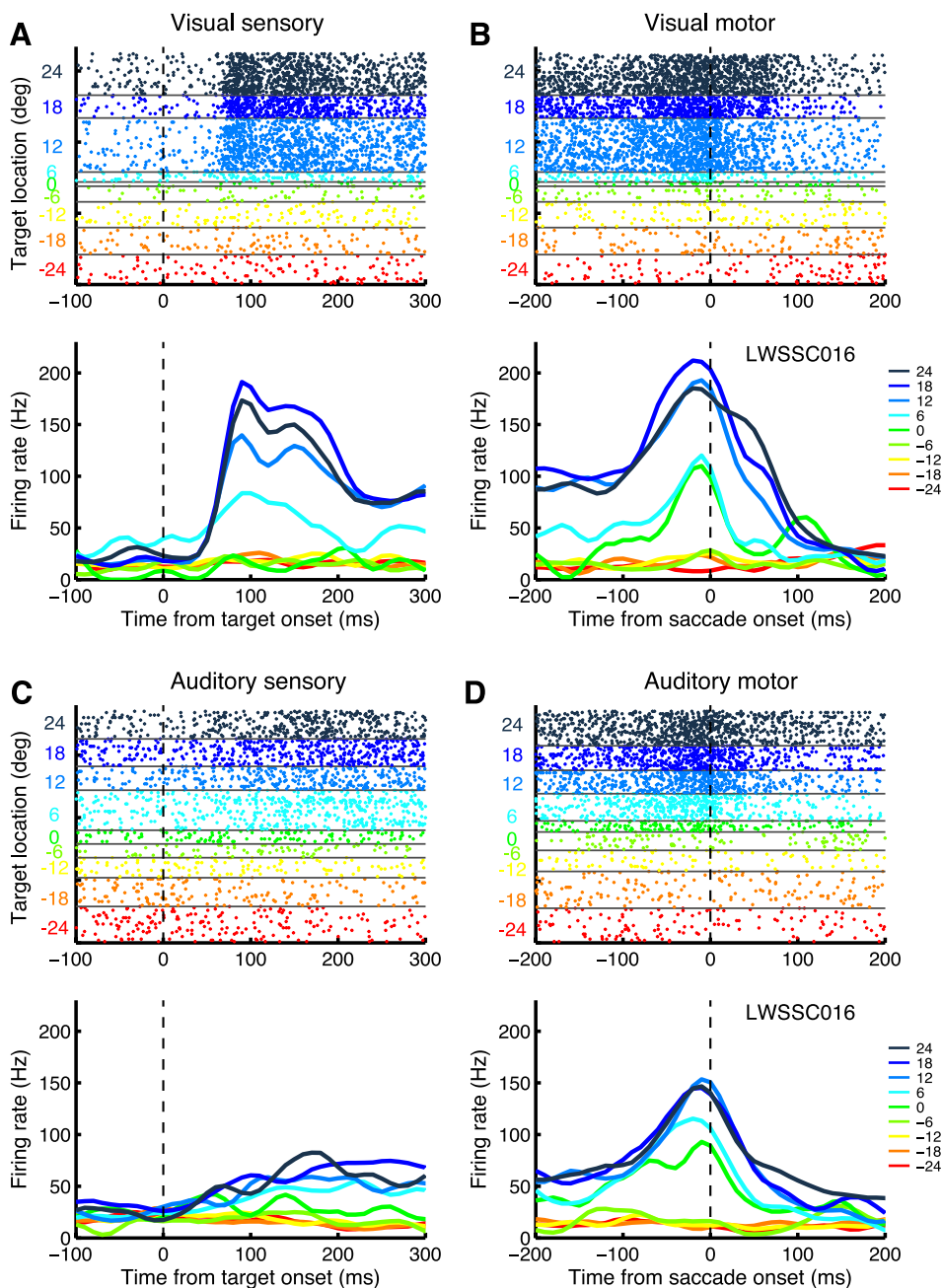
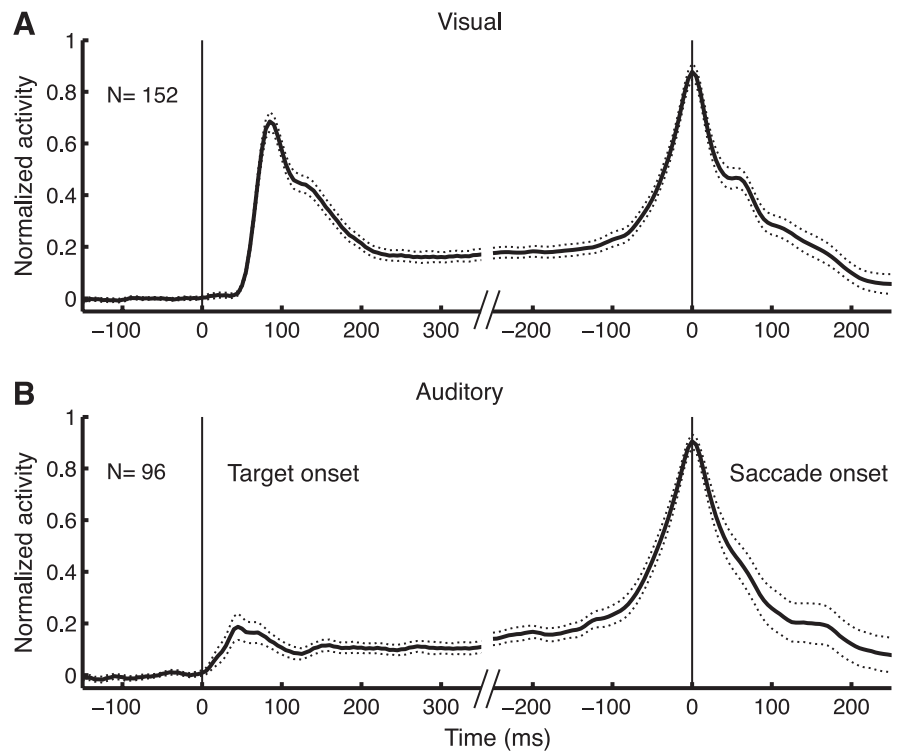


Fig. 2. Temporal profile of visual and auditory responses in an example neuron aligned on the target onset versus saccade onset. The neuron exhibited robust responses to both visual (*A* and *B*) and auditory (*C* and *D*) stimuli during the sensory (*A* and *C*) and motor (*B* and *D*) periods. *Top*: raster plots as a function of target location from the central fixation position. Data are aligned relative to the target onset (*left*) or saccade onset (*right*). The unequal spacing across the sections of the raster plot reflects different numbers of trials due to random selection with replacement of target location and fixation and due to exclusion of incorrectly performed trials. *Bottom*: perievent time histograms (PETHs) for data collected at each target location (central fixation position). The PETHs were smoothed by convolution with a filter with points [1/9 2/9 1/3 2/9 1/9]. The bin size was 10 ms.

Fig. 3. Population PETH of visual (A) and auditory (B) activity aligned at the target onset and saccade onset (vertical lines). Neurons with statistically significant activity for both sensory and motor-related periods (by one-tailed *t*-test and ANOVA; see MATERIALS AND METHODS) to either visual ( $N = 152$ ) or auditory ( $N = 96$ ) targets were included. Each contributing neuron's individual PETH was constructed for visual and auditory trials, pooling across target locations and fixation positions. The individual PETHs were made with 10-ms bins and smoothed with a triangular filter [1/9 2/9 1/3 2/9 1/9]. These PETHs were then normalized to the firing rate in their highest bins, and the firing rate during the baseline (500 ms before the target onset) was subtracted: (bin height - baseline)/(maximum bin height - baseline). They were then averaged to produce the population PETH. The dotted line indicates  $\pm$ SE. Note that because each contributing neuron was normalized individually and did not necessarily reach a peak in the same temporal bin as other neurons, the population PETH does not reach a value of 1.



different fixation positions. Figure 4 shows the activity patterns of an example neuron. For visual targets, response functions involving different initial fixation positions were much better aligned when plotted as a function of saccade amplitude (Fig. 4A, right) than the saccade end point (Fig. 4A, left). This was true of both sensory and motor response periods (Fig. 4A, top vs. bottom rows).

When the target was auditory, the sensory responses showed partial compensation for eye position (Fig. 4B): responses were partially aligned in both head- and eye-centered reference frames, similar to previous reports involving similar conditions (Jay and Sparks 1984, 1987b). By the time of the motor burst, however, the alignment of the response patterns improved in an eye-centered reference frame, suggesting an evolution of signals from hybrid toward more eye-centered coordinates.

#### Population Analysis

The patterns shown in this example neuron held true at the population level as well. Figure 5 graphically displays the reference frame by plotting the head-centered correlation coefficient versus the eye-centered correlation coefficient for each neuron by modality and response period. The majority of visually responsive SC neurons showed statistically significant eye-centered activity patterns for both the sensory and motor periods (Fig. 5, A and B): >86% of the neurons had correlation coefficients that were significantly higher in an eye-centered frame of reference than a head-centered frame of reference (86%,  $n = 134$  of 156, for sensory and 89%,  $n = 144$  of 162, for motor-related activity; orange symbols on graphs and orange areas on pie charts). Neurons were classified as exhibiting hybrid response patterns if the head- and eye-centered correlation coefficients were comparable, i.e., if their 95% confidence intervals overlapped (gray symbols on graphs and gray areas on pie charts). The proportion of hybrid visual

response patterns was minor and was similar for both the sensory and motor periods (13% and 11%, respectively). This proportion likely reflects a floor level of unclassifiable reference frame patterns that can occur due to noise or insufficient sampling of the receptive field.

