NIH Public Access
Author Manuscript
Nat Neurosci. Author manuscript; available in PMC 2009 October 1.

Published in final edited form as:
Nat Neurosci. 2008 October ; 11(10): 1201–1210. doi:10.1038/nn.2191.

Neural correlates of multi-sensory cue integration in macaque area MSTd

Yong Gu[1], **Dora E. Angelaki**[1,*], and **Gregory C. DeAngelis**[1,2,*]

[1]Department. of Anatomy & Neurobiology, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, MO 63110 USA

[2]Department of Brain and Cognitive Sciences, Center for Visual Science, University of Rochester, 245 Meliora Hall, Rochester NY 14627 USA

Abstract

Human observers combine multiple sensory cues synergistically to achieve greater perceptual sensitivity, but little is known about the underlying neuronal mechanisms. We recorded from neurons in the dorsal medial superior temporal area (MSTd) during a task in which trained monkeys combine visual and vestibular cues near optimally to discriminate heading. During bimodal stimulation, MSTd neurons combine visual and vestibular inputs linearly with sub-additive weights. Neurons with congruent heading preferences for visual and vestibular stimuli show improvements in sensitivity that parallel behavioral effects. In contrast, neurons with opposite preferences show diminished sensitivity under cue combination. Responses of congruent cells are more strongly correlated with monkeys' perceptual decisions than opposite cells, suggesting that the animal monitors the activity of congruent cells to a greater extent during cue integration. These findings demonstrate perceptual cue integration in non-human primates and identify a population of neurons that may form its neural basis.

Understanding how the brain combines different sources of sensory information to optimize perception is a fundamental problem in neuroscience. Information from different sensory modalities is often seamlessly integrated into a unified percept. Combining sensory inputs leads to improved behavioral performance in many contexts, including integration of texture and motion cues for depth perception[1], stereo and texture cues for slant perception[2,3], visual-haptic integration[4,5], visual-auditory localization[6], and object recognition[7]. Multisensory integration in human behavior often follows predictions of a quantitative framework that applies Bayesian statistical inference to the problem of cue integration[8–10]. An important prediction is that subjects show greater perceptual sensitivity when two cues are presented together than when either cue is presented alone. This improvement in sensitivity is largest (a factor of $\sqrt{2}$) when the two cues have equal reliability[5,10].

Despite intense recent interest in cue integration, the underlying neural mechanisms remain unclear. Improved perceptual performance during cue integration is thought to be mediated by neurons selective for multiple sensory stimuli[11]. Multi-modal neurons have been described in several brain areas[12,13], but these studies have typically been performed in anesthetized or passively viewing animals[14–17]. Multi-modal neurons have not been studied during

Address for correspondence: Gregory C. DeAngelis, Professor of Brain and Cognitive Sciences, Center for Visual Science, University of Rochester, 245 Meliora Hall, Rochester, NY 14627, Office Phone: (585)275-8677, FAX: (585)271-3043, Email: E-mail: gdeangelis@cvs.rochester.edu.

*equal author contributions

Author contributions: G.Y, D.E.A. and G.C.D. designed the experiments; G.Y. collected the data and performed data analyses; G.Y, D.E.A. and G.C.D. refined the analyses and wrote the paper.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

performance of multi-sensory tasks comparable to those used in human psychophysics. Because cue integration may only occur when cues have roughly matched perceptual reliabilities[5,6,18], it is critical to address the neural mechanisms of sensory integration under conditions in which cue combination is known to take place perceptually.

We trained macaque monkeys to report their direction of self-motion (heading) using both optic flow (visual) and inertial motion (vestibular) cues. A plausible neural substrate for this sensory integration is area MSTd, which contains neurons selective for optic flow[19–22] as well as inertial motion in darkness[23–27]. We show that monkeys combine visual and vestibular heading cues to improve perceptual sensitivity. By recording from single MSTd neurons, we address three important questions. First, do single MSTd cells show improved neuronal sensitivity under cue combination that parallels the change in behavior? Second, can bimodal responses be modeled as weighted linear sums of responses to the individual cues, as predicted by recent theory[28]? Third, do MSTd responses correlate with monkeys' perceptual decisions under cue combination and do such correlations depend on the congruency of tuning for visual and vestibular cues? Our findings establish a candidate neural substrate for visual/vestibular cue integration in macaque visual cortex.

Results

Psychophysical performance

We trained two monkeys to perform a multi-modal heading discrimination task in the horizontal plane (Fig 1a). Inertial motion (vestibular) signals were provided by translating the animal on a motion platform, and optic flow (visual) signals were provided by a projector that was mounted on the platform[25,29]. In each trial, the monkey experienced forward motion with a small leftward or rightward component, and made an eye movement to report the perceived direction (Fig 1b, Methods). Three stimulus conditions were randomly interleaved: (1) a *Vestibular* condition in which heading was defined solely by inertial motion cues[29]; (2) a *Visual* condition in which heading was defined solely by optic flow; and (3) a *Combined* condition consisting of congruent inertial motion and optic flow.

Discrimination performance was quantified by constructing psychometric functions for each stimulus condition (Fig. 1c), and psychophysical thresholds were estimated from cumulative Gaussian fits[30] (see Methods). In Fig. 1c, the monkey's heading threshold was between 3.5–4° for each single cue condition (white circles and gray squares). Note that the reliability of the individual cues was roughly equated during training by reducing the coherence of visual motion (see Methods). This balancing of cues is crucial, as it affords the maximal opportunity to observe improvement in performance under cue combination[5]. In the combined condition (Fig. 1c, black triangles), the monkey's heading threshold was substantially smaller (1.5°), as evidenced by the steeper slope of the curve. Threshold data for every recording session from both animals are provided in Suppl. Fig. 1.

Across 57 sessions, monkey C had average psychophysical thresholds of 3.1±0.09° s.e. and 3.2±0.1° in the vestibular and visual conditions, respectively (Fig. 1d). If the monkey combined the two cues optimally[5], thresholds should be reduced by ~30% (see Methods). The average threshold measured in the combined condition (2.1±0.06°, Fig. 1d, black bar) did not differ significantly ($p=0.4$, paired t-test) from the optimal prediction (2.2±0.06°) computed from the single cue data (Fig. 1d, light gray bar). Moreover, the combined threshold was significantly smaller than the single cue thresholds ($p<<0.001$, paired t-tests), indicating better performance using both cues. For monkey A (Fig. 1e), the combined threshold (0.9±0.02°) was again significantly smaller than the single-cue thresholds ($p<<0.001$, paired t-tests) and was close to the optimal prediction (0.8±0.02°), although the difference was significant ($p=0.004$, paired t-

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

test). These data suggest that monkeys, like humans, can combine multiple sensory cues near optimally to improve perceptual performance.

Neuronal sensitivity during cue combination

Having established robust cue integration behavior, we recorded from single neurons in area MSTd while monkeys performed the heading task. Previously, we reported[25] that ~60% of MSTd neurons signal heading for both vestibular and optic flow stimuli. To identify these multi-modal neurons, we measured heading tuning curves in the horizontal plane while animals maintained visual fixation. Fig. 2a shows data from an example neuron with clear tuning under both single-cue conditions ($p<<0.001$, ANOVA). This neuron preferred leftward headings for both stimuli, and was classified as a 'congruent' cell. Note that heading direction always refers to real or simulated body motion, such that congruent cells have visual and vestibular tuning curves with aligned peaks.

Among 340 neurons tested, 194 (57%) showed significant heading tuning under both single-cue conditions ($p<0.05$, ANOVA), and these cells were studied further. Over the narrow range of headings sampled during the discrimination task, the tuning of the example congruent neuron was monotonic in all three stimulus conditions (Fig. 2b). For this cell, mean firing rate in the combined condition was greater than that in each of the single-cue conditions (Fig. 2b). This difference was greater for leftward than for rightward headings, such that the slope of the tuning curve around straight ahead (0°) became steeper in the combined condition.

To compare neuronal and behavioral sensitivity, we used signal detection theory[31,32] to quantify the ability of an ideal observer to discriminate heading based on the activity of a single neuron and its presumed ‘anti-neuron’ (see Methods; Fig. 2c). We computed neuronal thresholds from the standard deviation of the best-fitting cumulative Gaussian functions (Fig. 2c, smooth curves). For the example cell in Fig. 2 a–c, the neuronal threshold was smaller in the combined condition (1.8°) than in the visual (2.6°) or vestibular (5.1°) conditions, indicating that the neuron could discriminate smaller variations in heading when both cues were provided.

Fig. 2d–f shows analogous data for an ‘opposite’ cell that preferred leftward headings during visual stimulation but rightward headings during vestibular stimulation (Fig. 2d). During the discrimination task, responses to combined stimulation were intermediate between responses to the visual and vestibular conditions (Fig. 2e, compare black triangles to white circles and gray squares). As a result, the cell’s tuning curve was rather flat in the combined condition (Fig. 2e) and its neurometric function (Fig. 2f) was substantially shallower than those for the single-cue conditions. Neuronal threshold increased from 5.7° and 2.6° under the vestibular and visual conditions, respectively, to 40.8° during combined stimulation (Fig. 2f). Thus, this opposite cell carried less precise information about heading under cue combination.

Among 194 neurons recorded during the discrimination task, we obtained sufficient data from 129 cells (see Methods). Only the most sensitive neurons rivaled behavioral performance, whereas most neurons were substantially less sensitive than the animal (Suppl. Fig 2). To perform the task based on MSTd activity, the monkey must therefore either pool responses across many neurons or rely more heavily on the most sensitive neurons[33]. Note that stimuli in our task were not tailored to the tuning of individual neurons, such that many neurons have large thresholds because their tuning curves were flat over the range of headings tested (Suppl. Fig. 3). Neuronal sensitivity was greatest for neurons with heading preferences 60–90 degrees away from straight forward, as these neurons have tuning curves with near-maximal slopes around zero heading (Suppl. Fig. 4 A,B).

The question of central interest is whether neuronal thresholds, like the behavioral thresholds, are significantly lower for the combined condition than the single-cue conditions. Unlike the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

behavior, average neuronal thresholds across the entire population of MSTd neurons were not significantly lower for the combined condition than for the best single-cue condition ($p>0.7$, paired t-tests). Moreover, combined neuronal thresholds were significantly larger than the optimal predictions ($p<0.001$, paired t-tests). However, this average result is not surprising given the example neurons in Fig. 2. Whereas congruent cells generally showed improved sensitivity in the combined condition (Fig. 2c), opposite cells typically became less sensitive (Fig. 2f).