Auditory sensory and auditory motor responses differed from each other. For auditory targets, the reference frame was mixed during the sensory period (Fig. 5C). Nearly two-thirds (64%) of the neurons could not be classified as eye or head centered (gray points locating along the line of slope 1), whereas 8% were classified as head centered (blue) and 29% were classified as eye centered (orange). This pattern suggests reasonably similar contributions of both eye- and head-centered reference frames (gray crosses and gray areas on pie charts). By the time the motor burst was generated, however, eye-centered neurons constituted a strong majority (69%), hybrid response patterns fell to 30%, and head-centered response patterns had nearly disappeared (<1%; Fig. 5D). In short, by the time a motor command was generated, the reference frame of the auditory responses (Fig. 5D) had evolved to become substantially more eye centered than is initially the case. As a corollary to this evolution, the discrepancy between visual and auditory coordinate systems had diminished by the time of the movement.

#### Time Course of the Reference Frame Change

In *Population Analysis*, we compared two discrete time periods: one immediately after the onset of the sensory stimulus and the other beginning slightly before the saccade. That analysis suggested that a transition in reference frame occurs between those two time periods but does not establish when this transition occurs or whether the transition is time locked to the target onset or saccade motor command.

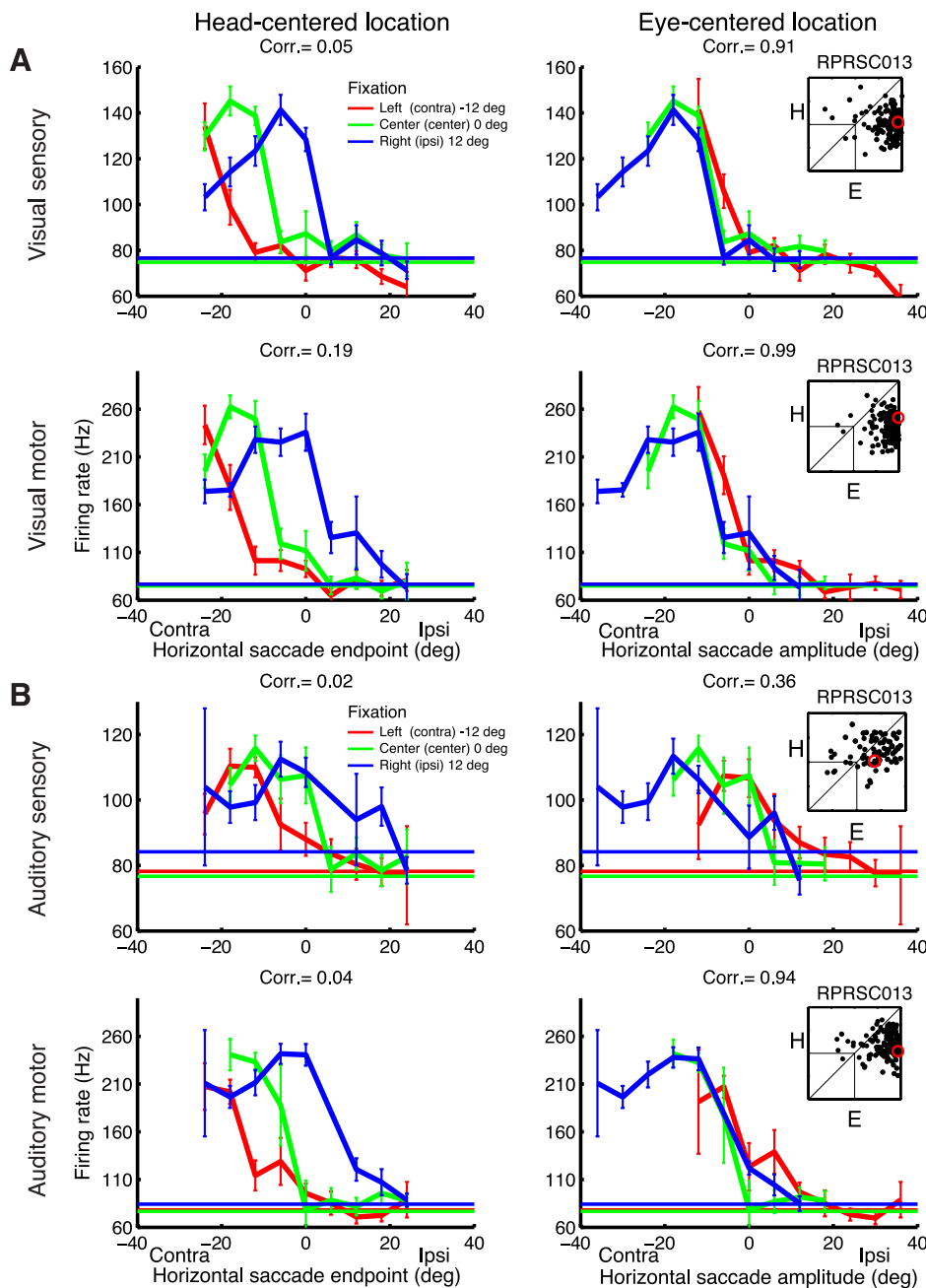


Fig. 4. Example neuron showing eye-centered visual sensory, visual motor, and auditory motor activity patterns. The auditory sensory activity was neither head nor eye centered. Responses for visual (A) and auditory (B) stimuli during the sensory (top) and motor (bottom) periods are shown. The response curves show activity as a function of the end point (left) or amplitude (right) of the saccade made by the animal in response to the target. For perfectly accurate saccades, these measures correspond to target location defined with respect to the head or eyes, respectively. Data were binned in  $6^\circ$  bins, and each point and error bar reflects the mean  $\pm$  SE for that bin. Red, green, and blue curves correspond to trials involving the contralateral, center, and ipsilateral initial fixation positions, respectively. The horizontal lines in each graph indicate the mean activity during the baseline period (500 ms before the target onset) for each fixation. The correlation coefficient for the quantitative analysis of reference frame (see RESULTS) is shown on the top of each graph. The insets indicate where this cell's head-center (H) and eye-centered (E) correlation values lie with respect to the population shown in Fig. 5 (red circle).

To investigate when the transformation occurs, we analyzed the time courses of the reference frame for the subset of 152 neurons (84.4% of the total population of 180 neurons) with visual activity and 96 neurons (53.3% of the 180 neurons) with auditory activity during both sensory and motor periods. We calculated the correlation coefficient for a 100-ms window stepping in increments of 50 ms during the time period between the target onset and saccade onset. The population average values of the eye- and head-centered correlation coefficients are shown as a function of time with respect to the target onset (Fig. 6, A and C) or saccade onset (Fig. 6, B and D).

For visual targets (Fig. 6, A and B), the average eye-centered correlation coefficient was high throughout the time interval between the target onset and saccade, averaging  $\sim 0.60$  and peaking at  $\sim 0.85$  at the time of the saccade (Fig. 6B, solid lines). The average head-centered correlation coefficient stayed

low throughout. When the target was auditory (Fig. 6, C and D), the average head- and eye-centered correlation coefficients were initially about equal,  $\sim 0.1$  to  $0.2$ . Starting at  $\sim 200$  ms before the saccade, the eye-centered correlation coefficient increased significantly, peaking at  $\sim 0.76$  at the saccade onset. The eye-centered correlation coefficient was significantly greater than the head-centered correlation coefficient for  $\sim 300$  ms centered around the saccade onset (Fig. 6, C and D, gray box and solid black symbols). The same basic pattern was evident regardless of whether we used saccade amplitude/end point as the measure of the subject's localization of the targets (solid lines) versus the location of the targets themselves (dotted lines).

This increase in the eye-centered correlation coefficient was better time locked to the saccade than to the target. (The target was presented for 600–900 ms before the fixation light offset



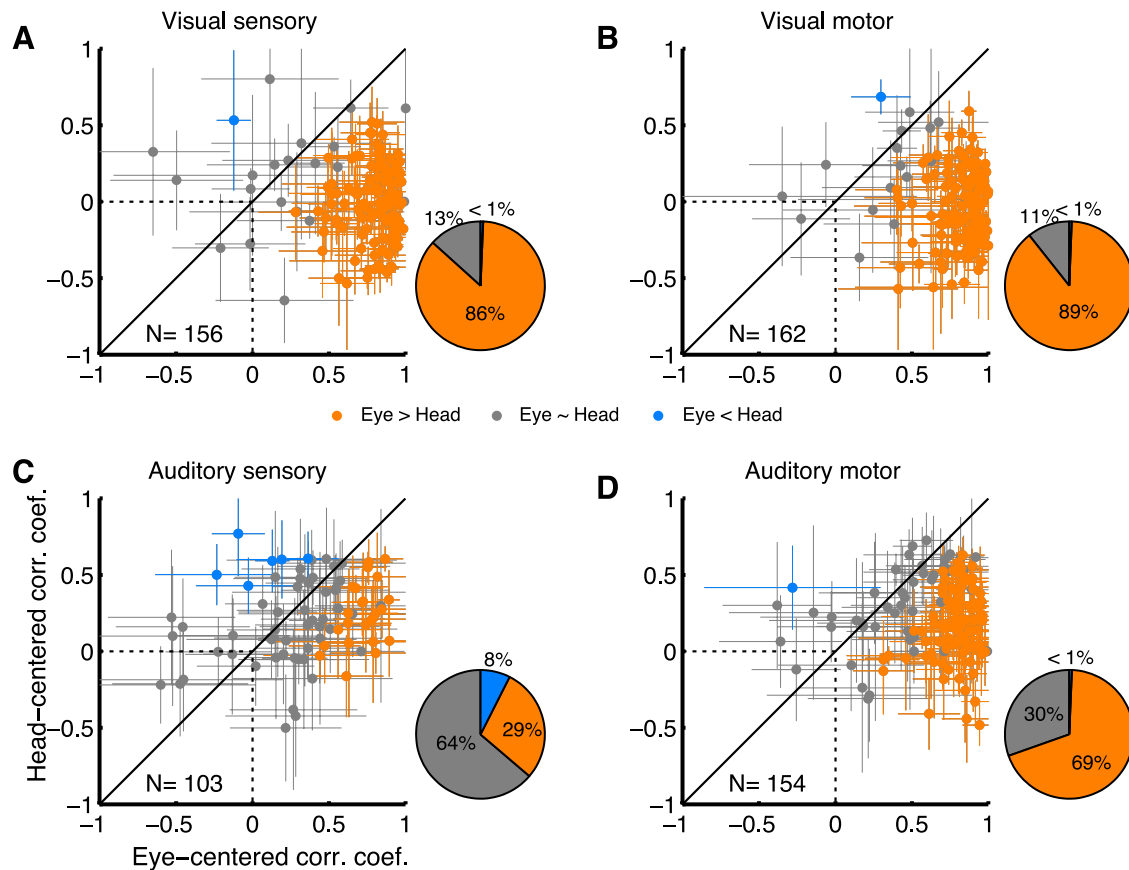


Fig. 5. Across the population of superior colliculus (SC) neurons, visual sensory, visual motor, and auditory motor activity patterns are predominantly eye centered, whereas auditory sensory activity is hybrid. These graphs plot the head-centered (y-axis) versus eye-centered (x-axis) correlation coefficients of visual (A and B) and auditory (C and D) responses in the population of neurons during sensory (A and C) and motor (B and D) periods. The error bars (crosses) indicate the 95% confidence intervals (CIs) using the bootstrap analysis. The diagonal line is a line of slope 1, and the dashed lines indicate values of 0 on the x- and y-axes. Neurons whose eye-centered correlation coefficients were significantly greater than 0 and significantly higher than their head-centered correlation coefficients (meaning that the eye-centered correlation coefficient's CIs did not include the head-centered correlation coefficient) lie below the line of slope 1 and to the right of 0 and are indicated in orange. Blue data points indicate significantly larger head-centered than eye-centered correlation coefficients and lie in the corresponding region above the line and above 0. The remainder of the data points are colored gray, indicating that neither eye-centered nor head-centered reference frame is predominant. The pie charts summarize the percentage of neurons in each reference frame category (orange: eye centered, blue: head centered, and gray: hybrid).

cued the saccade, for a total target-saccade intervals of  $\sim 800$ – $1,100$  ms.) The highest values of the eye-centered correlation coefficients ( $\pm$ SD) were observed when the data were aligned on the saccade rather than on the target onset (saccade onset alignment:  $0.76 \pm 0.26$  vs. target onset alignment:  $0.56 \pm 0.31$ ;  $P < 0.01$  by two-sample  $t$ -test). This suggests that the conversion of auditory signals to an eye-centered reference frame was related specifically to the saccade and not a merely delayed response to the target onset.

#### Actual Target Location Versus Saccade Metrics: Impact on the Observed Reference Frame

In the results shown in Figs. 4–6, we used saccade amplitude or end point as a proxy for the actual target location in eye- or head-centered coordinates. This is useful because it takes into account both systematic and trial-by-trial variation in the animal's assessment of target location. If the animals are performing the task well, the impact of using saccade metrics instead of target location should be relatively small, and, indeed, the population analysis of time course tended to differ mainly in degree rather than in qualitative pattern (e.g., dotted vs. solid lines in Fig. 6).

Here, we provide details of the monkeys' behavioral responses on visual (Fig. 7A) and auditory (Fig. 7B) trials. Figure 7, A–C, shows mean saccade end points and SDs for each target location for each fixation position. Figure 7, A and B, shows the results including only correct trials; Fig. 7C shows the results for auditory trials, including all trials in which the latency was shorter than 400 ms and the direction of the vertical component of the saccade was correct. Although auditory saccades were more variable than visual saccades and there was a greater tendency to undershoot the more eccentric targets, the average end point location was highly correlated with the true location of the target. This was true even when incorrect trials were included (Fig. 7C). Monkeys mostly, but not completely, compensated for differences in initial eye position when making saccades to sounds. The effect of initial eye position on auditory trials was a little larger than what we have previously observed (Metzger et al. 2004). The average separation between the saccade end points for a given auditory target location from different fixation positions was  $\sim 2.55^\circ$ , or  $\sim 21\%$  of the  $12^\circ$  separation between the fixation positions. For visual targets, the average separation was only  $0.81^\circ$ , or  $\sim 6.8\%$  of the fixation position difference.

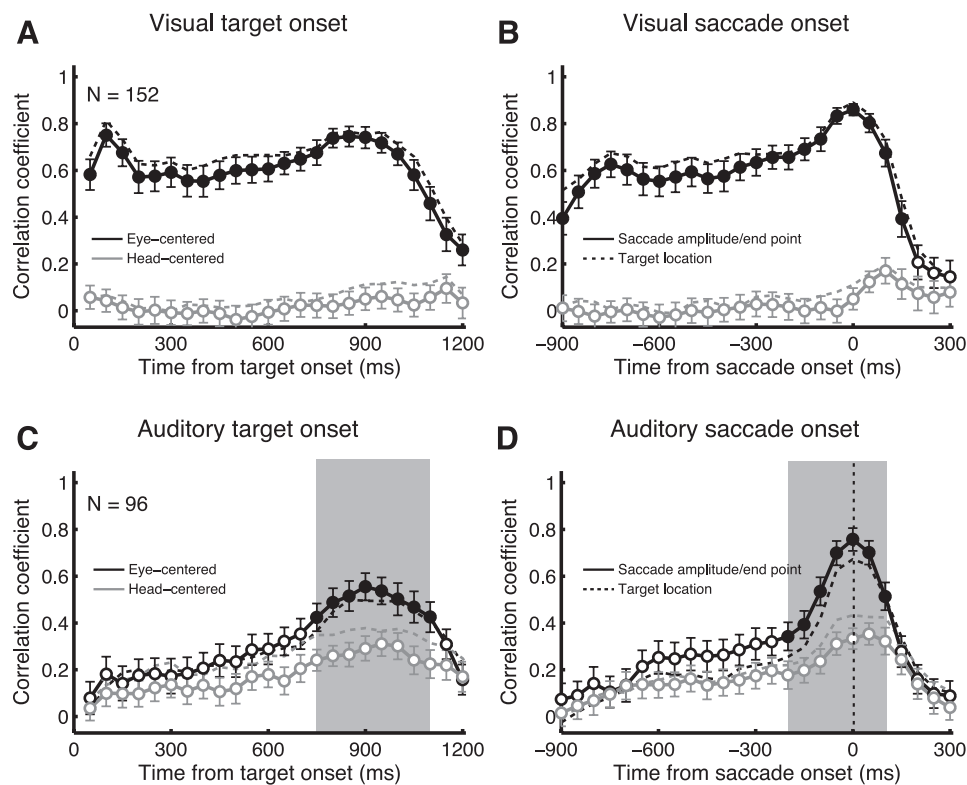


Fig. 6. Time course of the reference-frame transition. For visual trials (A and B), the eye-centered correlation coefficient was significantly larger than the head-centered correlation coefficient for all time periods and had peaks synchronized with both target onset and saccade onset. The solid black circles indicate time bins for which the eye-centered correlation value significantly exceeded the head-centered correlation value ( $P < 0.01$  by paired  $t$ -test). For auditory trials (C and D), the eye-centered and head-centered correlation coefficients were both initially lower than the best correlation coefficient for visual trials, but the eye-centered correlation coefficient reached a peak nearly comparable with that for visual trials around the time of the saccade. The peak was sharper in time for the motor-aligned graph (D), indicating that the coordinate transformation was time locked to the saccade and not the target onset. The gray box and solid black circles highlight the period of time during which the eye-centered correlation value was significantly larger than head-centered value ( $P < 0.01$  by paired  $t$ -test). The error bars indicate  $\pm$ SE. Response windows for this analysis were 100 ms in duration and stepped in 50-ms increments. The dotted lines indicate the time course of correlation coefficients calculated as a function of target location rather than saccade amplitude/end point.