To summarize this dependence on single-cue tuning, we computed a quantitative index of congruency between visual and vestibular responses (Congruency Index, CI, see Methods). The CI will be near +1 when visual and vestibular tuning functions have a consistent slope (Fig. 2b), near −1 when they have opposite slopes (Fig. 2f), and near zero when either tuning function is flat (or even-symmetric) over the range of headings tested. The ratio of combined to predicted thresholds is plotted against CI for all 129 neurons in Fig. 3a. A significant negative correlation is observed ($R=-0.45$, $p<<0.001$, Spearman rank correlation) such that neurons with large positive CIs have thresholds close to optimal predictions (ratios near unity). In contrast, neurons with large negative CIs generally have combined/predicted threshold ratios well in excess of unity. We defined neurons with CIs significantly larger than 0 ($p<0.05$) as '*CI-congruent*' cells (Fig. 3a, cyan filled symbols, $CI>0$). Average neuronal thresholds for CI-congruent cells follow a pattern similar to the monkeys' behavior (Fig. 3b). Combined thresholds were significantly lower than both single-cue thresholds ($p<0.001$, paired t-tests) and were not significantly different from optimal predictions ($p=0.9$, paired t-test). Analogously, we defined neurons with CIs significantly smaller than 0 ($p<0.05$) as '*CI-opposite*' cells (Fig. 3a, magenta filled symbols, $CI<0$). Combined thresholds for CI-opposite cells were significantly greater than both single-cue thresholds (Fig. 3c, $p<0.02$, paired t-tests), indicating that these neurons became less sensitive during cue combination.

Neuronal thresholds for heading discrimination depend on two main aspects of neuronal responses: the slope (steepness) of the tuning curve around straight-ahead and the variance of the response. Higher sensitivity could arise either due to a steepening of the slope or a reduction in response variance. Thus, to understand how differences in neuronal thresholds between congruent and opposite cells arise, we considered each of these factors separately. Tuning curve slopes were obtained by linear regression, which generally provided acceptable fits to the data (median R^2 values for CI-congruent cells: 0.80 (visual), 0.87 (vestibular), and 0.85 (combined); median R^2 values for CI-opposite cells: 0.76 (visual), 0.67 (vestibular), and 0.35 (combined)). Nonlinear fits were also evaluated (see Suppl. Fig. 3). The slope of the tuning curve in the combined condition is plotted as a function of the slopes for the vestibular (Fig. 4a) and visual (Fig. 4b) conditions. In both scatter plots, data from CI-congruent cells (cyan) generally fall above the unity-slope diagonal ($p<0.001$, sign test) whereas data from CI-opposite cells (magenta) typically fall below the diagonal ($p<0.02$, sign test). Thus, as illustrated in Fig. 2, tuning slopes in the combined condition were generally steeper for congruent cells and flatter for opposite cells, as compared to single-cue conditions. This is further illustrated by comparing the ratios of combined to single-cue slopes (Fig. 4c). CI-congruent cells tend to cluster in the upper right quadrant (both ratios > 1), while the majority of CI-opposite cells are located in the lower left quadrant (both ratios < 1).

In contrast to these changes in slope, there was no substantial difference in Fano factor (variance to mean ratio) between the combined condition and single-cue conditions (Fig. 4d–f). Thus, differences in neuronal sensitivity between congruent and opposite cells are due mainly to differences in slope with little contribution from changes in Fano factor. This should not be taken to imply that response variance makes no contribution to neuronal sensitivity. Indeed, for each stimulus condition, multiple regression reveals a strong correlation between neuronal threshold and slope, as well as a significant but weaker correlation between threshold and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

response variance (Suppl. Fig. 5). In general, response variance does play a role in determining neuronal sensitivity but the difference in sensitivity between congruent and opposite cells during cue combination cannot be attributed to differences in response variability.

Bimodal responses can be predicted by linear summation

How are neuronal firing rates in the combined condition mathematically related to responses in the single-cue conditions? A recent theoretical study proposed that a population of neurons with Poisson-like firing rate statistics can combine cues optimally when bimodal responses are simply the sum of responses elicited by unimodal stimuli [28]. We tested whether a linear model with independent visual and vestibular weights can fit the bimodal responses of MSTd neurons (other variants of linear models with the same or fewer parameters are described in Supplementary Methods and Suppl. Fig. 6). Examples of model fits are shown in Fig. 2b and e (gray Xs). For both example neurons, the weighted linear model provides a good fit, and the weights on the visual and vestibular responses are less than unity. Additional examples of tuning curves, along with superimposed model fits, are shown in Suppl. Fig. 7.

Results from these linear fits are summarized in Fig. 5. As shown in Fig. 5a, predicted responses from the weighted linear sum model were strongly correlated with responses measured in the combined condition ($R=0.99$, $p<<0.001$), with a slope that was not significantly different from unity (95% confidence interval: [0.988, 1.003]) and an offset that was not significantly different from zero (95% confidence interval: [−0.167, 0.070]). Across the population of neurons, correlation coefficients between predicted and measured responses had a median value of 0.83 and were individually significant ($p<0.05$) for 89/129 (69%) cells (Fig. 5b, filled bars). The remaining 40/129 cells (31%) did not show significant correlation coefficients (Fig. 5b, light gray bars) mainly because the combined responses were noisy or the tuning curve was flat over the range of headings tested. Among neurons with R^2 values smaller than 0.5, the average slope of the tuning curve was only 0.08±0.02 spikes/s/degree (n=26, mean ± sem) and only 23% (6/26) of these cells showed significant tuning in the combined condition (ANOVA, $p<0.05$). For cells with R^2 values larger than 0.5, the average tuning slope was 0.60±0.06 spikes/s/degree (n=103) and 89% (92/103) showed significant tuning in the combined condition.

Visual and vestibular weights derived from the linear model are shown for the population of MSTd neurons in Fig 5c, and a few aspects of these data are noteworthy. First, average weights were significantly less than unity ($p<<0.001$, t-tests), with a median vestibular weight of 0.6 and a median visual weight of 0.76. Thus, multisensory integration by MSTd neurons is typically sub-additive during heading discrimination. Second, there is a significant negative correlation between vestibular and visual weights ($R=-0.40$, $p<0.001$, Spearman rank correlation) such that neurons with large visual weights tend to have small vestibular weights, and vice-versa. This suggests that neurons vary continuously along a range from visually-dominant to vestibularly-dominant under the conditions of our experiment. Third, visual and vestibular weights do not depend on CI ($p>0.5$, Spearman rank correlation), indicating that the same linear weighting of visual and vestibular inputs is employed by congruent and opposite cells.

Correlations between neural activity and behavior

If monkeys rely on area MSTd to perform multisensory integration in the heading discrimination task, one should expect to find significant correlations between neuronal activity and behavior. To test this hypothesis, we computed choice probabilities (CPs)[34] to quantify whether trial-to-trial fluctuations in neural firing rates were correlated with fluctuations in the monkeys' perceptual decisions (for a constant physical stimulus). A significant CP>0.5 indicates that the monkey tends to choose the neuron's preferred sign of heading (leftward vs. rightward) when the neuron fires more strongly[29]. This result is thought to reflect a functional

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

link between the neuron and perception[33,35]. Of critical interest here is whether MSTd neurons show robust CPs in the combined condition and whether these effects depend on the congruency of visual and vestibular tuning. Across all 129 MSTd neurons tested under the combined condition, the average CP was modestly larger than chance (mean CP=0.52, p=0.003). Despite this significant bias in the expected direction, roughly equal numbers of neurons had individually significant CPs that were greater or less than 0.5 (Fig. 6a, filled symbols). Neurons with a significant CP < 0.5 paradoxically increase their firing rates when the monkey chooses their non-preferred direction.

This diversity of CPs in the combined condition is related to visual/vestibular congruency, as evidenced by a highly significant positive correlation between CP and CI (R=0.44, p<<0.001, Fig. 6a). The average CP for CI-congruent cells (cyan) is substantially greater than 0.5 (mean=0.58, p<<0.001, t-test), suggesting that these cells are strongly coupled to perceptual decisions under cue combination. For CI-opposite cells (magenta), however, the average CP is close to, and even slightly less than, 0.5 (mean=0.48, p=0.08). As a result, the average CP in the combined condition is significantly greater for CI-congruent than CI-opposite cells (p<<0.001, t-test). This result suggests that the animals may monitor the activity of congruent cells more closely than opposite cells during cue integration. Alternatively, this finding may reflect stronger correlations among congruent cells than opposite cells (see Discussion).

If significant CPs reflect a functional linkage between neurons and perception, then more sensitive neurons may show larger CPs. This relationship might arise either because the animal selectively monitors the most informative neurons[36] and/or because sensitive neurons are more strongly correlated with each other [37]. Indeed, CPs were negatively correlated with neuronal thresholds in the combined condition (Fig. 6b; R = −0.31, p<0.0003, Spearman rank correlation). By virtue of their steeper tuning slopes under cue combination, CI-congruent cells (cyan) tend to have low neuronal thresholds and high CPs, such that data from these neurons cluster in the upper left corner of Fig. 6b. In contrast, CI-opposite cells tend to have high neuronal thresholds and low CPs. Together, the data from CI-congruent and CI-opposite cells (Fig. 6b) form a cloud of points that shows a robust negative correlation, suggesting that the difference in CP between congruent and opposite cells may be driven largely by the difference in sensitivity.

How does this dependence of bimodal CPs on visual/vestibular congruency arise? Previously[29], we reported significant CPs for MSTd neurons under the vestibular condition (mean=0.55, p<<0.001). These CPs depend modestly on heading preference such that neurons with heading preferences 30–90 degrees away from straight ahead tend to have larger CPs (Suppl. Fig. 4 C,D). Unlike in the combined condition, the vast majority of neurons (96%, 23/24) with significant vestibular CPs had values larger than 0.5, and vestibular CPs were similar for CI-congruent and CI-opposite cells (p=0.06, Suppl. Fig. 8a). Both cell types showed mean vestibular CPs that were significantly greater than 0.5 (p <<0.001, t-tests). Thus, the strong dependence of combined CPs on congruency is not explained by a similar dependence under the vestibular condition.