When the head- and eye-centered correlation coefficients were calculated based on auditory target location rather than saccade amplitude/end point (Fig. 7, F and G), the incidence of the hybrid response patterns was about the same during the sensory period (63% vs. 64%, gray portion in the pie plots shown in Figs. 7F vs. 5C) and was slightly higher during the motor period (37% vs. 30%, gray portion in the pie plots shown in Figs. 7G vs. 5D). The incidence of statistically eye-centered response patterns fell slightly from 29% to 21% during the sensory period (orange portion in the pie plots shown in Figs. 5C vs. 7F) and from 69% to 52% during the motor period (orange portion in the pie plots shown in Figs. 5D vs. 7G). Classification of the visual response patterns changed by less than five percentage points in all categories (Figs. 5, A and B vs. 7, D and E). Despite these small changes in the classification of particular response patterns, the overall pattern remained the same: the early hybrid representation of auditory signals progressed to a more eye-centered representation at the time of the movement.

#### Response Strength and Reference Frame

We next considered possible alternative explanations for the auditory response patterns. Activity during the sensory response period was lower for auditory targets than for visual targets. This is evident in the example neuron shown in Fig. 2 as well as the population temporal response profile shown in Fig. 3. Could the hybrid reference frame observed during the sensory period be an artifact of either weaker responses or a lower signal-to-noise ratio on auditory trials during this period?

To investigate this, we first compared the responses during the sensory period for visual versus auditory trials on a neuron-by-neuron basis (Fig. 8, A and B). We used two measures of response strength: 1) the average response during the sensory

interval (expressed in spikes/s; Fig. 8A) and 2) a Z-score of the response relative to the activity during the baseline period [(response – baseline)/SD of the baseline; Fig. 8B]. Although visual responses were on average larger than auditory responses (combined across all target locations, mean response amplitude: visual, 75.0 spikes/s; and auditory: 64.3 spikes/s; mean Z-score: 1.02 and 0.44 for visual and auditory), many neurons had similar visual and auditory response strengths, lying along the line of slope 1. In short, in many cases, differences in response strength were larger across cells within a target modality than within a cell across target modalities.

We next selected a subset of neurons that had well-matched visual and auditory responses. For average responses, we selected neurons whose responses lay within 10 spikes/s of the line of slope 1. For the Z-score, we selected neurons within 0.5 units of the line of slope 1 (black symbols in Fig. 8, A and B;  $N = 46$  for the average response and  $N = 47$  for the Z-score).

We investigated whether there was any correlation between the responses and the observed reference frame within this subset of neurons. Figure 8, C and D, shows the results of correlation between the average responses or Z-scores during the sensory period for both visual and auditory targets and eye- or head-centered coordinates. If stronger responses were a key factor determining the results of the reference frame analysis, the relationship should be significant and positive. But this is not the case. Only one of these analyses (auditory Z-score vs. eye-centered correlation coefficient: 0.37,  $P < 0.01$ ; Fig. 8D, right, orange dots) exhibited a significant positive correlation between the correlation coefficient and response strength.

Finally, again using the subset of well-matched visual and auditory responses, we determined if these neurons were more likely to be classified as eye centered based on the correlation coefficient analysis (Fig. 8, E and F). The results for this subset

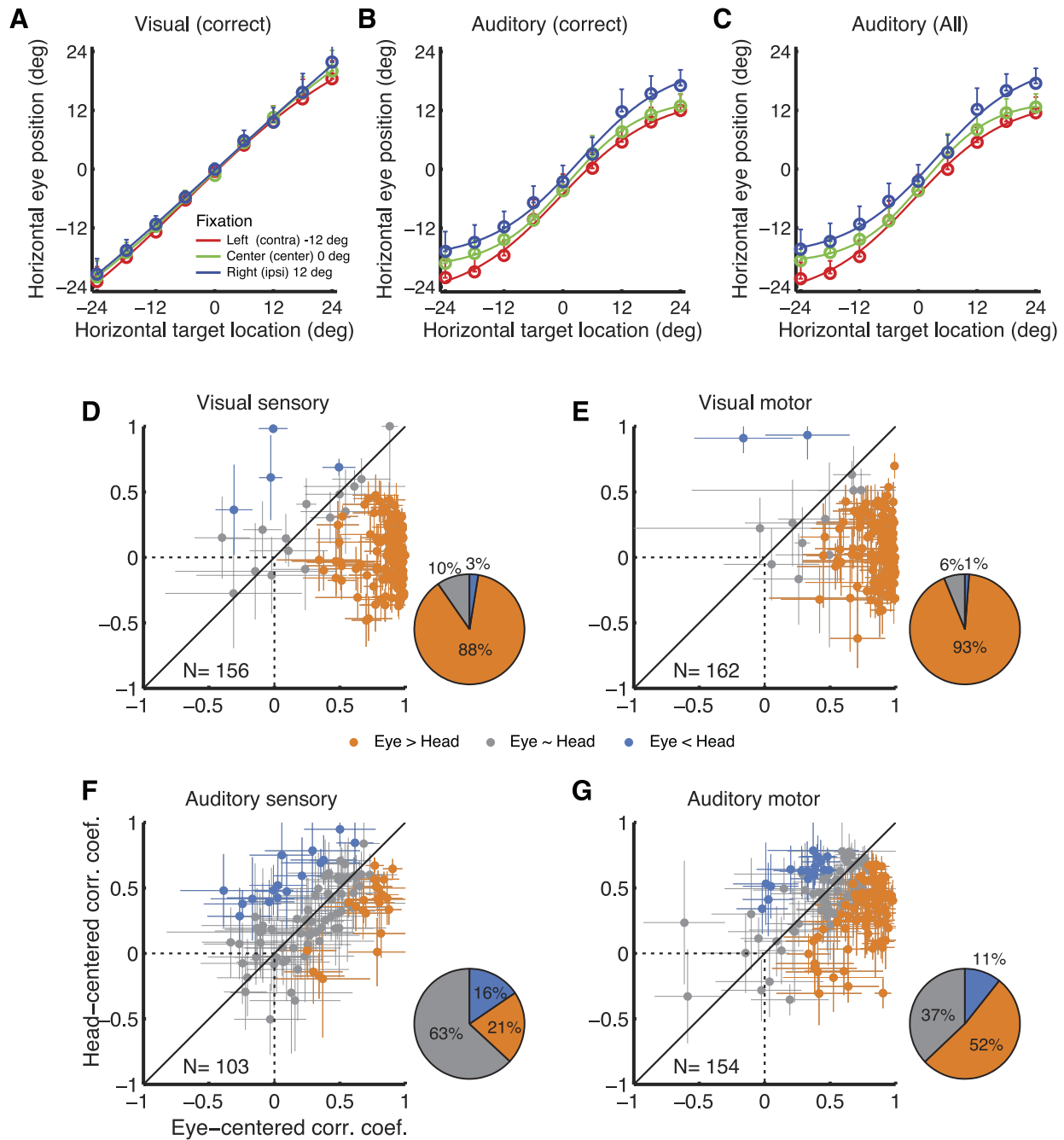


Fig. 7. Saccade accuracy and reference frame. *A–C*: horizontal saccade end point as a function of target locations for visual trials (*A*), auditory correct trials (*B*), or auditory trials including incorrect trials in which a saccade was initiated within 400 ms and had a vertical component in the correct direction (*C*). Colored lines indicate different eye fixations. A slight compression of auditory space relative to visual space was seen for both the correct trials (*B*) and all trials (*C*). The horizontal saccade end point also depended on the initial fixation position (mean shift: 2.55 and 0.81° for auditory and visual, respectively). The error bars indicate 0.5 SEs. *D–G*: correlation coefficient plots based on target location rather than saccade amplitude/end point. All conventions were similar to those shown in Fig. 5.

were very similar to the overall population, as shown in Fig. 5, *A* and *C*. In short, even the subset of neurons that had comparable visual and auditory responses showed different reference frames during the sensory period, with visual signals still predominantly eye centered and auditory signals still predominantly hybrid.

#### Cross-Correlation Analysis

As an alternative means of assessing reference frame, we also calculated how much the receptive fields shifted with eye

position (DI; see MATERIALS AND METHODS) (d'Avossa et al. 2007). To facilitate comparison with other studies (see *Comparison of the Reference Frame With Other Brain Areas* and Fig. 10), we used both target location as well as saccade amplitude/end point for this analysis. For visual sensory and motor activity, the values of DI were clustered around 1 (mean: 0.97 for both sensory and motor activity), suggestive of an eye-centered reference frame (Fig. 9*A*), and were indistinguishable regardless of whether target location or saccade ampli-

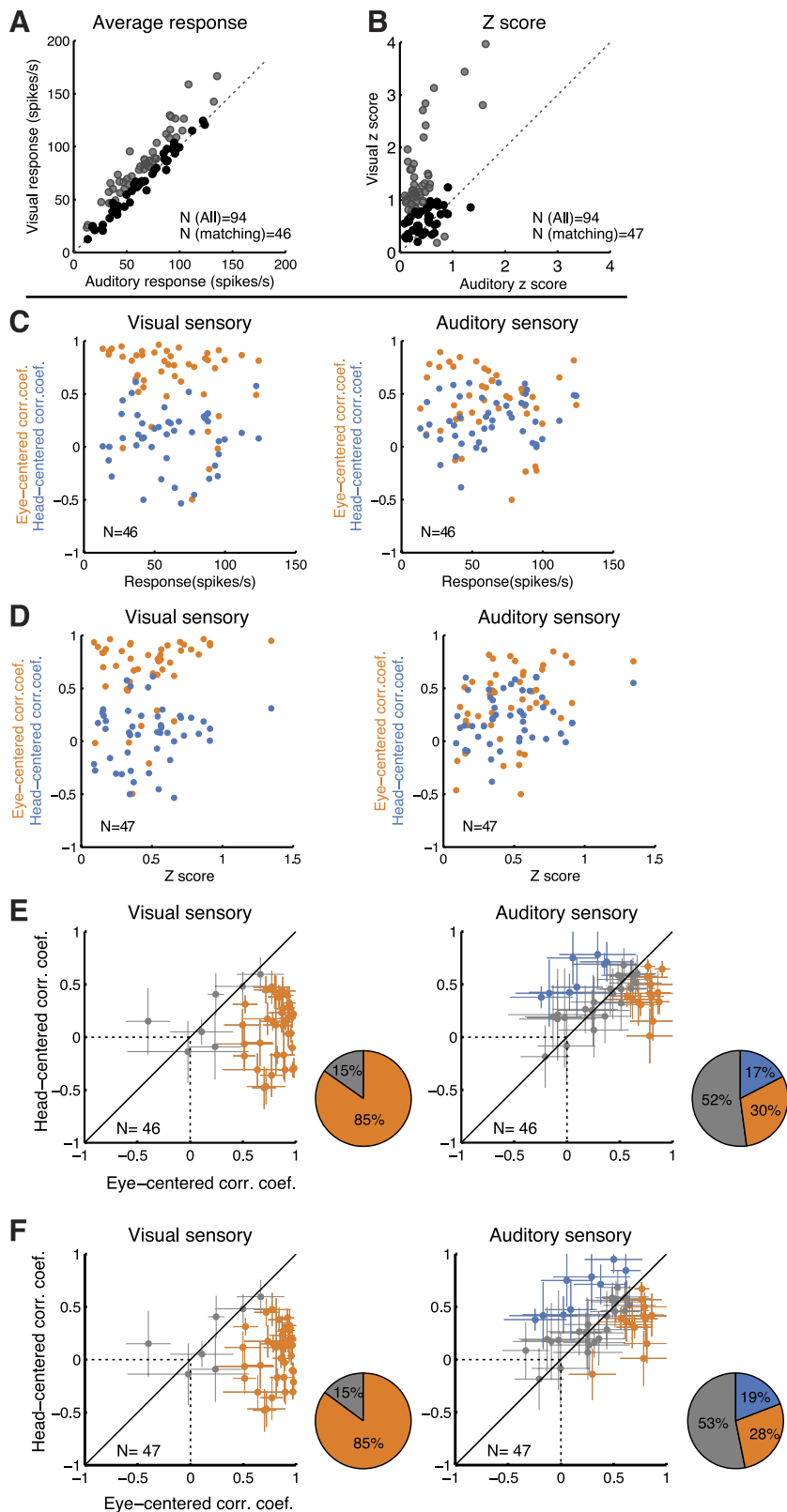


Fig. 8. Relationship between response strength and reference frame. *A*: comparison of average responses on visual versus auditory trials during the sensory period (only neurons with significant responses to both modalities are shown, i.e., neurons that were bimodal during the sensory period,  $N = 94$ ; see Table 1). Solid circles indicate the subset of these neurons in which the difference in responsiveness for the two modalities was  $< 10$  spikes/s ( $N = 46$ ). *B*: comparison of the Z-score of responsiveness [(response - baseline)/(SD of the baseline)] on visual versus auditory trials during the sensory period for bimodal neurons. Solid circles indicate the subset of neurons in which the difference in Z-score on visual versus auditory trials was  $< 0.5$  ( $N = 47$ ). *C*: eye-centered (orange) or head-centered (blue) correlation coefficients as a function of visual (*left*) or auditory (*right*) average responses for the subset of matching neurons in *A* ( $N = 46$ ). Relationships were not statistically significant (nonparametric Spearman correlation: visual eye centered,  $-0.31$ ; visual head centered,  $-0.02$ ; auditory eye centered,  $-0.16$ ; and auditory head centered,  $0.17$ ; all  $P > 0.03$  with Bonferroni correction). *D*: eye-centered (orange) or head-centered (blue) correlation coefficients as a function of visual (*left*) or auditory (*right*) Z-scores for the subset of matching neurons in *B* ( $N = 47$ ). Spearman correlations were as follows: visual eye centered,  $0.25$  (not significant); visual head centered,  $0.05$  (not significant); auditory eye centered,  $0.37$  ( $P < 0.01$ ); and auditory head centered,  $0.2$  (not significant). *E* and *F*: correlation coefficient plots involving the subset of neurons with similar visual and auditory responses in *A* (black circles,  $N = 46$ ) and *B* (black circles,  $N = 47$ ), respectively. All conventions were similar to those shown in Fig. 5.

tude/end point was used (open symbols compared with filled symbols). The DI for auditory sensory activity was intermediate between eye- and head-centered coordinates (target location mean:  $0.56$ , saccade amplitude/end point mean:  $0.56$ ; Fig. 9*B*). The DI for auditory motor activity was more eye centered

(target location mean:  $0.68$ , filled black triangles; saccade amplitude/end point mean:  $0.84$ , open black triangles). Thus, the mean DI for auditory motor activity was only slightly lower than the DI value for visual motor activity when using saccade amplitude/end point as the measure of localization.

The time course of DI relative to the saccade onset is shown in Fig. 9, *C* (visual) and *D* (auditory). The overall pattern was quite similar to the time course of the correlation coefficient (Fig. 6, *C* and *D*). For visual targets (Fig. 9*C*), the average DI was significantly above 0.5 (the midpoint between head- and eye-centered coordinates) throughout the time interval, averaging  $\sim 0.9$  and peaking at  $\sim 0.99$  at the time of the saccade. This was true regardless of whether target location or saccade amplitude/end point was used (dotted vs. solid lines). When the target was auditory (Fig. 9*D*), the average DI for saccade amplitude/end point was initially intermediate between head- and eye-centered coordinates (mean: 0.68). Starting at  $\sim 350$  ms before the saccade, the DI significantly increased above the index midpoint, peaking at  $\sim 0.86$  at the saccade onset. This peak in the auditory DI was better time locked to the saccade (peak value: 0.86) than to the target (peak value: 0.80; data not shown), again suggesting that the transformation of auditory

signals to an eye-centered reference frame occurred specifically in conjunction with generating the saccade. For the auditory DI using target location (dotted lines), the increase still occurred at the time of the saccade (peak value: 0.72), but it was less marked and reached a lower peak.

#### Comparison of the Reference Frame With Other Brain Areas

We have previously conducted very similar assessments of reference frame in the lateral (LIP) and medial banks of the intraparietal cortex (MIP) (Mullette-Gillman et al. 2005, 2009), primary auditory cortex (AC) (Werner-Reiss et al. 2003), and the inferior colliculus (Groh et al. 2001). As shown in Fig. 10, we compared the DI across these brain areas. The horizontal locations of the targets and fixation positions included in this analysis were approximately the same for all four brain regions (target range:  $-24$  to  $24^\circ$  and three  $12^\circ$  separated fixations).