In contrast, CPs in the visual condition depended strongly on congruency. Across all 129 neurons, the average CP was significantly lower in the visual condition (mean=0.52, p=0.01) than in the vestibular condition (p<0.008, paired t-tests). Moreover, roughly equal numbers of neurons had visual CPs significantly greater and less than 0.5 (Fig. 6c, black filled bars). This finding was linked to a highly significant correlation between CP and CI in the visual condition (Fig. 6c; R=0.51, p<<0.001). The average CP for CI-congruent cells (cyan) was substantially greater than 0.5 (0.59, p<<0.001), whereas CI-opposite cells (magenta) had an average visual CP significantly less than 0.5 (mean=0.45, p=0.02). Unlike vestibular and combined responses, there was no significant correlation between CPs and neuronal thresholds in the visual

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

condition (Fig. 6d, R=−0.1, p>0.2). Whereas CI-congruent and CI-opposite cells appear to fall along the same trend in the combined condition (Fig. 6b), this is not the case for the visual condition (Fig. 6d). Among the most sensitive neurons, CI-opposite cells tend to have CPs < 0.5 whereas CI-congruent cells have CPs consistently > 0.5.

Differences in CPs between stimulus conditions are shown, on a cell-by-cell basis, in Fig. 7. The difference in CP between visual and vestibular conditions is significantly correlated with CI (Fig. 7a, R=0.33, p<<0.001, Spearman rank correlation). A similar pattern of results holds for the difference in CP between combined and vestibular conditions (Fig. 7b, R=0.26, p<0.004).

In summary, our findings show that vestibular responses in MSTd were consistently correlated with heading percepts irrespective of congruency, whereas the correlation between visual signals and heading percepts changed sign with congruency. It is important to note that CPs were computed by identifying a neuron's preferred sign of heading (leftward or rightward) independently for each stimulus condition. Thus, in the visual condition, opposite cells with CP < 0.5 tend to fire more strongly when the monkeys report their non-preferred sign of heading. Of course, for these neurons, their non-preferred sign of heading in the visual condition is their preferred sign of heading in the vestibular condition. This suggests that, in the visual condition, responses of opposite cells may be decoded relative to their vestibular preference. Most importantly, in the combined condition, congruent cells were more strongly correlated with monkeys' heading judgments than were opposite cells (see Discussion).

Temporal evolution of thresholds and CPs

Population responses in all three stimulus conditions roughly follow the Gaussian velocity profile of the stimulus (Fig. 8a) [25,29]. The analyses summarized above were based on mean firing rates computed from the middle 1s of the 2s stimulus period (containing most of the velocity variation). We also examined how neuronal thresholds and CPs evolved as a function of time, focusing on the latter 1.5 seconds of the stimulus period when the neuronal responses were robust. Within this time range, we repeated threshold and CP analyses for several 500ms time windows spaced 100 ms apart.

Overall, neuronal thresholds were relatively high during early and late time windows, and were lowest during the middle windows. For CI-congruent cells, thresholds measured in the combined condition (Fig. 8b, solid blue curve) were close to those predicted by theory, with the closest agreement seen during the period of peak discharge (Fig. 8b, solid blue vs. orange curves). For CI-opposite cells, combined thresholds were consistently higher than predictions across time (Fig. 8b, dashed blue vs. orange curves). With regard to choice probabilities, CI-congruent cells consistently showed values significantly larger than 0.5 (p<0.008, t-tests, Fig. 8c, solid curves) for all time windows. For CI-opposite cells, vestibular CPs were consistently larger than 0.5 in most time windows, whereas visual CPs were consistently smaller than 0.5 across time, and combined CPs hovered around 0.5 (Fig. 8c, dashed curves). Thus, our main findings generalize to all time periods during which neural responses were robust. Significant CPs develop early during the response, at least for congruent cells, and remain elevated throughout and beyond the stimulus period.

Discussion

Our behavioral results indicate that monkeys integrate visual and vestibular signals to discriminate heading with greater precision than allowed by either cue alone. The improvement in performance is close to that predicted by optimal cue integration theory, as seen in human studies of multi-sensory integration[5,6]. This suggests that humans and non-human primates may employ similar strategies for multi-sensory integration, and establishes a model system

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

to explore the neural mechanisms underlying perceptual cue integration. By recording from single neurons in area MSTd, we identified a population of 'congruent' cells that also show improved sensitivity during cue combination, and could account for the improvement in behavioral performance. In contrast, 'opposite cells' show reduced sensitivity during cue combination. These findings suggest that behavior may rely more heavily on congruent cells than opposite cells when both cues are provided. Consistent with this idea, congruent cells are more strongly correlated with perceptual decisions during cue integration than opposite cells (Fig. 6a, Fig. 8c). Overall, these findings identify a candidate neural substrate that may integrate visual and vestibular cues to allow robust self-motion perception.

Neuronal sensitivity

We find that congruent cells show increased sensitivity during cue combination whereas opposite cells show reduced sensitivity. What aspects of MSTd responses account for this difference? In principle, increased sensitivity could be achieved by either increasing the slope of the tuning curve or by reducing the variance of responses. Our data show clearly that slopes are increased in the combined condition for congruent cells and reduced for opposite cells (Fig. 4c). This slope increase for congruent cells might arise simply because responses are larger in the combined condition than in the single-cue conditions. While this appears to be the case in Fig. 2b, it was not true generally. For other congruent cells (e.g., Suppl. Fig. 7b), tuning curve slope was greater in the combined condition despite the fact that combined responses were lower than those elicited by the most effective single cue. Across the population of CI-congruent cells, there is no significant correlation ($R=0.15$, $p>0.09$, spearman rank correlation) between the slope increase in the combined condition and the ratio of firing rates between the combined and single-cue conditions (Suppl. Fig. 9, cyan). Furthermore, the ratio of combined/single-cue firing rates is similar for CI-opposite cells, yet slopes in the combined condition are consistently reduced for these cells (Suppl. Fig. 9, magenta). Thus, improved sensitivity of MSTd neurons under cue combination is not simply a result of greater responses.

A reduction in response variance could also contribute to the increased sensitivity of congruent cells. However, Fano factor does not differ substantially between congruent and opposite cells (Fig. 4, d–f). Hence the difference in sensitivity between congruent and opposite cells arises primarily from differences in the slope of the response function, and these slope changes are well accounted for by linear weighted summation of single-cue responses. These findings are consistent with predictions of a recent theoretical study[28] which proposes that a population of neurons can implement Bayesian-optimal cue integration if they linearly sum their inputs and obey a family of Poisson-like statistics. Note, however, that linear combination of perceptual estimates at the level of behavior, as often seen in human studies, does not necessarily imply linear weighted summation at the level of single neurons. Indeed, congruent and opposite cells show comparable linear summation (Fig. 5c), but this produces greater sensitivity for congruent cells and poorer sensitivity for opposite cells. Thus, while neural cue combination may be well described by linear weighted summation, selective decoding and/or correlation of neuronal responses appears necessary to predict the behavioral effects from MSTd responses.

Choice probabilities

We previously reported significant choice probabilities for MSTd neurons in the vestibular condition of the heading discrimination task[29], suggesting that vestibular signals in MSTd contribute to perceptual decisions regarding heading. Our current findings show that this relationship holds for both congruent and opposite cells (Suppl. Fig. 8a), indicating that vestibular signals in MSTd are consistently correlated with perceptual decisions independent of the congruency with visual selectivity.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

In contrast, CPs in the visual condition depend strongly on congruency. For CI-congruent cells, the average visual CP (0.59) is substantially greater than chance (0.5). When CPs have been observed previously[34,36,38–42], this is the relationship typically seen—stronger firing when the animal reports a preferred stimulus for the neuron. By contrast, CI-opposite cells have an average visual CP (0.45) significantly less than 0.5, indicating that they tend to fire more strongly when the monkey reports their non-preferred sign of heading. This suggests that both CI-congruent and CI-opposite cells contribute to purely visual judgments of heading, but that their activities bear different relationships to behavior.

How might we explain this difference between CI-congruent and CI-opposite cells in the visual condition? If heading is represented in area MSTd in the form of a place code (or labeled-line code), then the activity level of each neuron may be interpreted as the degree to which a stimulus matches the heading preference of the neuron. Typically, in such codes, each neuron is assumed to have a fixed preference for each particular stimulus dimension. By definition, the heading preference of opposite neurons is different between vestibular and visual conditions. Thus, in a place code, strong activity of an opposite cell could represent how well stimuli match either the vestibular preference or the visual preference. Thus, one possible explanation for our findings is that responses of opposite cells in the visual condition are decoded with respect to their vestibular preference. This would account for significant CPs less than 0.5 for CI-opposite cells. Why might activity of opposite cells be decoded as evidence in favor of their vestibular, not visual, preference? We speculate that this decoding strategy may have arisen due to our protocol for training animals to perform the heading task. We initially trained animals to discriminate heading based solely on vestibular cues, and subsequently added the random-dot stimuli and gradually increased motion coherence until thresholds in the combined condition were reduced relative to the vestibular condition. Only then were visual condition trials introduced. Thus, animals may have learned to decode MST responses with respect to their vestibular preferences because the task initially relied upon vestibular cues. By this logic, the monkey should monitor opposite cells during the visual condition because these neurons frequently provide reliable information (Fig. 6d).

An alternative explanation for the visual CP results is that CI-opposite cells are selectively correlated with CI-congruent cells that have the same vestibular preference. Trial-to-trial correlations in response among neurons are thought to be important—and may be essential—for observing CPs[37], such that differences in CP across stimulus conditions may be due to variations in correlated noise among the neurons. If CI-opposite cells are correlated with CI-congruent cells that have matching vestibular preferences, this could also explain why CI-opposite cells show CP values below 0.5 in the visual condition.