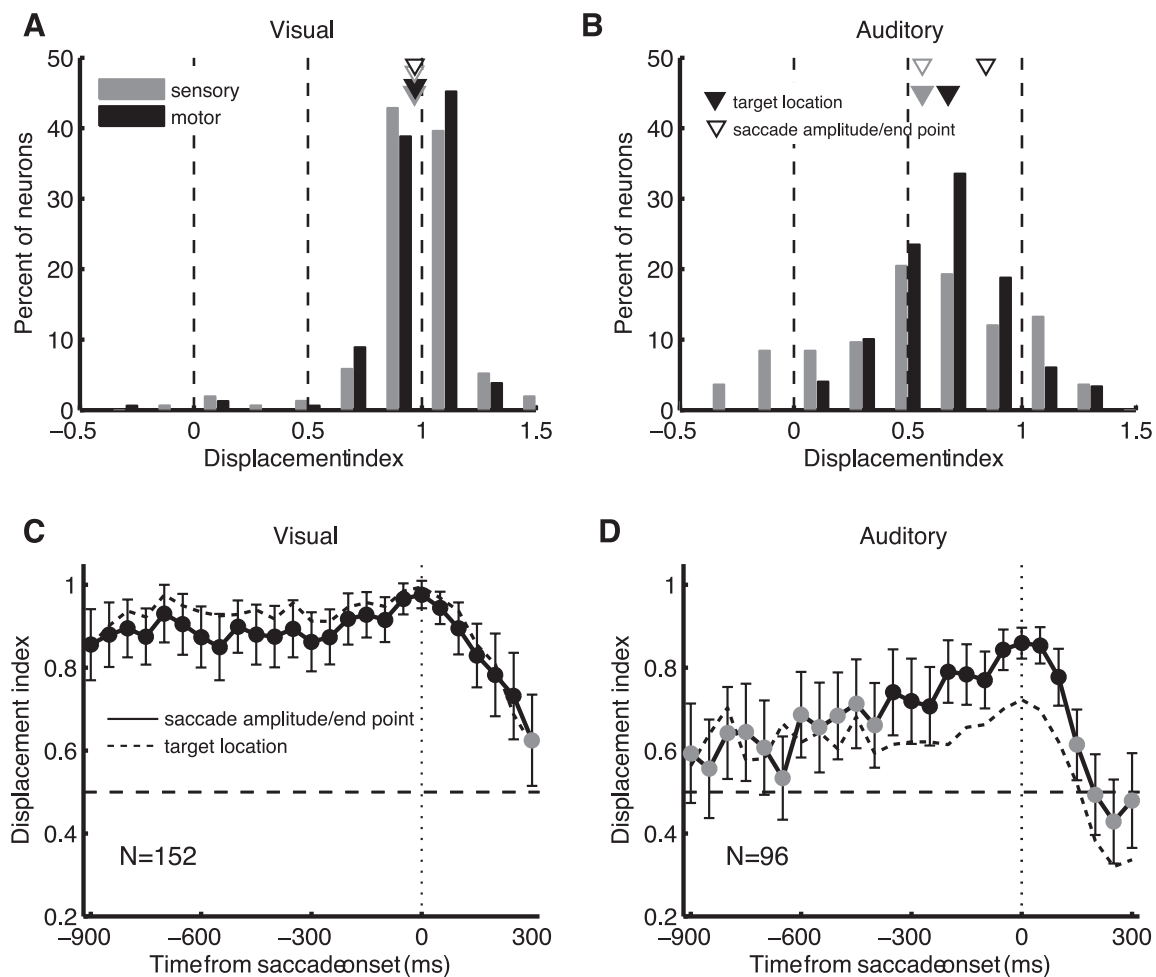


Fig. 9. Distribution and time course of the displacement index (DI). The percentage of neurons was plotted as a function of DI value for sensory (gray) and motor (black) periods for visual (*A*) and auditory (*B*) targets. The dashed vertical lines mark DI values of 0, 0.5, and 1, corresponding to a perfect head-centered reference frame, response functions that shift half of the amount of the eye movement, and perfect eye-centered reference frame, respectively. The full distributions of values are shown based on target location (solid symbols), and the means are shown for saccade amplitude/end point (open symbols). Mean DIs for each distribution based on target location were as follows: visual sensory (mean  $\pm$  SD:  $0.97 \pm 0.21$ ), visual motor ( $0.97 \pm 0.20$ ), auditory sensory ( $0.56 \pm 0.42$ ), and auditory motor ( $0.68 \pm 0.26$ ). Mean DIs using saccade amplitude/end point were as follows: visual sensory (mean  $\pm$  SD:  $0.97 \pm 0.21$ ), visual motor ( $0.97 \pm 0.21$ ), auditory sensory ( $0.56 \pm 0.42$ ), and auditory motor ( $0.84 \pm 0.24$ ). *C* and *D*: time course of DI based on saccade amplitude/end point for visual (*C*) and auditory (*D*) trials aligned on the saccade onset. The DI was calculated for 100-ms windows stepping in increments of 50 ms during the time period from 900 ms before the saccade onset to 300 ms after the saccade onset and was averaged across cells showing significant responses for both sensory and motor periods. The black circles indicate mean values statistically significant larger than 0.5 (horizontal dashed lines,  $P < 0.01$  by one-tailed *t*-test). The vertical dotted line indicates the saccade onset. The error bar represents  $\pm$ SE. The dotted lines indicate the time course of DI calculated as a function of target location for comparison. At the saccade onset, the peak auditory DI based on saccade amplitude/end point was larger (0.86) than that based on the target location (0.72).

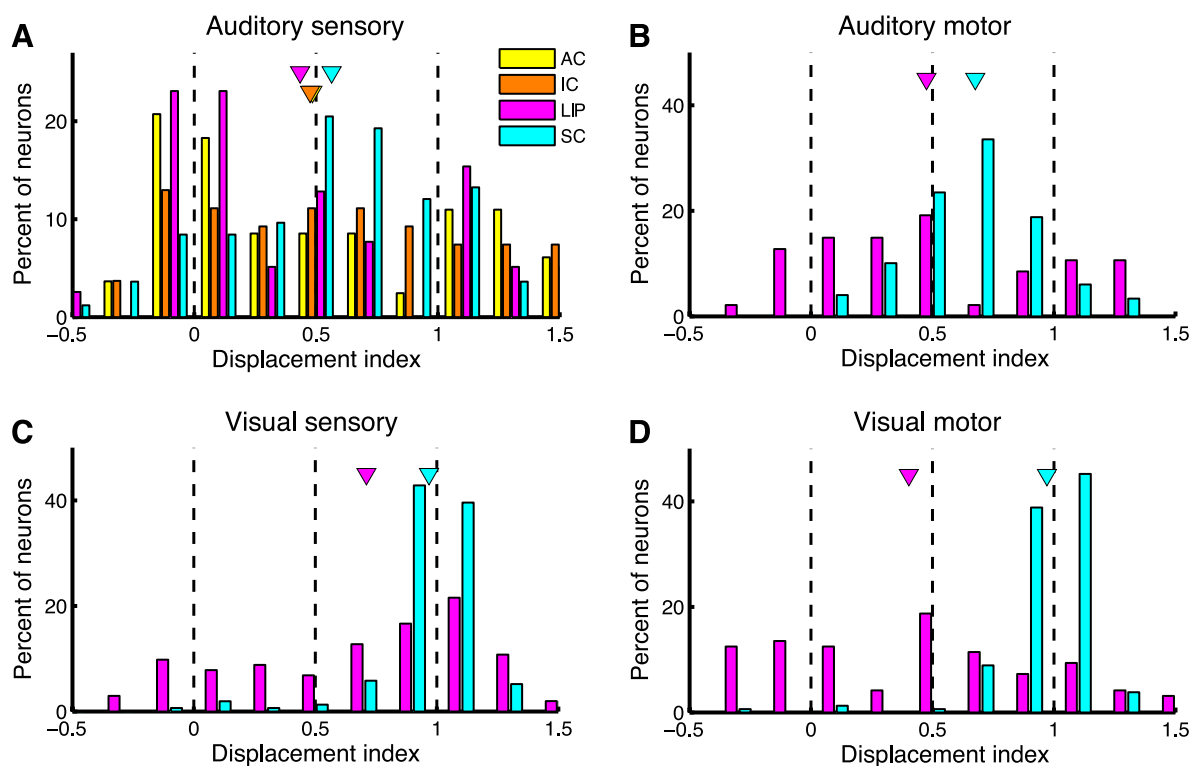


Fig. 10. Comparison of DI values in the SC to other brain areas. *A*: auditory sensory activity was similar in all brain areas tested [the inferior colliculus (IC; orange), auditory cortex (AC; yellow), lateral bank of the intraparietal cortex (LIP; pink), and SC (cyan)]. Note that all DIs were calculated based on the actual target location and were calculated from data limited as appropriate to match the approximate range of SC data in this study (i.e., from  $-24$  to  $24^\circ$  and 3 fixations with  $12^\circ$  separation). The three vertical lines correspond to DI values of 0 (head-centered reference frame), 0.5 (hybrid – 50% response function shift), and 1 (eye-centered reference frame). The mean value for each distribution is marked with an upside-down triangle. Mean auditory sensory DIs in these four brain areas were close to 0.5 [means  $\pm$  SD; IC:  $0.48 \pm 0.58$ , AC:  $0.48 \pm 0.56$ , LIP/medial bank of the intraparietal cortex (MIP):  $0.43 \pm 0.53$ , and SC:  $0.56 \pm 0.42$ ]. (Note that despite the similarity in the mean values, there were differences in the shapes of the distributions, with the other brain areas including more neurons with head-centered response functions in addition to a range of hybrid coding, whereas the SC's hybrid representation was more unimodal, with a peak in the hybrid range.) *B*: auditory motor activity in the SC and LIP/MIP. Auditory motor activity differed in the SC compared with the parietal cortex. The mean DI value for auditory motor activity in the LIP/MIP was quite hybrid (mean  $\pm$  SD:  $0.48 \pm 0.49$ ), whereas the value for the SC was more eye centered ( $0.68 \pm 0.26$ ,  $P < 0.0001$  by two-sample *t*-test). *C* and *D*: similar differences were seen for visual sensory activity (*C*) and visual motor activity (*D*) between the SC (cyan) and LIP/MIP (pink). For visual sensory responses, the mean DI for the SC was  $0.97 \pm 0.21$ , whereas for LIP/MIP it was  $0.71 \pm 0.47$ . For visual motor activity, the mean DI for the SC was  $0.97 \pm 0.20$ , whereas for LIP/MIP it was  $0.40 \pm 0.53$ . These both differed significantly ( $P < 0.0001$  by two-sample *t*-test).

Neurons were screened for inclusion in this analysis using the same methods as SC data [see *Responsiveness and Spatial Tuning* in MATERIALS AND METHODS: 1) testing the significance of the response from baseline activity using a one-tailed *t*-test and 2) testing the spatial selectivity using one-way ANOVA]. For the intraparietal cortex and AC experiments, monkeys were not required to make saccades to the sounds, whereas they did for the study involving the parietal cortex. Target location (rather than saccade metrics) was therefore used for calculating the DI index for all four areas.

The results shown in Fig. 10 demonstrate that the intraparietal cortex, AC, and LIP/MIP use a hybrid reference frame for both auditory and, when present, visual sensory and motor signals: the DI values for these brain areas, target modalities, and response periods lay in the intermediate range. For auditory sensory signals, the DI values in these areas were not significantly different from that of the SC ( $P = 0.37$  by one-way ANOVA; Fig. 10A). In contrast, the visual sensory, visual motor, and auditory motor signals in the SC differed significantly from their counterparts in the LIP/MIP ( $P < 0.0001$  for all comparisons by two-sample *t*-tests; Fig. 10, B–D). These results reveal that a significant transformation of

signals occurs between the LIP/MIP and SC, which is situated later than the LIP/MIP in the audiovisual-oculomotor pathway.

## DISCUSSION

Our study addresses a long-standing question in the literature: what happens to the discrepant reference frames initially observed in response to visual and auditory stimuli? That monkeys largely compensate for eye position when making saccades to sounds (Fig. 7) (Metzger et al. 2004) strongly suggested that somewhere in the brain, visual and auditory signals must join together in a common reference frame, but the brain area and particular reference frame used were not known.

Our study shows that the difference in reference frame observed in the SC shortly after targets are presented is much reduced by the time saccades to those targets are programmed. Saccade-related bursts encode auditory saccade metrics in a reference frame that is predominantly eye centered, but still not quite as eye centered as the pattern of activity observed when the target is visual. In short, a change in reference frame is evident as the auditory response patterns of individual neurons unfold in time, and this change reduces the discrepancy between the activities evoked by targets of different modalities.

This outcome did not have to be the case. Although the conventional wisdom holds that the SC's representation is eye centered, the evidence for this was neither complete nor free of nuance. Quantitative recording studies of the reference frame of visual or oculomotor signals in the SC have been limited (Jay and Sparks 1987b). Other recent recording studies (Van Opstal et al. 1995; Pare and Munoz 2001; Campos et al. 2006) revealed that eye position modulates the activity of SC neurons. It should not be assumed that such eye position modulations are irrelevant to the reference frame used by a neural representation; rather, the presence of such modulations suggests that the reference frame should be quantitatively verified. Using analyses largely robust to modulatory eye position influences, our study thus provides confirmation that visual/oculomotor receptive/movement fields of SC neurons are stable in location across changes in eye position and is consistent with recent findings by DeSouze et al. (2011) for head-unrestrained monkeys. Taken together, the emerging portrait is that the SC's representation is predominantly, but not purely, eye centered: SC responses are a function of both the eye-centered location of the target (a dominant influence) and the position of the eyes (a weaker contributor). Strong electrical stimulation of the SC appears to override such eye position modulation in evoking eye-centered saccades (Pare et al. 1994; Freedman et al. 1996; Freedman and Sparks 1997; Klier et al. 2001), but weak electrical stimulation reveals that such modulations can indeed contribute to evoked saccades (Groh 2011).

Our findings also differ markedly from our previous study in the parietal cortex using the same behavioral paradigm with similar target locations. In that study, neither auditory nor visual signals were predominantly eye centered, and there was no evolution of reference frame between the sensory and motor periods (Mullette-Gillman et al. 2005, 2009). The discrepancy between the SC and parietal cortex implies that the functions of these two brain areas may be quite different.

The presence of eye-centered auditory activity only in the motor-related discharges of SC neurons suggests that the reference frame of sensory signals may be underconstrained, and an eye-centered coordinate system only emerges when a specific behavior requiring the computation of target location in that reference frame must be generated. Although the classical view of the SC emphasizes its role in the control and execution of eye movements, several recent studies have revealed that the SC also plays a role in controlling attention to locations in space (Goldberg and Wurtz 1972; Kustov and Robinson 1996; Carello and Krauzlis 2004; Muller et al. 2005). Thus, an eye-centered representation of sounds at the level of the SC has the potential to facilitate a range of multimodal interactions beyond the immediate generation of an eye movement to the source of a sound.

Whether the emergence of an eye-centered reference frame in the saccade-related bursts of SC neurons reflects computations that occur within the SC or are received by the SC from elsewhere is uncertain. The main sources of auditory input to the SC are the inferior colliculus (IC), AC, frontal eye fields (FEF), LIP, and MIP (Kunzle et al. 1976; Moore et al. 1977; FitzPatrick and Imig 1978; Fries 1984; Sparks and Hartwich-Young 1989; John and Van Lieshout 2000; Hanes and Wurtz 2001). The available evidence for the IC, AC, and LIP/MIP

suggests that auditory signals are encoded in a hybrid format in these areas (Fig. 10) (Groh et al. 2001; Werner-Reiss et al. 2003; Mullette-Gillman et al. 2005, 2009) (note that the reference frame at the time of a saccade has not been investigated in the IC or AC). In the FEF, activity is known to depend on the sound location with respect to the eyes, but it is not known if the movement fields are predominantly or only partially eye centered (Russo and Bruce 1994). Information on whether auditory-related signals in an eye-centered frame of reference exist anywhere else in the brain would help establish whether the computation is likely to occur within the SC itself or might possibly occur elsewhere.

That the reference frame may be dictated by the motor effector and that such coordinate systems may emerge most clearly later in motor pathways has been observed for head and reaching movements as well: visual and auditory signals in quite clear head- and limb-centered coordinates have been observed in the ventral premotor cortex (Graziano and Gross 1998; Graziano et al. 1999; Graziano 2001). Crawford and colleagues have proposed a "conversion-on-demand" model in which targets are retained in sensory coordinates and only those relevant to a specific action are made available to downstream motor systems (Henriques et al. 1998; Klier et al. 2001). Thus, "better late than never" may be a general principle for the brain's strategy of resolving signals into an eye-centered reference frame only at the point where and when a command begins to be prepared to direct the eyes, and thus focus visual processing, on a particular location in space.

#### ACKNOWLEDGMENTS

The authors thank Jessi Cruger for animal care and Tom Heil for technical assistance. The authors also thank Valeria Caruso, Kurtis Gruters, and Daniel Pages for helpful comments on the manuscript.

#### GRANTS

This work was supported by National Institute of Neurological Disorders and Stroke Grant NS-50942. J. Lee was also supported by Korea Research Foundation Grant KRF-2008-356-H00003 funded by the Korean Government.

#### DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

#### AUTHOR CONTRIBUTIONS

Author contributions: J.L. and J.M.G. conception and design of research; J.L. performed experiments; J.L. and J.M.G. analyzed data; J.L. and J.M.G. interpreted results of experiments; J.L. prepared figures; J.L. and J.M.G. drafted manuscript; J.L. and J.M.G. edited and revised manuscript; J.L. and J.M.G. approved final version of manuscript.

#### REFERENCES

- Andersen RA, Snyder LH, Li CS, Stricanne B. Coordinate transformations in the representation of spatial information. *Curr Opin Neurobiol* 3: 171–176, 1993.
- Avillac M, Deneve S, Olivier E, Pouget A, Duhamel JR. Reference frames for representing visual and tactile locations in parietal cortex. *Nat Neurosci* 8: 941–949, 2005.
- Bahill AT, Bahill KA, Clark MR, Stark L. Closely spaced saccades. *Invest Ophthalmol* 14: 317–321, 1975.
- Campos M, Cherian A, Segraves MA. Effects of eye position upon activity of neurons in macaque superior colliculus. *J Neurophysiol* 95: 505–526, 2006.