In the combined condition, CI-congruent cells have an average CP (0.58) that is substantially greater than chance ($p<<0.001$, t-test), whereas CI-opposite cells have an average CP (0.48) slightly less than chance ($p=0.08$, t-test). While we cannot firmly exclude the possibility that CI-opposite cells have some correlation with perceptual decisions during cue combination, it seems clear that CI-congruent cells are more strongly linked to perceptual decisions. This may be sensible given that congruent cells are substantially more sensitive than opposite cells under cue combination (Fig. 3 b,c). The CP and neuronal threshold data from CI-congruent and CI-opposite cells combine to form a cloud of data with a clear negative correlation (Fig. 6b), suggesting that the difference in CPs between these cell classes may be driven by the difference in sensitivity. Similar negative correlations between CPs and neuronal thresholds have been observed in other studies[34,36,41–43]. Thus, one interpretation of the data is that CI-opposite cells are given less weight during cue integration because they carry less reliable heading information.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

An alternative possibility is that CI-congruent cells exhibit greater interneuronal correlations than CI-opposite cells in the combined condition, thus leading to larger CPs for CI-congruent cells. We examined this possibility by computing noise correlations between single-unit and multi-unit activity recorded from the same electrode, as these noise correlations have been reported to be predictive of significant CPs[39]. We found robust noise correlations between single-unit and multi-unit responses, but did not observe any dependence of these correlations on congruency (CI) for any of the stimulus conditions (Suppl. Fig. 10). While these data do not exclude a contribution of noise correlations to our findings, they support the possibility that variations in CP may be linked to selective decoding of MSTd neurons[36]. It has not been demonstrated directly, however, that selective decoding can lead to variations in CP without corresponding changes in the structure of noise correlations. Indeed, Shadlen et al.[37] have provided evidence to the contrary based on simulations of how populations of MT neurons contribute to motion discrimination.

Given that CI-opposite neurons do not contribute to improved sensitivity during cue combination, what purpose do these neurons serve? CI-opposite cells may play important roles in parsing retinal image motion into components related to self-motion and object motion. Specifically, differences in activity between populations of CI-congruent and CI-opposite neurons may help to identify retinal image motion that is inconsistent with self-motion and therefore results from moving objects in the scene.

In closing, our findings implicate area MSTd in sensory integration for heading perception and establish an excellent model system for studying the mechanisms by which neurons combine different sensory signals to optimize performance. Future experiments can probe for causal links between MSTd neurons and heading perception during cue integration and can test whether neurons change their weighting of visual and vestibular cues dynamically as the reliability of cues varies. Ultimately, these studies should lead to a deeper understanding of how populations of neurons mediate probabilistic (e.g., Bayesian) inference.

Methods

Motion stimuli

Two rhesus monkeys (*Macaca mulatta*) weighing ~6 kg were trained using a virtual reality system[25]. Translation of the monkey in the horizontal plane was accomplished by a motion platform (MOOG 6DOF2000E; Moog, East Aurora, NY). To activate vestibular otolith organs, each inertial motion stimulus followed a smooth trajectory with a Gaussian velocity profile and a peak-acceleration of ~1 m/s^2.

A projector (Christie Digital Mirage 2000) was mounted on the motion platform and rear-projected images (90 × 90° of visual angle) onto a tangent screen. Visual stimuli depicted movement through a 3D cloud of ‘stars’ that occupied a virtual space 100 cm wide, 100 cm tall, and 50 cm deep. Star density was $0.01/cm^3$, with each star being a 0.15cm × 0.15cm triangle. Stimuli were presented stereoscopically as red/green anaglyphs, viewed through Kodak Wratten filters (red #29, green #61). The stimulus contained a variety of depth cues, including horizontal disparity, motion parallax and size information. Motion coherence was manipulated by randomizing the 3D location of a percentage of stars on each display update while the remaining stars moved according to the specified heading. This manipulation degrades optic flow as a heading cue and was used to reduce psychophysical sensitivity in the visual condition such that it matched vestibular sensitivity.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Behavioral task

Monkeys were trained to perform a heading discrimination task around psychophysical threshold. In each trial, the monkey experienced forward motion with a small leftward or rightward component (angle α, Fig 1a). Monkeys were required to maintain fixation on a head-fixed visual target located at the center of the display screen. Trials were aborted if conjugate eye position deviated from a 2 × 2° electronic window around the fixation point. At the end of the 2s trial, the fixation spot disappeared, two choice targets appeared, and the monkey made a saccade to one of the targets to report his perceived motion as leftward or rightward relative to straight ahead (Fig. 1b). Across trials, heading was varied in fine steps around straight ahead. The range of headings was chosen based on extensive psychophysical testing using a staircase paradigm[29]. Nine logarithmically spaced heading angles were tested for each monkey including an ambiguous straight-forward direction (monkey A: ±9°, ±3.5°, ±1.3°, ±0.5° and 0°; monkey C: ±16°, ±6.4°, ±2.5°, ±1° and 0°). These values were chosen carefully to obtain near-maximal psychophysical performance while allowing neural sensitivity to be estimated reliably for most neurons. All animal procedures were approved by the Institutional Animal Care and Use Committee at Washington University and were in accordance with National Institutes of Health guidelines.

The experiment consisted of three randomly-interleaved stimulus conditions: (1) In the *Vestibular* condition, the monkey was translated by the motion platform while fixating a head-fixed target on a blank screen. There was no optic flow, except for that produced by small fixational eye movements. Performance in this condition depends heavily on vestibular signals[29]. (2) In the *Visual* condition, the motion platform remained stationary while optic flow simulated the same range of headings. (3) In the *Combined* condition, congruent inertial motion and optic flow were provided [25]. Each of the 27 unique stimulus conditions (9 headings × 3 cue conditions) was typically repeated ~30 times, for a total of ~800 discrimination trials per recording session.

Electrophysiology

Extra-cellular single-unit recording was carried out as described previously[25,29]. MSTd was located by structural MRI imaging and mapping of physiological response properties, such as direction selectivity for visual motion and visual receptive fields encompassing a large proportion of the contra-lateral visual field including the fovea[44–48].

Once the action potential of a single neuron was isolated, we measured heading tuning in the horizontal plane (10 directions relative to straight-ahead: 0°, ±22.5°, ±45°, ±90°, ±135° and 180°) under both the vestibular and visual conditions (Fig. 2a, d). For this measurement, monkeys were simply required to fixate a head-centered target while 4–5 repetitions were collected for each stimulus. Only MSTd neurons with significant tuning in both vestibular and visual conditions ($p<0.05$, ANOVA) were tested during the heading discrimination task.

Data analysis

To quantify behavioral performance, we plotted the proportion of 'rightward' decisions as a function of heading (Fig. 1c), and we fit these psychometric functions with a cumulative Gaussian[30]. The psychophysical threshold for each stimulus condition was taken as the standard deviation parameter of the Gaussian fit.

Predicted thresholds for the combined condition, assuming optimal (maximum likelihood) cue integration, were computed as[5]:

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

$$\sigma_{\text{prediction}} = \sqrt{\frac{\sigma^2_{\text{vestibular}} \times \sigma^2_{\text{visual}}}{\sigma^2_{\text{vestibular}} + \sigma^2_{\text{visual}}}} \quad (1)$$

where $\sigma_{\text{vestibular}}$ and σ_{visual} represent psychophysical thresholds in the vestibular and visual conditions, respectively.

Neural responses were quantified as mean firing rates over the middle 1 s interval of each stimulus presentation (see Fig. 8 for other time windows). To characterize neuronal sensitivity, we used receiver operating characteristic (ROC) analysis to compute the ability of an ideal observer to discriminate between two oppositely-directed headings (e.g. +1° vs. −1°) based solely on the firing rate of the recorded neuron and a presumed 'anti-neuron' with opposite tuning[29,31]. Neurometric functions were constructed from these ROC values and were fit with cumulative Gaussian functions to determine neuronal thresholds.

To quantify the relationship between MSTd responses and perceptual decisions, we computed "choice probabilities" using ROC analysis[29,34]. For each heading direction, neuronal responses were sorted into two groups based on the animal's choice at the end of each trial (i.e., 'preferred' versus 'null' choices). ROC values were calculated from these two distributions whenever there were at least 3 choices in each group, and this yielded a choice probability (CP) for each heading direction. We combined data across headings (following z-score normalization) to compute a grand CP for each cue condition[29]. The statistical significance of CPs (relative to the chance level of 0.5) was determined using permutation tests (1000 permutations).

Note that, for opposite cells, the definition of "preferred" and "null" choices is different for the vestibular and visual conditions. In computing CPs, we defined preferred and null choices according to the tuning of the neuron in each particular stimulus condition. Thus, if the opposite neuron of Fig. 2d consistently responds more strongly when the monkey reports rightward movement, it will have a CP > 0.5 for the vestibular condition and a CP < 0.5 for the visual condition.

To quantify the congruency between visual and vestibular tuning functions measured during discrimination, we calculated a congruency index (CI). A Pearson correlation coefficient was first computed for each single-cue condition. This quantified the strength of the linear trend between firing rate and heading for vestibular ($R_{\text{vestibular}}$) and visual (R_{visual}) stimuli. CI was defined as the product of these two correlation coefficients:

$$CI = R_{\text{vestibular}} \times R_{\text{visual}} \quad (2)$$

CI ranges from −1 to 1 with values near 1 indicating that visual and vestibular tuning functions have a consistent slope (Fig. 2b), whereas values near −1 indicate opposite slopes (Fig. 2d). Note that CI reflects both the congruency of tuning and the steepness of the slopes of the tuning curves around straight ahead. CI was considered to be significantly different from zero when both of the constituent R values were significant ($p < 0.05$). We denote neurons having values of CI significantly different from zero as CI-congruent (CI > 0) or CI-opposite (CI < 0). We also examined a global measure of visual-vestibular congruency (see Supplementary Methods) and obtained similar results using this measure (Suppl. Figure 11 and Suppl. Figure 12).

We used a linear weighted summation model to predict responses during cue combination from responses to each single cue condition:

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

$$R_{\text{prediction}} = w_{\text{vestibular}} \times R_{\text{vestibular}} + w_{\text{visual}} \times R_{\text{visual}} \quad (3)$$

where $R_{vestibular}$ and R_{visual} are responses from the single-cue conditions, and $w_{vestibular}$ and w_{visual} represent weights applied to the vestibular and visual responses, respectively. The weights were determined by minimizing the sum squared error between predicted responses and measured responses in the combined condition. Weights were constrained to lie between −20 and +20. The correlation coefficient (R) from a linear regression fit, which ranges from −1 to 1, was used to assess goodness of fit. We also evaluated three variants of the linear model, as described in Supplementary Methods and Suppl. Fig. 6.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Amanda Turner and Erin White for excellent monkey care and training. We are grateful to L. Snyder, A. Pouget, and R. Moreno Bote for helpful comments on an earlier version of the manuscript. This work was supported by NIH EY017866 (to DEA) and by NIH EY016178 and an EJLB Foundation grant (to GCD).

REFERENCES

1. Jacobs RA. Optimal integration of texture and motion cues to depth. Vision Res 1999;39:3621–3629. [PubMed: 10746132]
2. Hillis JM, Watt SJ, Landy MS, Banks MS. Slant from texture and disparity cues: optimal cue combination. J Vis 2004;4:967–992. [PubMed: 15669906]
3. Knill DC, Saunders JA. Do humans optimally integrate stereo and texture information for judgments of surface slant? Vision Res 2003;43:2539–2558. [PubMed: 13129541]
4. van Beers RJ, Sittig AC, Gon JJ. Integration of proprioceptive and visual position-information: An experimentally supported model. J Neurophysiol 1999;81:1355–1364. [PubMed: 10085361]
5. Ernst MO, Banks MS. Humans integrate visual and haptic information in a statistically optimal fashion. Nature 2002;415:429–433. [PubMed: 11807554]
6. Alais D, Burr D. The ventriloquist effect results from near-optimal bimodal integration. Curr Biol 2004;14:257–262. [PubMed: 14761661]
7. Molholm S, Ritter W, Javitt DC, Foxe JJ. Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. Cereb Cortex 2004;14:452–465. [PubMed: 15028649]
8. Yuille, AL.; Bulthoff, HH. Bayesian decision theory and psychophysics. In: Knill, D.; Richards, W., editors. Perception as Bayesian Inference. Cambridge University Press; 1996. p. 123-161.
9. Mamassian, P.; Landy, MS.; Maloney, LT. Bayesian modelling of visual perception. In: Rao, PN., et al., editors. Probabilistic Models of the Brain. 2002. p. 13-36.
10. Knill DC, Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci 2004;27:712–719. [PubMed: 15541511]
11. Stein, BE.; Meredith, MA. The merging of the senses. The MIT Press; 1993.
12. Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex 1991;1:1–47. [PubMed: 1822724]
13. Ghazanfar AA, Schroeder CE. Is neocortex essentially multisensory? Trends Cogn Sci 2006;10:278–285. [PubMed: 16713325]
14. Wallace MT, Ramachandran R, Stein BE. A revised view of sensory cortical parcellation. Proc Natl Acad Sci U S A 2004;101:2167–2172. [PubMed: 14766982]
15. Stanford TR, Quessy S, Stein BE. Evaluating the operations underlying multisensory integration in the cat superior colliculus. J Neurosci 2005;25:6499–6508. [PubMed: 16014711]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

16. Sugihara T, Diltz MD, Averbeck BB, Romanski LM. Integration of auditory and visual communication information in the primate ventrolateral prefrontal cortex. J Neurosci 2006;26:11138–11147. [PubMed: 17065454]
17. Avillac M, Ben Hamed S, Duhamel JR. Multisensory integration in the ventral intraparietal area of the macaque monkey. J Neurosci 2007;27:1922–1932. [PubMed: 17314288]
18. Wuerger SM, Hofbauer M, Meyer GF. The integration of auditory and visual motion signals at threshold. Percept Psychophys 2003;65:1188–1196. [PubMed: 14710954]
19. Tanaka K, et al. Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. J Neurosci 1986;6:134–144. [PubMed: 3944614]
20. Tanaka K, Fukada Y, Saito HA. Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. J Neurophysiol 1989;62:642–656. [PubMed: 2769352]
21. Duffy CJ, Wurtz RH. Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. J Neurophysiol 1991;65:1329–1345. [PubMed: 1875243]
22. Duffy CJ, Wurtz RH. Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. J Neurosci 1995;15:5192–5208. [PubMed: 7623145]
23. Duffy CJ. MST neurons respond to optic flow and translational movement. J Neurophysiol 1998;80:1816–1827. [PubMed: 9772241]
24. Page WK, Duffy CJ. Heading representation in MST: sensory interactions and population encoding. J Neurophysiol 2003;89:1994–2013. [PubMed: 12686576]
25. Gu Y, Watkins PV, Angelaki DE, DeAngelis GC. Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. J Neurosci 2006;26:73–85. [PubMed: 16399674]
26. Takahashi K, et al. Multimodal coding of three-dimensional rotation and translation in area MSTd: comparison of visual and vestibular selectivity. J Neurosci 2007;27:9742–9756. [PubMed: 17804635]
27. Bremmer F, Kubischik M, Pekel M, Lappe M, Hoffmann KP. Linear vestibular self-motion signals in monkey medial superior temporal area. Ann N Y Acad Sci 1999;871:272–281. [PubMed: 10372078]
28. Ma WJ, Beck JM, Latham PE, Pouget A. Bayesian inference with probabilistic population codes. Nat Neurosci 2006;9:1432–1438. [PubMed: 17057707]
29. Gu Y, Deangelis GC, Angelaki DE. A functional link between area MSTd and heading perception based on vestibular signals. Nat Neurosci 2007;10:1038–1047. [PubMed: 17618278]
30. Wichmann FA, Hill NJ. The psychometric function: I. Fitting, sampling, and goodness of fit. Percept Psychophys 2001;63:1293–1313. [PubMed: 11800458]
31. Britten KH, Shadlen MN, Newsome WT, Movshon JA. The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci 1992;12:4745–4765. [PubMed: 1464765]
32. Green, DM.; Swets, JA. Signal detection theory and psychophysics. New York: Wiley; 1966.
33. Parker AJ, Newsome WT. Sense and the single neuron: probing the physiology of perception. Annu Rev Neurosci 1998;21:227–277. [PubMed: 9530497]
34. Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA. A relationship between behavioral choice and the visual responses of neurons in macaque MT. Vis Neurosci 1996;13:87–100. [PubMed: 8730992]
35. Krug K. A common neuronal code for perceptual processes in visual cortex? Comparing choice and attentional correlates in V5/MT. Philos Trans R Soc Lond B Biol Sci 2004;359:929–941. [PubMed: 15306408]
36. Purushothaman G, Bradley DC. Neural population code for fine perceptual decisions in area MT. Nat Neurosci 2005;8:99–106. [PubMed: 15608633]
37. Shadlen MN, Britten KH, Newsome WT, Movshon JA. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. J Neurosci 1996;16:1486–1510. [PubMed: 8778300]
38. Nienborg H, Cumming BG. Psychophysically measured task strategy for disparity discrimination is reflected in V2 neurons. Nat Neurosci 2007;10:1608–1614. [PubMed: 17965712]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

39. Nienborg H, Cumming BG. Macaque V2 neurons, but not V1 neurons, show choice-related activity. J Neurosci 2006;26:9567–9578. [PubMed: 16971541]
40. Dodd JV, Krug K, Cumming BG, Parker AJ. Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. J Neurosci 2001;21:4809–4821. [PubMed: 11425908]
41. Celebrini S, Newsome WT. Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. J Neurosci 1994;14:4109–4124. [PubMed: 8027765]
42. Uka T, DeAngelis GC. Contribution of area MT to stereoscopic depth perception: choice-related response modulations reflect task strategy. Neuron 2004;42:297–310. [PubMed: 15091344]
43. Parker AJ, Krug K, Cumming BG. Neuronal activity and its links with the perception of multi-stable figures. Philos Trans R Soc Lond B Biol Sci 2002;357:1053–1062. [PubMed: 12217173]
44. Tanaka K, Sugita Y, Moriya M, Saito H. Analysis of object motion in the ventral part of the medial superior temporal area of the macaque visual cortex. J Neurophysiol 1993;69:128–142. [PubMed: 8433128]
45. Van Essen DC, Maunsell JH, Bixby JL. The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographic organization. J Comp Neurol 1981;199:293–326. [PubMed: 7263951]
46. Komatsu H, Wurtz RH. Relation of cortical areas MT and MST to pursuit eye movements. III. Interaction with full-field visual stimulation. J Neurophysiol 1988;60:621–644. [PubMed: 3171645]
47. Komatsu H, Wurtz RH. Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. J Neurophysiol 1988;60:580–603. [PubMed: 3171643]
48. Desimone R, Ungerleider LG. Multiple visual areas in the caudal superior temporal sulcus of the macaque. J Comp Neurol 1986;248:164–189. [PubMed: 3722457]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