- Carello CD, Krauzlis RJ. Manipulating intent: evidence for a causal role of the superior colliculus in target selection. *Neuron* 43: 575–583, 2004.
- Colby CL, Goldberg ME. Space and attention in parietal cortex. *Annu Rev Neurosci* 22: 319–349, 1999.
- d'Avossa G, Tosetti M, Crespi S, Biagi L, Burr DC, Morrone MC. Spatiotopic selectivity of BOLD responses to visual motion in human area MT. *Nat Neurosci* 10: 249–255, 2007.
- DeSouza JF, Keith GP, Yan X, Blohm G, Wang H, Crawford JD. Intrinsic reference frames of superior colliculus visuomotor receptive fields during head-unrestrained gaze shifts. *J Neurosci* 31: 18313–18326, 2011.
- FitzPatrick KA, Imig TJ. Projections of auditory cortex upon the thalamus and midbrain in the owl monkey. *J Comp Neurol* 177: 573–555, 1978.
- Freedman EG, Sparks DL. Activity of cells in the deeper layers of the superior colliculus of the rhesus monkey: evidence for a gaze displacement command. *J Neurophysiol* 78: 1669–1690, 1997.
- Freedman EG, Stanford TR, Sparks DL. Combined eye-head gaze shifts produced by electrical stimulation of the superior colliculus in rhesus monkeys. *J Neurophysiol* 76: 927–952, 1996.
- Fries W. Cortical projections to the superior colliculus in the macaque monkey: a retrograde study using horseradish peroxidase. *J Comp Neurol* 230: 55–76, 1984.
- Gardner JL, Merriam EP, Movshon JA, Heeger DJ. Maps of visual space in human occipital cortex are retinotopic, not spatiotopic. *J Neurosci* 28: 3988–3999, 2008.
- Glimcher PW, Sparks DL. Movement selection in advance of action in the superior colliculus. *Nature* 355:542–545, 1992.
- Goldberg ME, Wurtz RH. Activity of superior colliculus in behaving monkey. II. Effect of attention on neuronal responses. *J Neurophysiol* 35:560–574, 1972.
- Graziano MS. A system of multimodal areas in the primate brain. *Neuron* 29: 4–6, 2001.
- Graziano MS, Gross CG. Spatial maps for the control of movement. *Curr Opin Neurobiol* 8: 195–201, 1998.
- Graziano MS, Reiss LA, Gross CG. A neuronal representation of the location of nearby sounds. *Nature* 397: 428–430, 1999.
- Groh JM. Effects of initial eye position on saccades evoked by microstimulation in the primate superior colliculus: implications for models of the SC read-out process. *Front Integr Neurosci* 4: 130, 2011.
- Groh JM, Sparks DL. Saccades to somatosensory targets. I. behavioral characteristics. *J Neurophysiol* 75: 412–427, 1996.
- Groh JM, Trause AS, Underhill AM, Clark KR, Inati S. Eye position influences auditory responses in primate inferior colliculus. *Neuron* 29: 509–518, 2001.
- Guitton D, Crommelinck M, Roucoux A. Stimulation of the superior colliculus in the alert cat. I. Eye movements and neck EMG activity evoked when the head is restrained. *Exp Brain Res* 39: 63–73, 1980.
- Hanes DP, Wurtz RH. Interaction of the frontal eye field and superior colliculus for saccade generation. *J Neurophysiol* 85: 804–815, 2001.
- Hartline PH, Vimal RL, King AJ, Kurylo DD, Northmore DP. Effects of eye position on auditory localization and neural representation of space in superior colliculus of cats. *Exp Brain Res* 104: 402–408, 1995.
- Henriques DY, Klier EM, Smith MA, Lowy D, Crawford JD. Gaze-centered remapping of remembered visual space in an open-loop pointing task. *J Neurosci* 18: 1583–1594, 1998.
- Jay MF, Sparks DL. Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature* 309: 345–347, 1984.
- Jay MF, Sparks DL. Sensorimotor integration in the primate superior colliculus. I. Motor convergence. *J Neurophysiol* 57: 22–34, 1987a.
- Jay MF, Sparks DL. Sensorimotor integration in the primate superior colliculus. II. Coordinates of auditory signals. *J Neurophysiol* 57: 35–55, 1987b.
- Jay MF, Sparks DL. Localization of auditory, and visual targets for the initiation of saccadic eye movements. In: *Comparative Perception: Basic Mechanism*, edited by Berkely MA, Stebbins WC. New York: Wiley, 1990, vol. 1, p. 527.
- John KH, Van Lieshout DP. Projections from the rostral pole of the inferior colliculus to the cat superior colliculus. *Brain Res* 881: 244–247, 2000.
- Judge SJ, Richmond BJ, Chu FC. Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res* 20: 535–538, 1980.
- Klier EM, Wang H, Crawford JD. The superior colliculus encodes gaze commands in retinal coordinates. *Nat Neurosci* 4: 627–632, 2001.
- Knudsen EL, Knudsen PF. Vision guides the adjustment of auditory localization in young barn owls. *Science* 230: 545–548, 1985.
- Kopco N, Lin IF, Shinn-Cunningham BG, Groh JM. Reference frame of the ventriloquism aftereffect. *J Neurosci* 29: 13809–13814, 2009.
- Kunzle H, Akert K, Wurtz RH. Projection of area 8 (frontal eye field) to superior colliculus in the monkey. An autoradiographic study. *Brain Res* 117: 487–492, 1976.
- Kustov AA, Robinson DL. Modified saccades evoked by stimulation of the macaque superior colliculus account for properties of the resettable integrator. *J Neurophysiol* 73: 1724–1728, 1995.
- Kustov AA, Robinson DL. Shared neural control of attentional shifts and eye movements. *Nature* 384: 74–77, 1996.
- Lee C, Rohrer WH, Sparks DL. Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature* 332: 357–360, 1988.
- Metzger RR, Mulette-Gillman OA, Underhill AM, Cohen YE, Groh JM. Auditory saccades from different eye positions in the monkey: implications for coordinate transformations. *J Neurophysiol* 92: 2622–2627, 2004.
- Moore JK, Karapas F, Moore RY. Projections of the inferior colliculus in insectivores and primates. *Brain Behav Evol* 14: 301–327, 1977.
- Muller JR, Philiastides MG, Newsome WT. Microstimulation of the superior colliculus focuses attention without moving the eyes. *Proc Natl Acad Sci USA* 102: 524–529, 2005.
- Mulette-Gillman OA, Cohen YE, Groh JM. Eye-centered, head-centered, and complex coding of visual and auditory targets in the intraparietal sulcus. *J Neurophysiol* 94: 2331–2352, 2005.
- Mulette-Gillman OA, Cohen YE, Groh JM. Motor-related signals in the intraparietal cortex encode locations in a hybrid, rather than eye-centered reference frame. *Cereb Cortex* 19: 1761–1775, 2009.
- Munoz DP, Wurtz RH. Saccade-related activity in monkey superior colliculus. I. Characteristics of burst and buildup cells. *J Neurophysiol* 73: 2313–2333, 1995.
- Pare M, Crommelinck M, Guitton D. Gaze shifts evoked by stimulation of the superior colliculus in the head-free cat conform to the motor map but also depend on stimulus strength and fixation activity. *Exp Brain Res* 101: 123–139, 1994.
- Pare M, Munoz DP. Expression of a re-centering bias in saccade regulation by superior colliculus neurons. *Exp Brain Res* 137: 354–368, 2001.
- Peck CK, Baro JA, Warder SM. Effects of eye position on saccadic eye movements and on the neuronal responses to auditory and visual stimuli in cat superior colliculus. *Exp Brain Res* 103: 227–242, 1995.
- Populin LC, Tollin DJ, Yin TC. Effect of eye position on saccades and neuronal responses to acoustic stimuli in the superior colliculus of the behaving cat. *J Neurophysiol* 92: 2151–2167, 2004.
- Porter KK, Metzger RR, Groh JM. Representation of eye position in primate inferior colliculus. *J Neurophysiol* 95: 1826–1842, 2006.
- Recanzone GH. Rapidly induced auditory plasticity: the ventriloquism aftereffect. *Proc Natl Acad Sci USA* 95: 869–875, 1998.
- Robinson DA. A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Trans Biomed Eng* 10: 137–145, 1963.
- Robinson DA. Eye movements evoked by collicular stimulation in the alert monkey. *Vision Res* 12: 1795–1808, 1972.
- Russo GS, Bruce CJ. Frontal eye field activity preceding aurally guided saccades. *J Neurophysiol* 71: 1250–1253, 1994.
- Schiller PH, Stryker M. Single-unit recording and stimulation in superior colliculus of the alert rhesus monkey. *J Neurophysiol* 35: 915–924, 1972.
- Schlack A, Sterbing-D'Angelo SJ, Hartung K, Hoffmann KP, Bremmer F. Multisensory space representations in the macaque ventral intraparietal area. *J Neurosci* 25: 4616–4625, 2005.
- Schlag-Rey M, Schlag J, Shook B. Interactions between natural and electrically evoked saccades. I. Differences between sites carrying retinal error and motor error signals in monkey superior colliculus. *Exp Brain Res* 76: 537–547, 1989.
- Sparks DL, Hartwich-Young R. The deep layers of the superior colliculus. *Rev Oculomot Res* 3: 213–255, 1989.
- Sparks DL, Holland R, Guthrie BL. Size and distribution of movement fields in the monkey superior colliculus. *Brain Res* 113: 21–34, 1976.
- Sparks DL, Mays LE. Spatial localization of saccade targets. I. Compensation for stimulation-induced perturbations in eye position. *J Neurophysiol* 49: 45–63, 1983.
- Stricanne B, Andersen RA, Mazzoni P. Eye-centered, head-centered, and intermediate coding of remembered sound locations in area LIP. *J Neurophysiol* 76: 2071–2076, 1996.
- Van Opstal AJ, Hepp K, Suzuki Y, Henn V. Influence of eye position on activity in monkey superior colliculus. *J Neurophysiol* 74: 1593–1610, 1995.



- Waitzman DM, Ma TP, Optican LM, Wurtz RH.** Superior colliculus neurons provide the saccadic motor error signal. *Exp Brain Res* 72: 649–652, 1988.
- Werner-Reiss U, Kelly KA, Trause AS, Underhill AM, Groh JM.** Eye position affects activity in primary auditory cortex of primates. *Curr Biol* 13: 554–562, 2003.
- Witten IB, Knudsen EI.** Why seeing is believing: merging auditory and visual worlds. *Neuron* 48: 489–496, 2005.
- Woods TM, Recanzone GH.** Visually induced plasticity of auditory spatial perception in macaques. *Curr Biol* 14: 1559–1564, 2004.
- Wurtz RH, Albano JE.** Visual-motor function of the primate superior colliculus. *Annu Rev Neurosci* 3: 189–226, 1980.
- Wurtz RH, Goldberg ME.** Superior colliculus cell responses related to eye movements in awake monkeys. *Science* 171: 82–84, 1971.
- Yao L, Peck CK.** Saccadic eye movements to visual and auditory targets. *Exp Brain Res* 115: 25–34, 1997.
- Zella JC, Brugge JF, Schnupp JW.** Passive eye displacement alters auditory spatial receptive fields of cat superior colliculus neurons. *Nat Neurosci* 4: 1167–1169, 2001.