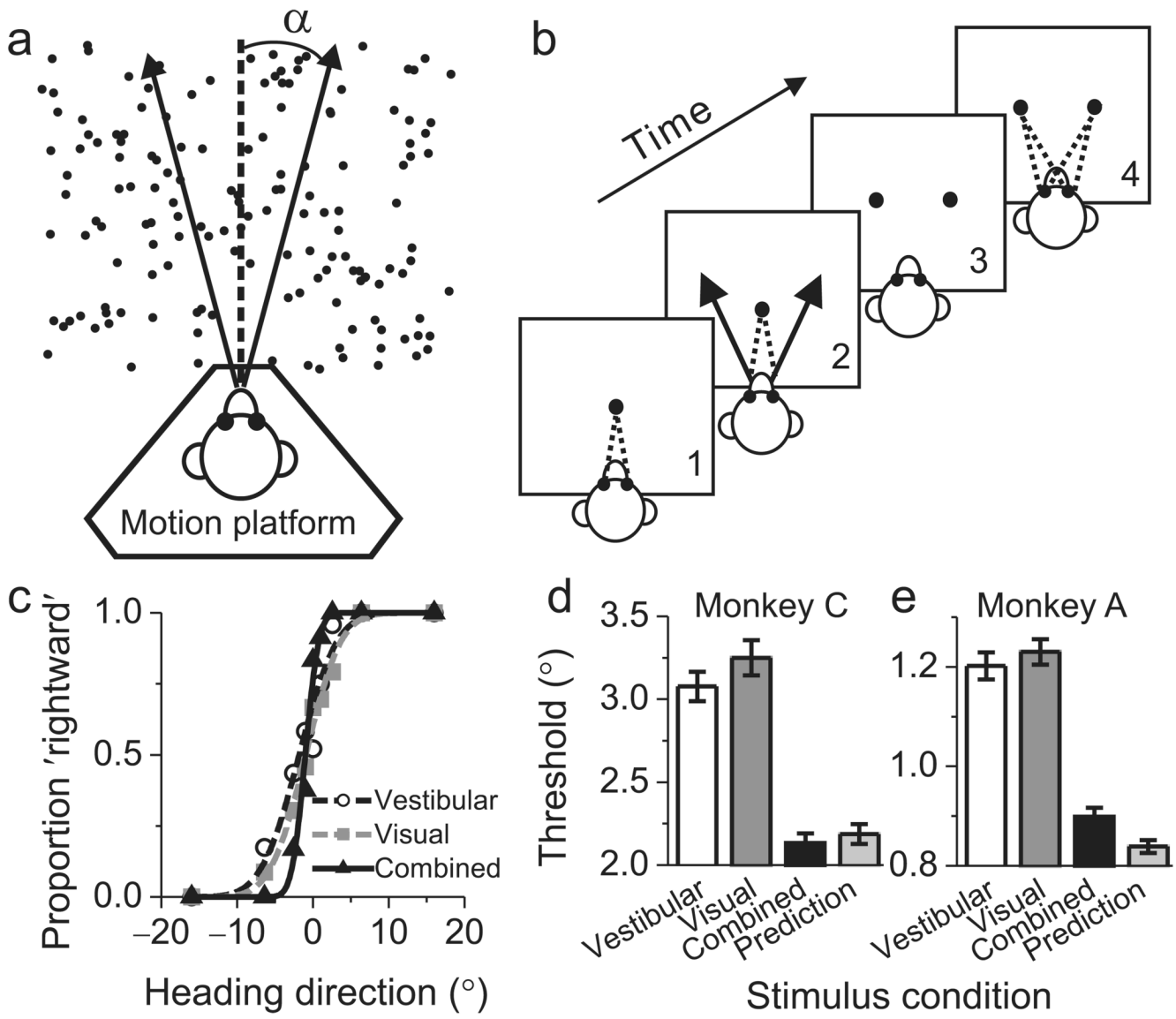

Figure 1.

Heading task and behavioral performance. (**a**) Monkeys were seated on a motion platform and translated within the horizontal plane. A projector mounted on the platform displayed images of a 3D star field, thus providing optic flow. (**b**) After fixating a visual target, the monkey experienced forward motion with a small leftward or rightward (arrow) component, and subsequently reported his perceived heading ('left' vs. 'right') by making a saccadic eye movement to one of two targets. (**c**) Example psychometric functions from one session (~30 stimulus repetitions). The proportion of 'rightward' decisions is plotted as a function of heading, and smooth curves are best fitting cumulative Gaussian functions. White circles, gray squares, and black triangles represent data from the vestibular, visual and combined conditions, respectively. (**d, e**) Average psychophysical thresholds from two monkeys (Monkey C, N=57; Monkey A, N=72) for the 3 stimulus conditions, and predicted thresholds computed from optimal cue integration theory (light gray bars). Error bars: SEM.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

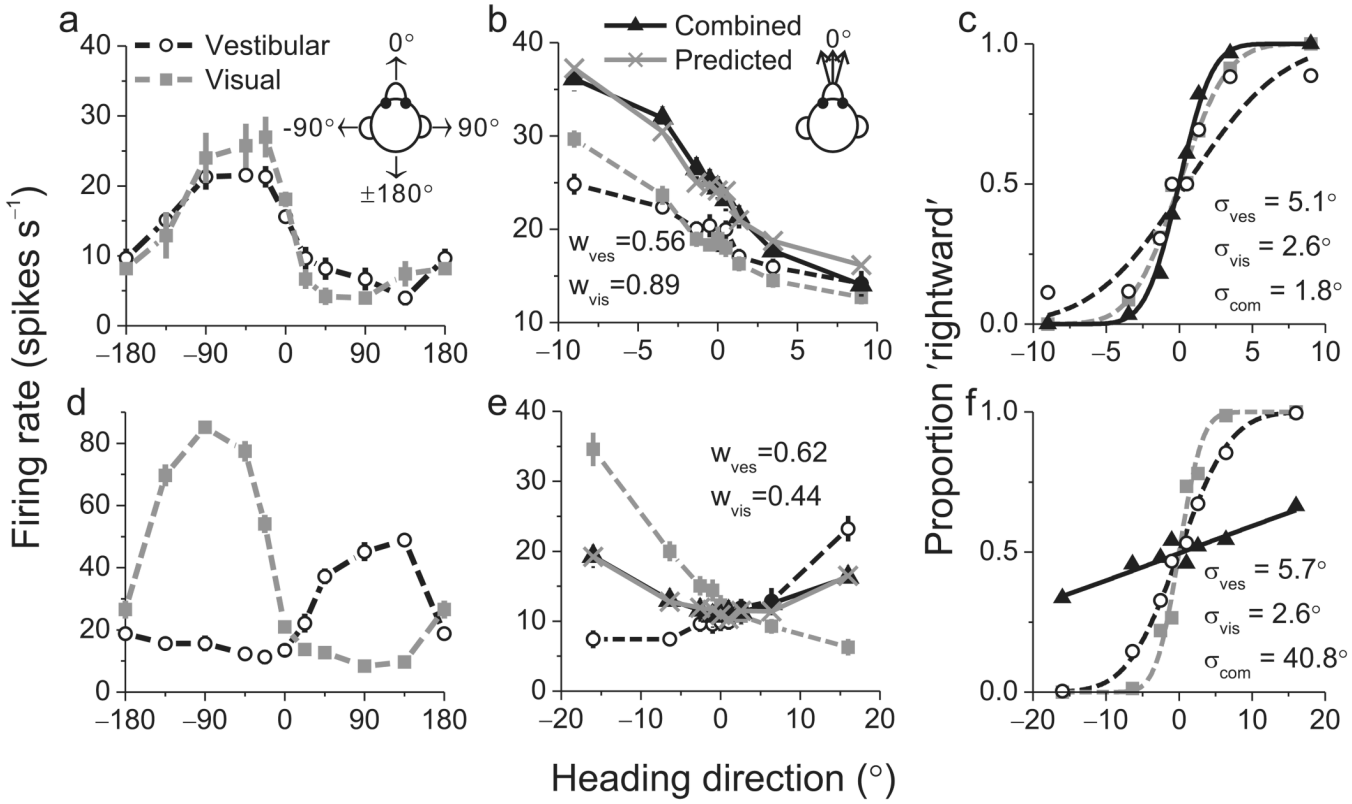

Figure 2.

Examples of neuronal tuning and neurometric functions for one congruent cell (a–c) and one opposite cell (d–f). (**a**) Heading tuning curves measured in the horizontal plane for the congruent cell. This neuron preferred leftward headings (near −90°) in both the vestibular (white circles) and visual (gray squares) conditions. (**b**) Responses of the same neuron to a narrow range of headings presented during the discrimination task. Responses in the combined condition (black triangles) were very similar to responses predicted from a weighted linear summation model (gray Xs). (**c**) Neurometric functions for the congruent cell computed by ROC analysis. Smooth curves show best-fitting cumulative Gaussian functions, with neuronal thresholds of 5.1°, 2.6° and 1.8° in the vestibular (white circles), visual (gray squares), and combined (black triangles) conditions, respectively. (**d**) Tuning curves for an opposite cell that preferred rightward motion in the vestibular condition and leftward motion in the visual condition. (**e**) Responses of the opposite cell during the discrimination task. Note that combined responses (black triangles) and predictions (gray Xs) fall between the single-cue responses. (**f**) Neurometric functions for the opposite cell, which has thresholds of 5.7°, 2.6° and 40.8° in the vestibular, visual, and combined conditions, respectively.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

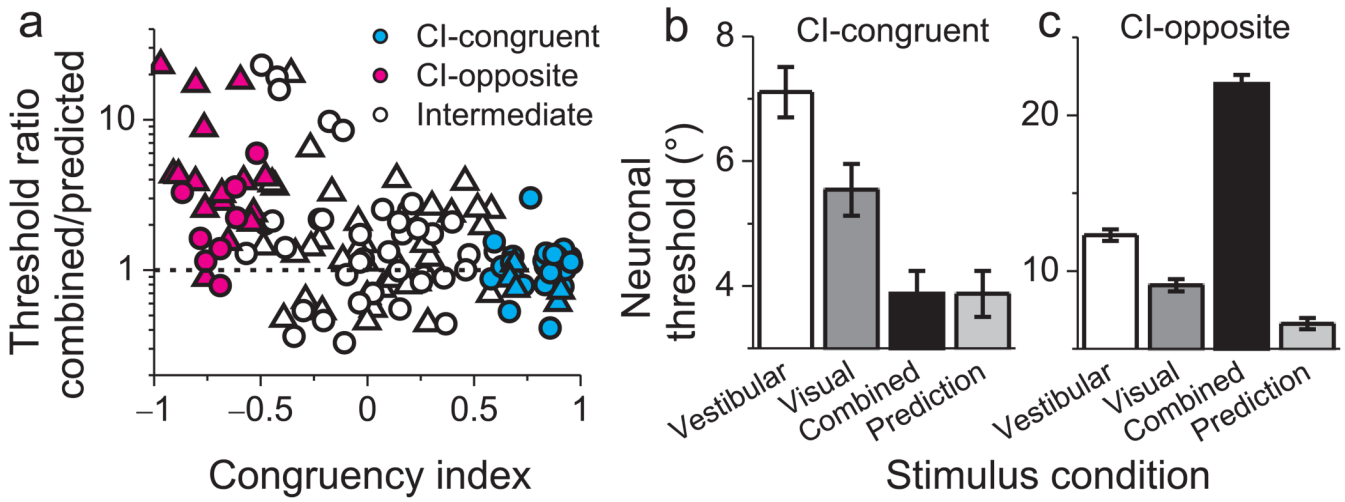

Figure 3.

Neuronal sensitivity under cue combination depends on visual/vestibular congruency. (**a**) The vertical axis in this scatter plot represents the ratio of the threshold measured in the combined condition to that predicted by optimal cue integration. The horizontal axis represents the Congruency Index of heading tuning for visual and vestibular responses (CI, see Methods). Filled symbols denote neurons for which CI is significantly different from zero. Cyan and magenta symbols represent ‘CI-congruent’ and ‘CI-opposite’ cells, respectively. Triangles and circles denote data from monkeys C and A, respectively. (**b**) Average neuronal thresholds (geometric mean ± geometric SE) for 'CI-congruent' cells (significant CI > 0, N = 30). Note that the average combined threshold (black bar) is very similar to the optimal prediction (light gray bar). (**c**) Average neuronal thresholds for 'CI-opposite' cells (significant CI < 0, N = 24), which become less sensitive under cue combination. Note that the vertical scale differs between panels b and c to clearly show the cue combination effect for each group of neurons.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

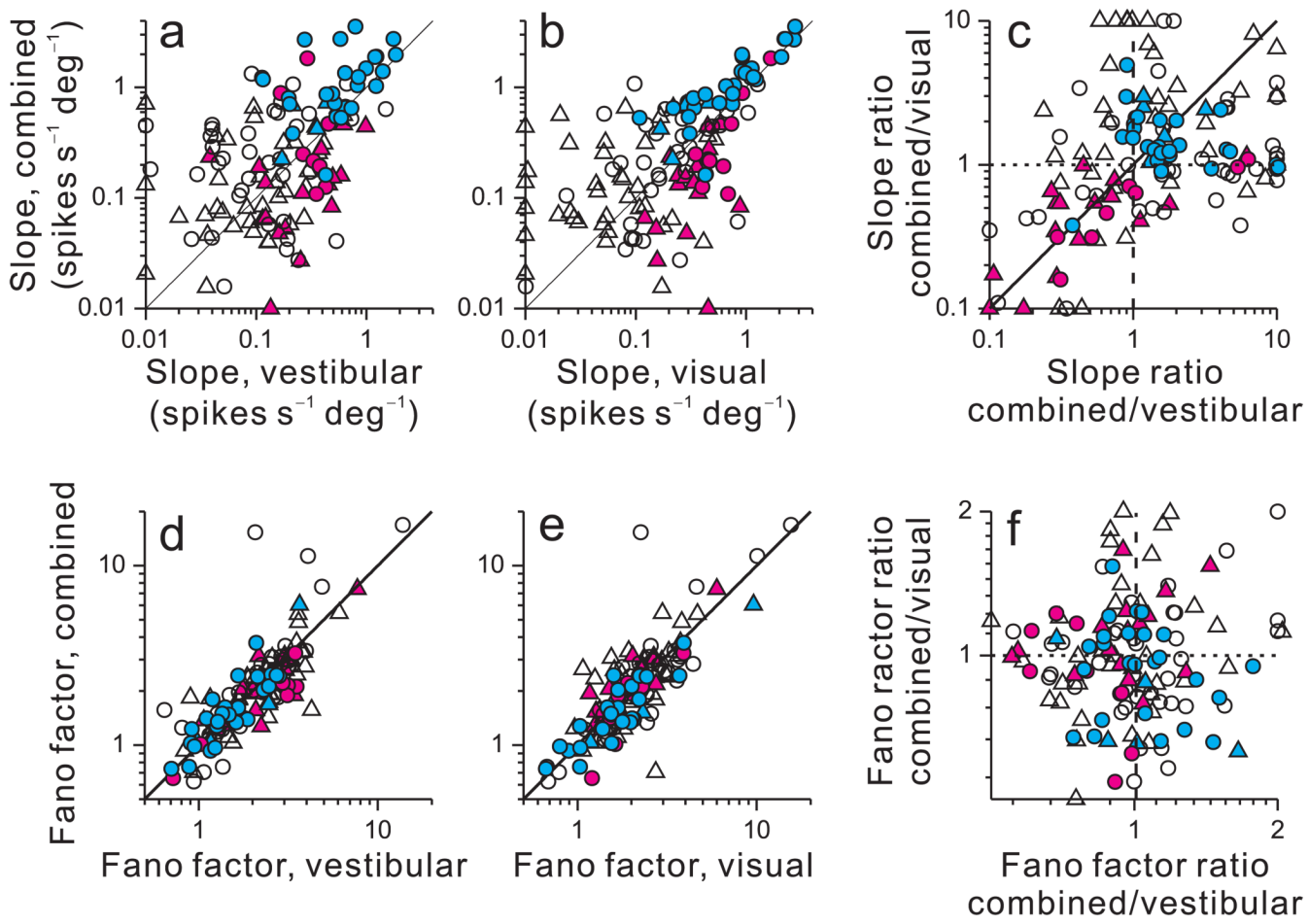

Figure 4.

Effect of cue integration on tuning curve slopes and Fano factors. (**a, b**) Tuning curve slope in the combined condition is plotted against slope in the vestibular (a) and visual (b) conditions. For CI-congruent cells (cyan), slopes were steeper during cue combination than in either single-cue condition ($p<0.001$, sign tests). For CI-opposite cells (magenta), slopes were flatter in the combined condition ($p<0.02$, sign tests). For intermediate cells (black), slopes were not significantly different ($p>0.1$, sign tests). Data points with slopes <0.01 spikes/s/deg were plotted at 0.01 for clarity. (**c**) The ratio of combined/visual slopes is plotted again the ratio of combined/vestibular slopes. Data points with values <0.1 (or >10) were plotted at 0.1 (or 10) for clarity. (**d, e, f**) Variance to mean ratios (Fano factors) are plotted in the same format as a–c. Fano factors were marginally smaller for CI-opposite cells in the combined condition as compared to the vestibular condition ($p=0.02$, sign test) but not significantly different as compared to the visual condition ($p>0.8$, sign test). There were no significant differences for all other comparisons ($p>0.2$, sign tests). Circles: monkey A; Triangles: monkey C.

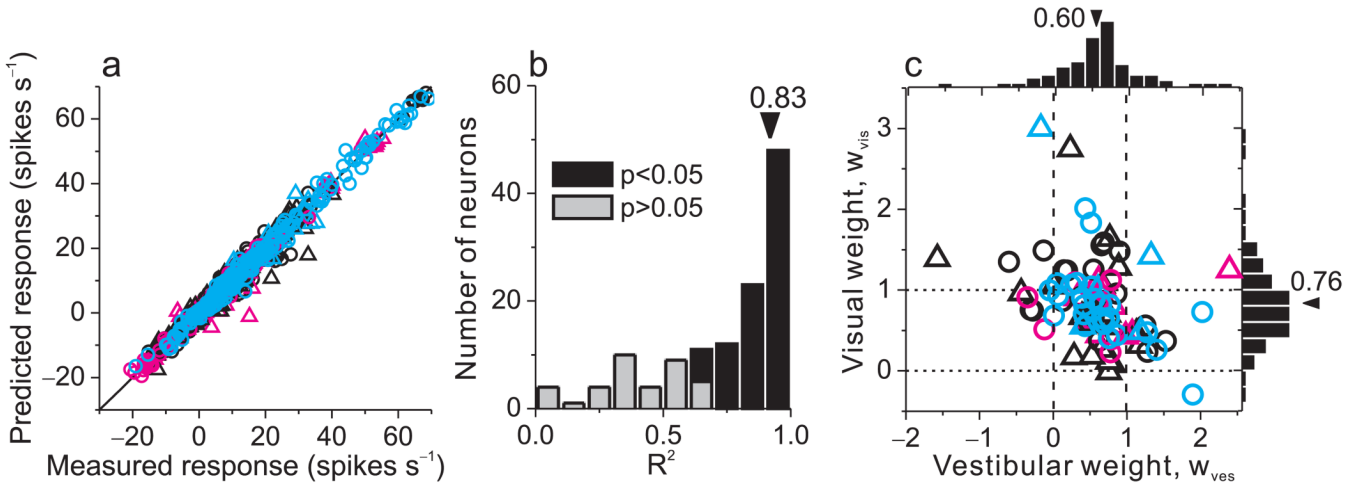

Figure 5.

Combined condition responses are well approximated by linear weighted summation. (**a**) Predicted responses from weighted linear summation were strongly correlated with measured responses in the combined condition (R=0.99, p<<0.001). Each datum represents the response of one neuron at one heading angle; spontaneous activity was subtracted. (**b**) A correlation coefficient was computed, for each neuron, from a linear regression fit to the predicted and measured responses. The median R^2 value was 0.83 and 89 cases (69%) were significant (p<0.05). Three cases with negative (but not significant) R^2 values are not shown. (**c**) Visual and vestibular weights derived from the best fit of the linear weighted sum model for each neuron with significant R^2 values (black bars in panel b). Median weights for vestibular and visual inputs were 0.6 and 0.76, respectively, which are significantly smaller than 1 (p<<0.001, t-tests). There is a significant negative correlation between vestibular and visual weights (R= −0.40, p<0.001, spearman rank correlation). Circles: monkey A; Triangles: monkey C.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

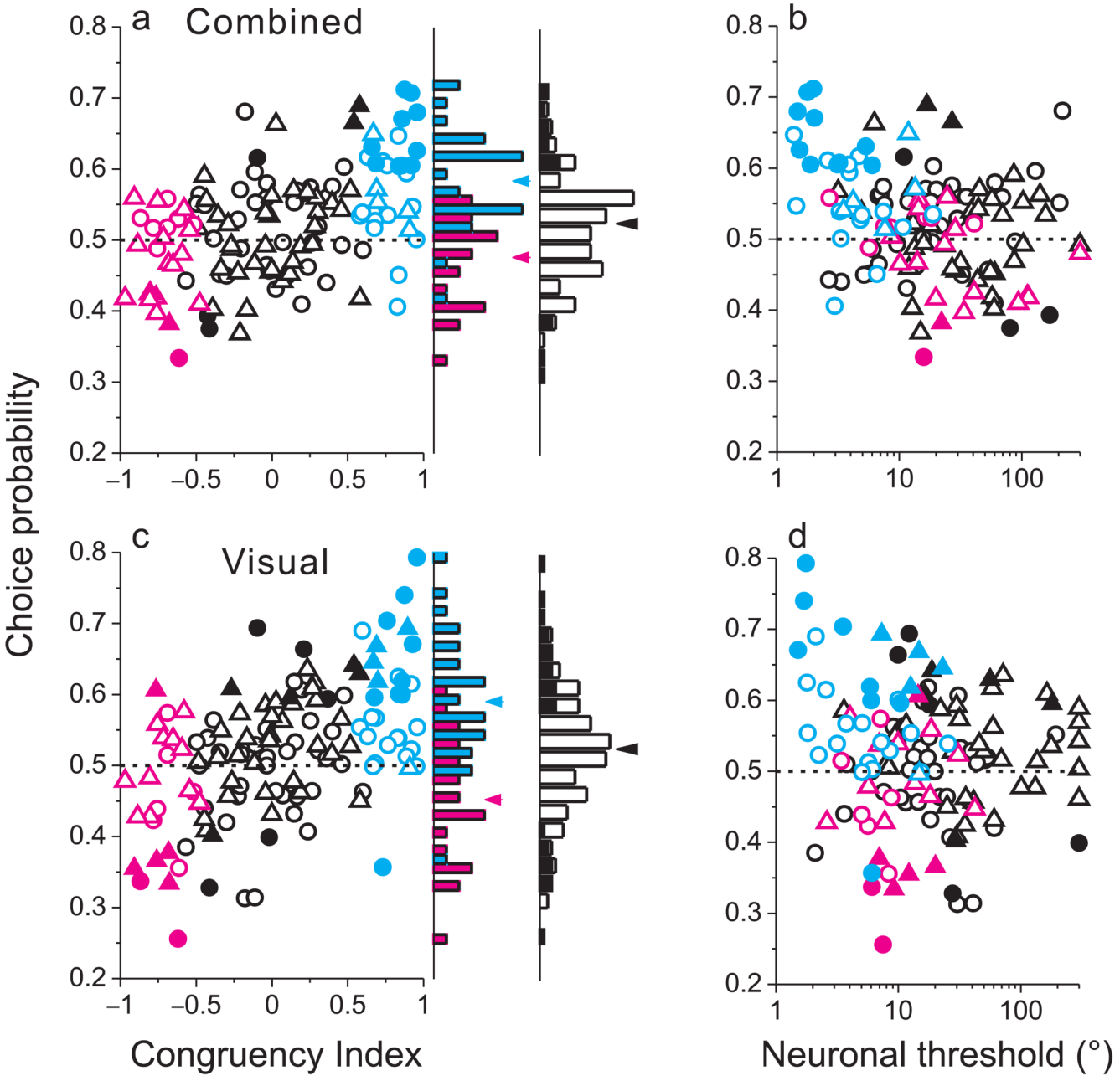

Figure 6.

Correlations between MSTd responses and perceptual decisions depend on congruency of tuning. **(a)** Choice probability (CP) is plotted against congruency index (CI) for all129 MSTd neurons tested during cue combination (triangles: monkey C; circles: monkey A). Cyan and magenta data points represent 'CI-congruent' and 'CI-opposite' cells, respectively. Filled symbols denote CPs significantly different from 0.5. The rightmost marginal histogram shows the distribution of CP values for all neurons, with filled bars denoting CPs significantly different from 0.5. The adjacent marginal histogram shows distributions of CP values for CI-congruent (cyan) and CI-opposite (magenta) cells. **(b)** CP is significantly anti-correlated with neuronal threshold during cue combination ($R = -0.31$, $p<0.0003$, Spearman rank correlation). **(c)** CPs in the visual condition, presented in the same format as panel a. CP was significantly

correlated with CI (R=0.51, p<<0.001). **(d)** Visual condition CPs plotted as a function of neuronal thresholds.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

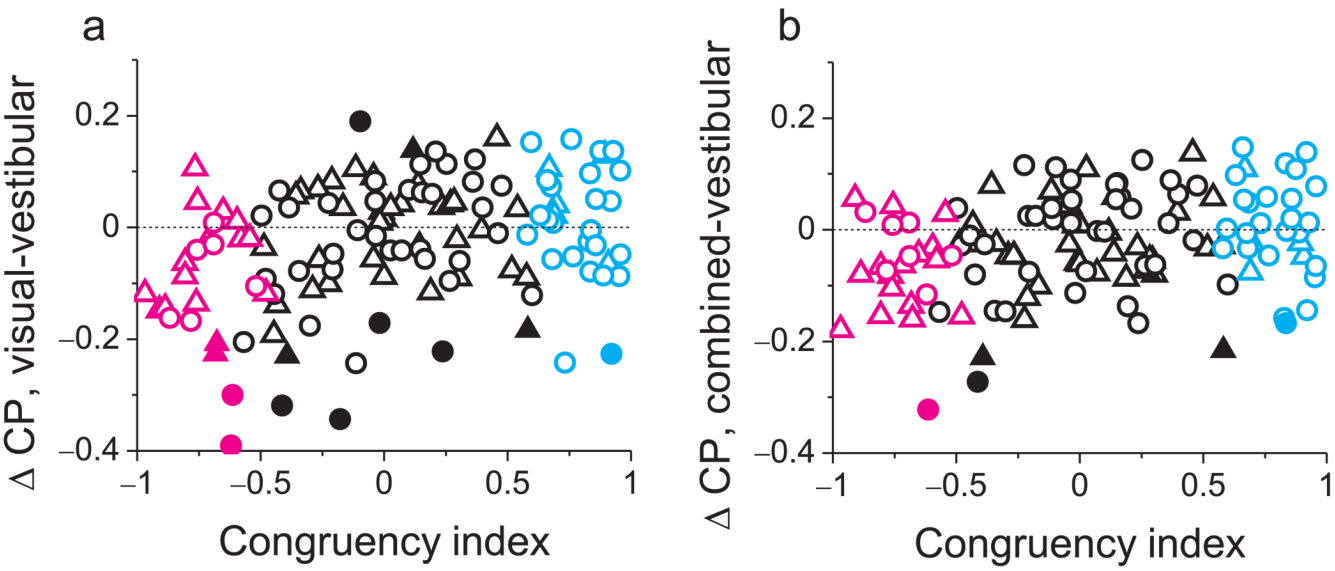

Figure 7.

Summary of effects of congruency on CP values across stimulus conditions. **(a)** The difference in CP between visual and vestibular conditions is plotted, for each neuron, against CI Filled symbols denote differences in CP that are significantly different from zero for individual neurons. Circles: monkey A; triangles: monkey C. Cyan and magenta data points represent 'CI-congruent' and 'CI-opposite' cells, respectively. **(b)** Differences in CP between the combined and vestibular conditions are plotted as a function of CI.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

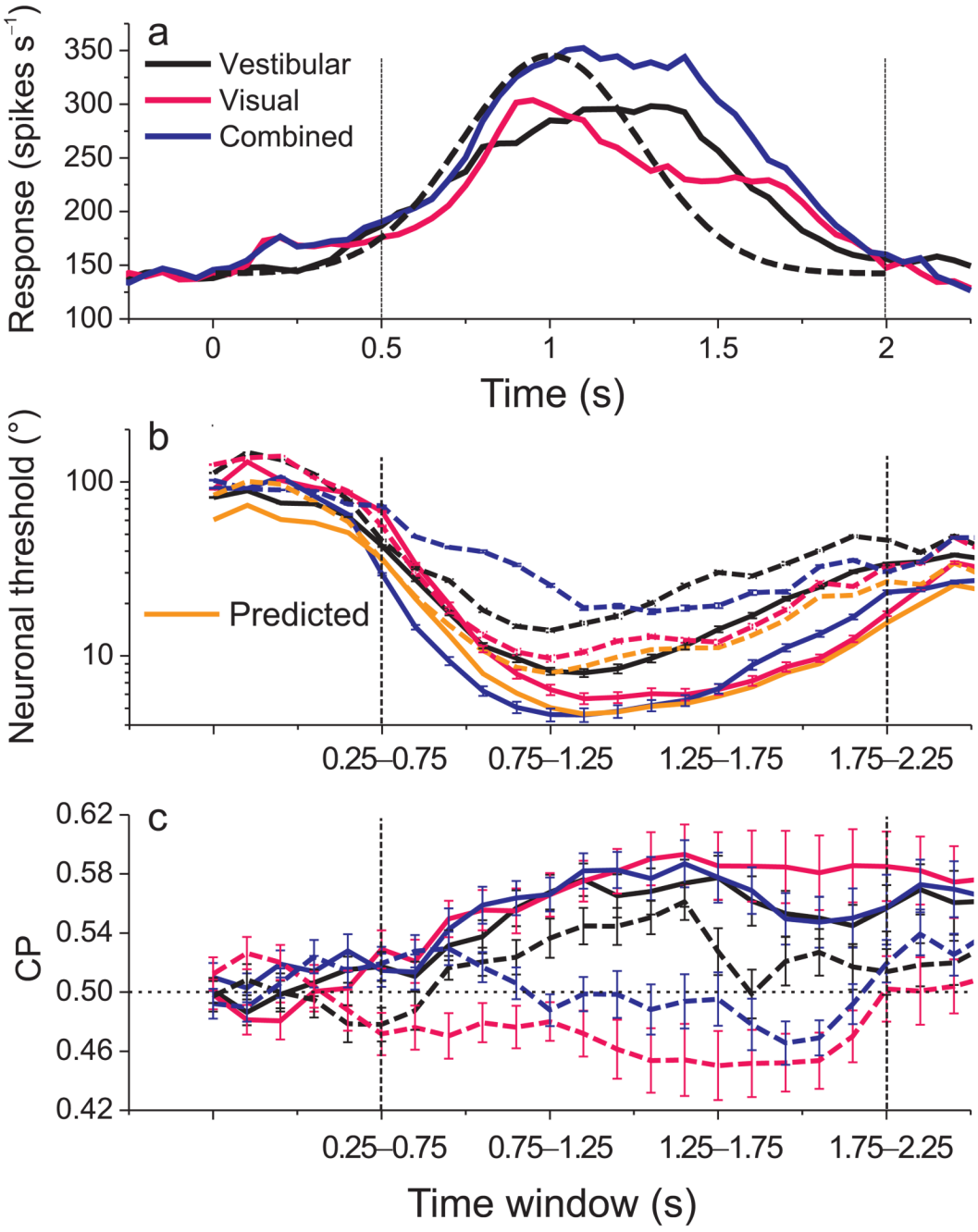

Figure 8.

Temporal evolution of population responses, neuronal thresholds, and CPs. (**a**) Average responses across all 129 neurons are shown for each stimulus condition (black: vestibular; pink: visual; blue: combined). For each neuron, responses were taken from the heading that elicited maximum firing rate. Dashed curve: velocity profile of stimulus. Time zero represents stimulus onset and 2s represents offset. Vertical dashed lines: time range for temporal analysis. (**b**) Average neuronal thresholds for CI-congruent (solid) and CI-opposite (dashed) cells as a function of time from 0.50 to 2s. Thresholds were computed from responses within a 500ms sliding time window that was advanced in steps of 100ms. The orange curves show the time course of predicted thresholds for the combined condition. Error bars: SEM. (**c**) Average CPs

for CI-congruent (solid) and CI-opposite (dashed) cells as a function of time, format as in panel b. Dashed horizontal line: chance level (CP=0.5).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript