

Neural correlates of multisensory cue integration in macaque MSTd

Yong Gu¹, Dora E Angelaki^{1,3} & Gregory C DeAngelis¹⁻³

Human observers combine multiple sensory cues synergistically to achieve greater perceptual sensitivity, but little is known about the underlying neuronal mechanisms. We recorded the activity of neurons in the dorsal medial superior temporal (MSTd) area during a task in which trained monkeys combined visual and vestibular cues near-optimally to discriminate heading. During bimodal stimulation, MSTd neurons combined visual and vestibular inputs linearly with subadditive weights. Neurons with congruent heading preferences for visual and vestibular stimuli showed improvements in sensitivity that parallel behavioral effects. In contrast, neurons with opposite preferences showed diminished sensitivity under cue combination. Responses of congruent cells were more strongly correlated with monkeys' perceptual decisions than were responses of opposite cells, suggesting that the monkey monitored the activity of congruent cells to a greater extent during cue integration. These findings show that perceptual cue integration occurs in nonhuman 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 perception¹, integration of stereo and texture cues for slant perception^{2,3}, visual-haptic integration^{4,5}, visual-auditory localization⁶ and object recognition⁷. Multisensory integration in human behavior often follows predictions of a quantitative framework that applies Bayesian statistical inference to the problem of cue integration⁸⁻¹⁰. 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 stimuli¹¹. Multimodal neurons have been described in several brain areas^{12,13}, but these studies have typically used anesthetized or passively viewing animals¹⁴⁻¹⁷. Multimodal neurons have not been studied during performance of multisensory 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 crucial 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 flow¹⁹⁻²² and inertial motion in darkness²³⁻²⁷. 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 theory²⁸? 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 and vestibular cue integration in macaque visual cortex.

RESULTS

Psychophysical performance

We trained two monkeys to perform a multimodal 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**). Three stimulus conditions were randomly interleaved: a vestibular condition in which heading was defined solely by inertial motion cues²⁹, a visual condition in which heading was defined solely by optic flow and a combined condition consisting of congruent inertial motion and optic flow.

¹Department of Anatomy and Neurobiology, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, Missouri 63110, USA. ²Department of Brain and Cognitive Sciences, Center for Visual Science, University of Rochester, 245 Meliora Hall, Rochester, New York 14627, USA. ³These authors contributed equally to this work. Correspondence should be addressed to G.C.D. (gdeangelis@cvs.rochester.edu).

Received 12 March; accepted 24 July; published online 7 September 2008; doi:10.1038/nn.2191

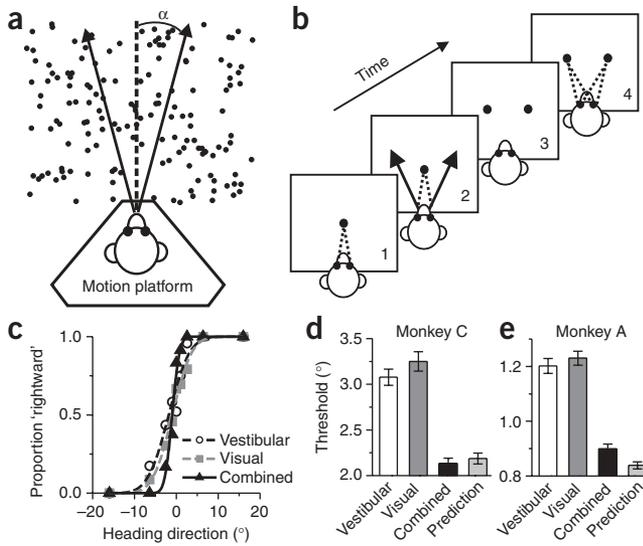


Figure 1 Heading task and behavioral performance. (a) Monkeys were seated on a motion platform and translated within the horizontal plane. A projector mounted on the platform displayed images of a three-dimensional star field, thus providing optic flow. (b) After fixating a visual target, the monkey experienced forward motion with a small leftward or rightward (arrow) component and subsequently reported his perceived heading ('left' versus '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. Smooth curves represent best-fitting cumulative Gaussian functions. (d,e) Average psychophysical thresholds from two monkeys (monkey C, $n = 57$; monkey A, $n = 72$) for the three stimulus conditions, and predicted thresholds computed from optimal cue integration theory. Error bars indicate s.e.m.

We quantified discrimination performance by constructing psychometric functions for each stimulus condition (Fig. 1c), and we estimated psychophysical thresholds from cumulative Gaussian fits³⁰ (see Methods). In this example, the monkey's heading threshold was between 3.5° and 4° for each single-cue condition (Fig. 1c). 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 combination⁵. In the combined condition, 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 monkeys are provided in Supplementary Fig. 1 online).

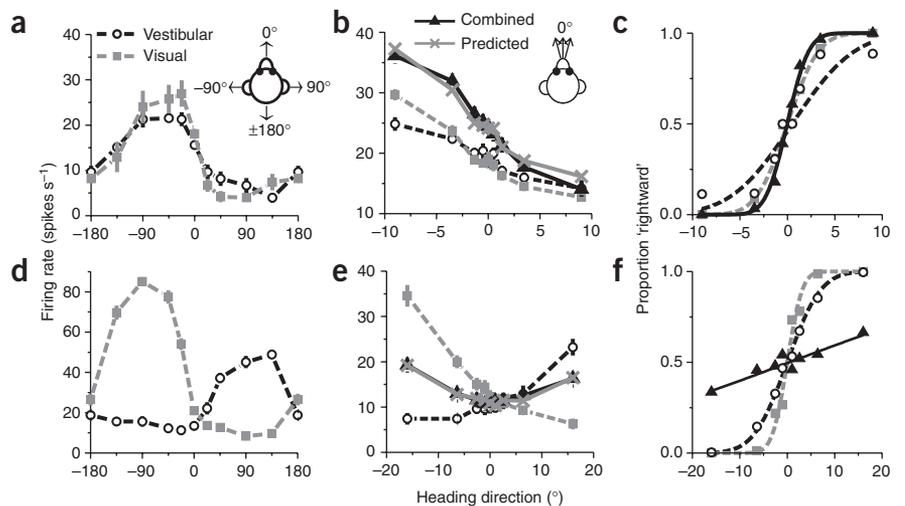
Across 57 sessions, monkey C had psychophysical thresholds of $3.1 \pm 0.09^\circ$ and $3.2 \pm 0.1^\circ$ (mean \pm s.e.m.) in the vestibular and visual conditions, respectively (Fig. 1d). If the monkey combined the two cues optimally⁵, thresholds should be reduced by ~30% (see Methods). The average threshold measured in the combined condition ($2.1 \pm 0.06^\circ$) did not differ significantly from the optimal prediction ($2.2 \pm 0.06^\circ$; $P = 0.4$,

paired t -test) computed from the single-cue data. Moreover, the combined threshold was significantly smaller than the single-cue thresholds ($P \ll 0.001$, paired t -tests), indicating better performance using both cues. For monkey A (Fig. 1e), the combined threshold ($0.9 \pm 0.02^\circ$) was again significantly smaller than the single-cue thresholds ($P \ll 0.001$, paired t -tests) and was close to the optimal prediction ($0.8 \pm 0.02^\circ$), although the difference was significant ($P = 0.004$, paired t -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 the activity of single neurons in area MSTd while monkeys performed the heading task. We previously reported that ~60% of MSTd neurons signal heading for both vestibular and optic flow stimuli²⁵. To identify these multimodal neurons, we measured heading tuning curves in the horizontal plane while animals maintained visual fixation (data from an example neuron with clear tuning under both single-cue conditions ($P \ll 0.001$, analysis of variance (ANOVA)) is shown; Fig. 2a). This neuron preferred leftward headings for both stimuli and was classified as a 'congruent' cell. Heading direction always refers to real or simulated body motion, such that congruent cells have visual and vestibular tuning curves with aligned peaks.

Figure 2 Examples of neuronal tuning and neurometric functions for one congruent cell and one opposite cell. (a) Heading tuning curves measured in the horizontal plane for the congruent cell. This neuron preferred leftward headings (near -90°) under both the vestibular and visual conditions. (b) Responses of the same neuron to a narrow range of headings presented during the discrimination task. Responses under the combined condition were very similar to responses predicted from a weighted linear summation model with weights of 0.56 and 0.89 applied to the vestibular and visual responses, respectively. (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° under the vestibular, visual and combined conditions, respectively. (d) Tuning curves for an opposite cell that preferred rightward motion under the vestibular condition and leftward motion under the visual condition. (e) Responses of the opposite cell during the discrimination task. Combined responses and predictions fell between the single-cue responses. Weights of 0.62 and 0.44 were applied to the vestibular and visual responses, respectively. (f) Neurometric functions for the opposite cell, which has thresholds of 5.7°, 2.6° and 40.8° under the vestibular, visual and combined conditions, respectively.



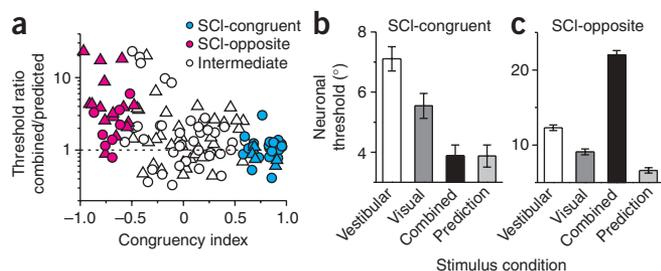


Figure 3 Neuronal sensitivity under cue combination depends on visual and vestibular congruency. **(a)** Vertical axis represents the ratio of the threshold measured under the combined condition to that predicted by optimal cue integration. Horizontal axis represents the congruency index of heading tuning for visual and vestibular responses. Filled symbols denote neurons for which congruency index was significantly different from 0. Triangles and circles denote data from monkeys C and A, respectively. **(b)** Average neuronal thresholds (geometric mean \pm geometric s.e.m.) for SCI-congruent cells (significant congruency index > 0 , $n = 30$). The average combined threshold was very similar to the optimal prediction. **(c)** Average neuronal thresholds for SCI-opposite cells (significant congruency index < 0 , $n = 24$), which became less sensitive under cue combination. The vertical scale differs between **b** and **c** to clearly show the cue-combination effect for each group of neurons.

Among 340 neurons tested, 194 (57%) showed significant heading tuning under both single-cue conditions ($P < 0.05$, ANOVA), and we studied these cells further. Over the narrow range of headings sampled during the discrimination task, the tuning of the example congruent neuron was monotonic under all three stimulus conditions (Fig. 2b). For this cell, the mean firing rate in the combined condition was greater than that under 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 under 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 'antireneur' (see Methods; Fig. 2c). We computed neuronal thresholds from the s.d. of the best-fitting cumulative Gaussian functions (Fig. 2c). For the example cell (Fig. 2a–c), the neuronal threshold was smaller under the combined condition (1.8°) than under the visual (2.6°) or vestibular (5.1°) conditions, indicating that the neuron could discriminate smaller variations in heading when both cues were provided.

We collected analogous data for an 'opposite' cell that preferred leftward headings during visual stimulation but rightward headings during vestibular stimulation (Fig. 2d–f). During the discrimination task, responses to combined stimulation were intermediate between responses to the visual and vestibular conditions (Fig. 2e). As a result, the cell's tuning curve was rather flat under 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 monkey (Supplementary Fig. 2 online). 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 neurons³³. Stimuli in our task were not tailored to the tuning of individual neurons, such that many neurons had large thresholds because their tuning curves were flat over the range of headings tested (Supplementary Fig. 3 online). Neuronal sensitivity was greatest for neurons with heading preferences $60\text{--}90^\circ$ away from straight forward, as these neurons had tuning curves with near-maximal slopes around the 0° heading (Supplementary Fig. 4 online).

The question of central interest is whether neuronal thresholds, like the behavioral thresholds, are significantly lower for the combined condition than for the single-cue conditions. Unlike the behavioral thresholds, 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 described above (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 (see Methods). The congruency index is 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 0 when either tuning function is flat (or evenly symmetric) over the range of headings tested. We plotted the ratio of combined to predicted thresholds against congruency index for all 129 neurons (Fig. 3a). A significant negative correlation was observed ($R = -0.45$, $P \ll 0.001$, Spearman rank correlation), such that neurons with large positive congruency indices had thresholds close to optimal predictions (ratios near 1). In contrast, neurons with large negative congruency indices generally had combined/predicted threshold ratios well in excess of 1. We defined neurons with congruency indices significantly larger than 0 ($P < 0.05$) as 'significant congruency index (SCI)-congruent' cells (Fig. 3a). Average neuronal thresholds for SCI-congruent cells followed a pattern similar to that of 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 congruency indices significantly smaller than 0 ($P < 0.05$) as 'SCI-opposite' cells (Fig. 3a). Combined thresholds for SCI-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 result from either 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 SCI-congruent cells: 0.80 (visual), 0.87 (vestibular) and 0.85 (combined); median R^2 values for SCI-opposite cells: 0.76 (visual), 0.67 (vestibular) and 0.35 (combined)). Nonlinear fits were also evaluated (see Supplementary Fig. 3). We plotted the slope of the tuning curve in the combined condition as a function of the slopes for the vestibular (Fig. 4a) and visual (Fig. 4b) conditions. In both scatter plots, data from SCI-congruent cells

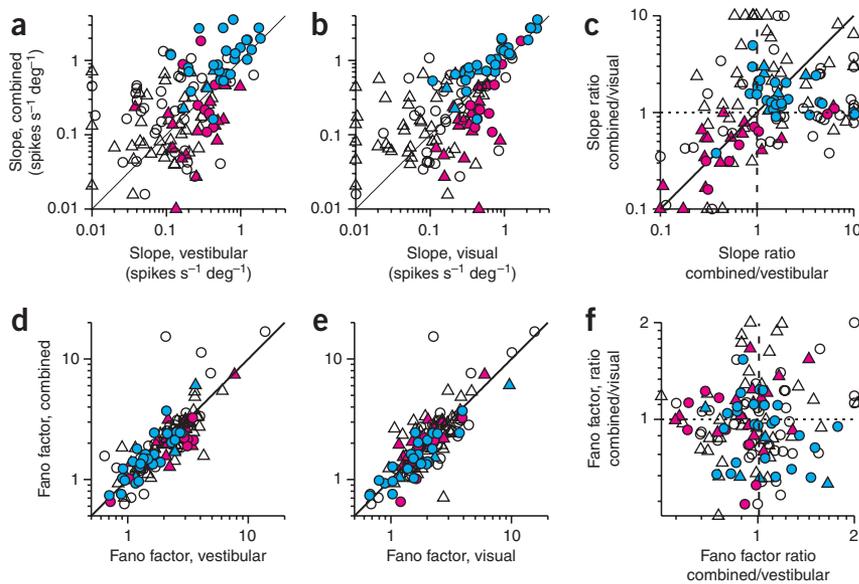


Figure 4 Effect of cue integration on tuning curve slopes and Fano factors. (**a,b**) Tuning curve slope under the combined condition is plotted against slope under the vestibular (**a**) and visual (**b**) conditions. For SCI-congruent cells, slopes were steeper during cue combination than under either single-cue condition ($P < 0.001$, sign tests). For SCI-opposite cells, slopes were flatter under the combined condition ($P < 0.02$, sign tests). For intermediate cells, slopes were not significantly different ($P > 0.1$, sign tests). Data points with slopes < 0.01 spikes per second per degree were plotted at 0.01 for clarity. Blue, SCI-congruent; red, SCI-opposite; white, intermediate. (**c**) The ratio of combined/visual slopes is plotted against the ratio of combined/visual slopes. Data points with values < 0.1 (or > 10) were plotted at 0.1 (or 10) for clarity. (**d–f**) Variance-to-mean ratios (Fano factors) are plotted as in **a–c**. Fano factors were marginally smaller for SCI-opposite cells under the combined condition compared to the vestibular condition ($P = 0.02$, sign test) but not significantly different compared to the visual condition ($P > 0.8$, sign test). There were no significant differences for any other comparisons ($P > 0.2$, sign tests). Circles, monkey A; triangles, monkey C.

generally fell above the unity-slope diagonal ($P < 0.001$, sign test), whereas data from SCI-opposite cells typically fell below the diagonal ($P < 0.02$, sign test). Thus, as shown in **Figure 2**, tuning slopes under the combined condition were generally steeper for congruent cells and flatter for opposite cells compared to single-cue conditions. This is underscored by comparing the ratios of combined to single-cue slopes (**Fig. 4c**). SCI-congruent cells tended to cluster in the upper right quadrant (both ratios > 1), whereas the majority of SCI-opposite cells were 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 and single-cue conditions (**Fig. 4d–f**). Thus, differences in neuronal sensitivity between congruent and opposite cells resulted mainly from differences in slope, with little contribution from changes in Fano factor. However, 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 response variance (**Supplementary Fig. 5** online). In general, response variance does have 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

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 stimuli²⁸. We tested whether a linear model with independent visual and vestibular weights could fit the bimodal responses of MSTd neurons (other variants of linear models with the same or fewer parameters are described in **Supplementary Methods** and **Supplementary Fig. 6** online; examples of model fits are shown in **Fig. 2b,e**). For both example neurons, the weighted linear model provided a good fit, and the weights on the visual and vestibular responses were less than 1 (additional examples of tuning curves, along with superimposed model fits, are shown in **Supplementary Fig. 7** online).

We examined the results of the linear fits (**Fig. 5**). Predicted responses from the weighted linear sum model were strongly correlated with responses measured in the combined condition ($R = 0.99$, $P < 0.001$; **Fig. 5a**), with a slope that was not significantly different from 1 (95% confidence interval = 0.988–1.003) and an offset that was not significantly different from 0 (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 (69%) of the 129 cells (**Fig. 5b**). The remaining 40 (31%) of the 129 cells did not show significant correlation coefficients, mainly because the combined responses were noisy or the tuning

curve was flat over the range of headings tested. Among neurons with R values less than 0.5, the average slope of the tuning curve was only 0.08 ± 0.02 spikes per second per degree (mean \pm s.e.m.; $n = 26$), and only 6 (23%) of 26 cells showed significant tuning under the combined condition ($P < 0.05$, ANOVA). For cells with R values greater than 0.5, the average tuning slope was 0.60 ± 0.06 spikes per second per degree ($n = 103$), and 92 (89%) of 103 cells showed significant tuning under the combined condition.

We determined the visual and vestibular weights derived from the linear model for a population of MSTd neurons (shown in **Fig. 5c**), and a few aspects of these data are notable. First, average weights were significantly less than 1 ($P < 0.001$, t -tests), with median vestibular and visual weights of 0.6 and 0.76, respectively. Thus, multisensory integration by MSTd neurons is typically subadditive during heading discrimination. Second, there was 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 tended to have small vestibular weights, and vice-versa. This suggests that neurons varied continuously along a range from visual to vestibular dominance under the conditions of our experiment. Third, visual and vestibular weights did not depend on congruency index ($P > 0.5$, Spearman rank correlation), indicating that the same linear weighting of visual and vestibular inputs is used by congruent and opposite cells.

Correlations between neural activity and behavior

If monkeys rely on area MSTd for multisensory integration in the heading discrimination task, one should expect to find significant correlations between neuronal activity and behavior. To test this

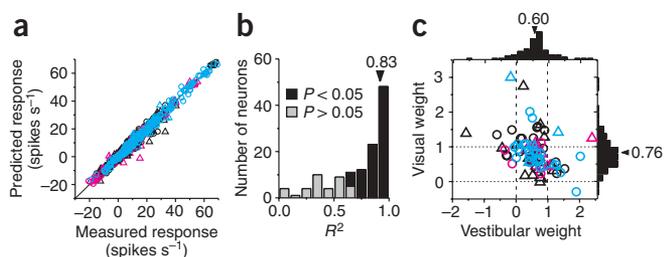


Figure 5 Combined-condition responses are well approximated by linear weighted summation. **(a)** Predicted responses from weighted linear summation were strongly correlated with measured responses under the combined condition ($R = 0.99$, $P \ll 0.001$). Each symbol represents the response of one neuron at one heading angle; spontaneous activity was subtracted. Blue, SCI-congruent; red, SCI-opposite; white, intermediate. **(b)** A correlation coefficient was computed, for each neuron, from a linear regression fit to the predicted and measured responses. The median R value was 0.83, and 89 cases (69%) were significant ($P < 0.05$). Three cases with negative (but not significant) R 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 values (black bars in **b**). Median weights for vestibular and visual inputs were 0.6 and 0.76, respectively, which are significantly smaller than 1 ($P \ll 0.001$, t -tests). There was a significant negative correlation between vestibular and visual weights ($R = -0.40$, $P < 0.001$, Spearman rank correlation). Circles, monkey A; triangles, monkey C.

hypothesis, we computed choice probabilities³⁴ 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 choice probability greater than 0.5 indicates that the monkey tends to choose the neuron's preferred sign of heading (leftward versus rightward) when the neuron fires more strongly²⁹. This result is thought to reflect a functional link between the neuron and perception^{33,35}. Of crucial interest is whether MSTd neurons show robust choice probabilities under 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 choice probability was modestly larger than chance (mean choice probability = 0.52, $P = 0.003$). Despite this significant bias in the expected direction, roughly equal numbers of neurons had individually significant choice probabilities that were greater or less than 0.5 (Fig. 6a). Neurons with a significant choice probability less than 0.5 paradoxically increased their firing rates when the monkey chose their nonpreferred direction.

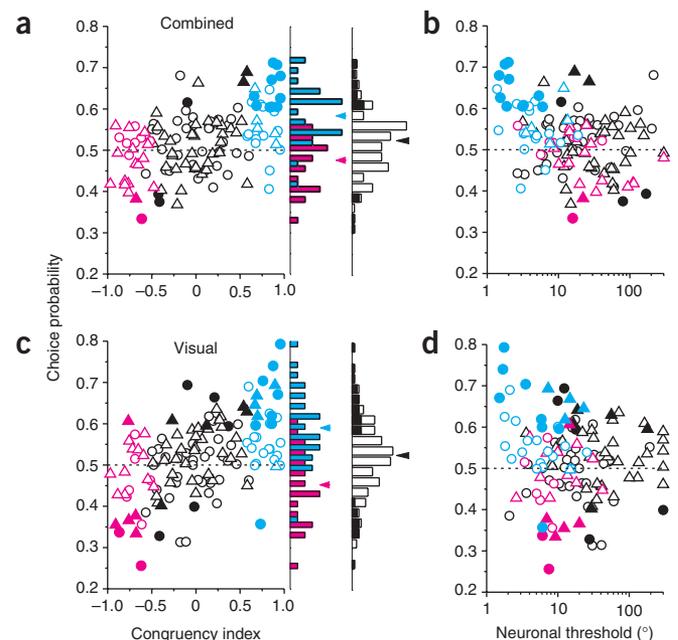
This diversity of choice probabilities under the combined condition is related to visual-vestibular congruency, as evidenced by a highly

Figure 6 Correlations between MSTd responses and perceptual decisions depend on congruency of tuning. **(a)** Choice probability is plotted against congruency index for all 129 MSTd neurons tested during cue combination (triangles, monkey C; circles, monkey A). Filled symbols denote choice probabilities significantly different from 0.5. The right histogram shows the distribution of choice probability values for all neurons, with filled bars denoting choice probabilities significantly different from 0.5. The left marginal histogram shows distributions of choice probability values for SCI-congruent and SCI-opposite cells. Blue, SCI-congruent; red, SCI-opposite; white, intermediate. **(b)** Choice probability was significantly anticorrelated with neuronal threshold during cue combination ($R = -0.31$, $P < 0.0003$, Spearman rank correlation). **(c)** Choice probabilities under the visual condition, presented in the same format as in **a**. Choice probability was significantly correlated with congruency index ($R = 0.51$, $P \ll 0.001$). **(d)** Visual-condition choice probabilities plotted as a function of neuronal thresholds.

significant positive correlation between choice probability and congruency index ($R = 0.44$, $P \ll 0.001$; Fig. 6a). The average choice probability for SCI-congruent cells was significantly greater than 0.5 (mean = 0.58, $P \ll 0.001$, t -test), suggesting that these cells were strongly coupled to perceptual decisions during cue combination. For SCI-opposite cells, however, the average choice probability was close to, and even slightly less than, 0.5 (mean = 0.48, $P = 0.08$). As a result, the average choice probability under the combined condition was significantly greater for SCI-congruent than for SCI-opposite cells ($P \ll 0.001$, t -test). This result suggests that the monkeys monitored 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 choice probabilities reflect a functional linkage between neurons and perception, then more sensitive neurons may show larger choice probabilities. This relationship might arise because the animal selectively monitors the most informative neurons³⁶ and/or because sensitive neurons are more strongly correlated with each other³⁷. Indeed, choice probabilities were negatively correlated with neuronal thresholds under the combined condition ($R = -0.31$, $P < 0.0003$, Spearman rank correlation; Fig. 6b). By virtue of their steeper tuning slopes under cue combination, SCI-congruent cells tended to have low neuronal thresholds and high choice probabilities (data from these neurons clustered in the upper left corner of Fig. 6b). In contrast, SCI-opposite cells tended to have high neuronal thresholds and low choice probabilities. Together, the data from SCI-congruent and SCI-opposite cells (Fig. 6b) formed a cloud of points that showed a robust negative correlation, suggesting that the difference in choice probability between congruent and opposite cells is driven largely by the difference in sensitivity.

We previously reported significant choice probabilities for MSTd neurons under the vestibular condition²⁹ (mean = 0.55, $P \ll 0.001$). These choice probabilities depended modestly on heading preference, such that neurons with heading preferences 30–90° away from straight ahead tended to have larger choice probabilities (Supplementary Fig. 4). Unlike in the combined condition, the vast majority of neurons (23 (96%) of 24) with significant vestibular choice probabilities had values larger than 0.5, and vestibular choice probabilities were similar



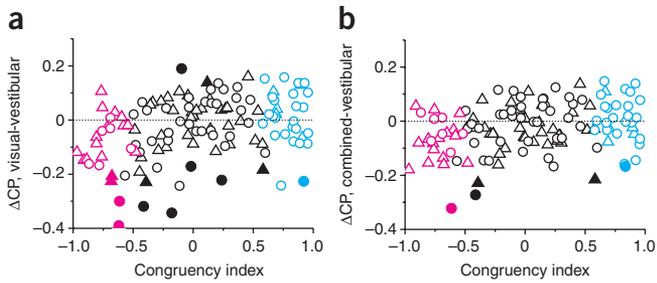


Figure 7 Summary of effects of congruency on choice probability values across stimulus conditions. **(a)** The difference in choice probability (ΔCP) between visual and vestibular conditions is plotted, for each neuron, against congruency index. Filled symbols denote differences in choice probability that are significantly different from 0 for individual neurons. Circles, monkey A; triangles, monkey C. Blue, SCI-congruent; red, SCI-opposite; white, intermediate. **(b)** Differences in choice probability between the combined and vestibular conditions are plotted as a function of congruency index.

for SCI-congruent and SCI-opposite cells ($P = 0.06$; **Supplementary Fig. 8** online). Both cell types showed mean vestibular choice probabilities that were significantly greater than 0.5 ($P \ll 0.001$, t -tests). Thus, the strong dependence of combined choice probabilities on congruency cannot be explained by a similar dependence under the vestibular condition.

In contrast, choice probabilities under the visual condition depended strongly on congruency. Across all 129 neurons, the average choice probability was significantly lower under the visual condition (mean = 0.52, $P = 0.01$) than under the vestibular condition ($P < 0.008$, paired t -tests). Moreover, roughly equal numbers of neurons had visual choice probabilities significantly greater and less than 0.5 (**Fig. 6c**). This finding was linked to a highly significant correlation between choice probability and congruency index under the visual condition ($R = 0.51$, $P \ll 0.001$; **Fig. 6c**). The average choice probability for SCI-congruent cells was substantially greater than 0.5 (0.59, $P \ll 0.001$), whereas SCI-opposite cells had an average visual choice probability significantly less than 0.5 (mean = 0.45, $P = 0.02$). Unlike vestibular and combined responses, there was no significant correlation between choice probabilities and neuronal thresholds under the visual condition ($R = -0.1$, $P > 0.2$; **Fig. 6d**). SCI-congruent and SCI-opposite cells seemed to fall along the same trend under the combined condition (**Fig. 6b**), but this was not the case for the visual condition (**Fig. 6d**). Among the most sensitive neurons, SCI-opposite cells tended to have choice probabilities less than 0.5, whereas SCI-congruent cells had choice probabilities consistently greater than 0.5.

We determined the differences in choice probabilities between stimulus conditions on a cell-by-cell basis (**Fig. 7**). The difference in

choice probability between visual and vestibular conditions was significantly correlated with congruency index ($R = 0.33$, $P \ll 0.001$, Spearman rank correlation; **Fig. 7a**). A similar pattern of results held for the difference in choice probability between combined and vestibular conditions ($R = 0.26$, $P < 0.004$; **Fig. 7b**).

In summary, we found 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. Notably, choice probabilities were computed by identifying a neuron's preferred sign of heading (leftward or rightward) independently for each stimulus condition. Thus, under the visual condition, opposite cells with choice probability less than 0.5 tended to fire more strongly when the monkeys reported their nonpreferred sign of heading. Of course, for these neurons, their nonpreferred sign of heading under the visual condition was their preferred sign of heading under the vestibular condition. This suggests that, under the visual condition, responses of opposite cells are decoded relative to their vestibular preference. Most important, under the combined condition, congruent cells were more strongly correlated with monkeys' heading judgments than were opposite cells (see Discussion).

Temporal evolution of thresholds and choice probabilities

Population responses under all three stimulus conditions roughly followed 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 1 s of the 2-s stimulus period (containing most of the velocity variation). We also examined how neuronal thresholds and choice probabilities evolved as a function of time, focusing on the latter 1.5 s of the stimulus period, when the neuronal responses were robust. Within this time range, we repeated threshold and choice probability analyses for several 500-ms 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 SCI-congruent cells, thresholds measured under the combined condition were close to those predicted by theory, with the closest agreement

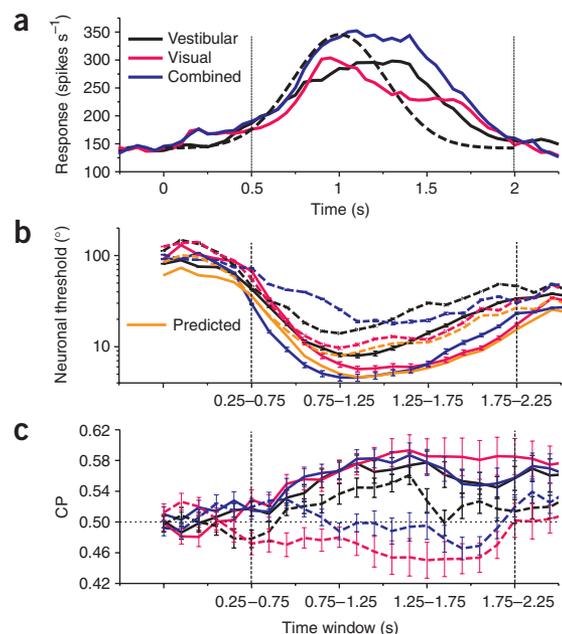


Figure 8 Temporal evolution of population responses, neuronal thresholds and choice probabilities. **(a)** Average responses across all 129 neurons are shown for each stimulus condition. For each neuron, responses were taken from the heading that elicited maximum firing rate. Dashed curve indicates velocity profile of stimulus. Time 0 represents stimulus onset; 2 s represents offset. Vertical lines indicate time range for temporal analysis. **(b)** Average neuronal thresholds for SCI-congruent (solid lines) and SCI-opposite (dashed lines) cells as a function of time from 0.50 to 2 s. Thresholds were computed from responses within a 500-ms sliding time window that was advanced in steps of 100 ms. Time courses of predicted thresholds for the combined condition are also shown. **(c)** Average choice probabilities (CP) for SCI-congruent (solid lines) and SCI-opposite (dashed lines) cells as a function of time, as in **b**. Dotted horizontal line indicates chance level (choice probability = 0.5). Error bars indicate s.e.m.

seen during the period of peak discharge (**Fig. 8b**). For SCI-opposite cells, combined thresholds were consistently higher than predictions across time. With regard to choice probabilities, SCI-congruent cells consistently showed values significantly larger than 0.5 ($P < 0.008$, t -tests) for all time windows (**Fig. 8c**). For SCI-opposite cells, vestibular choice probabilities were consistently larger than 0.5 in most time windows, whereas visual choice probabilities were consistently smaller than 0.5 across time, and combined choice probabilities hovered around 0.5. Thus, our main findings generalized to all time periods during which neural responses were robust. Significant choice probabilities developed early during the response, at least for congruent cells, and remained 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 was close to that predicted by optimal cue integration theory, as seen in human studies of multisensory integration^{5,6}. This suggests that humans and nonhuman primates use similar strategies for multisensory integration and establishes a model system to explore the neural mechanisms underlying perceptual cue integration. By recording the activity of single neurons in area MSTd, we identified a population of congruent cells that also showed improved sensitivity during cue combination and could account for the improvement in behavioral performance. In contrast, opposite cells showed reduced sensitivity during cue combination. These findings suggest that behavior relies more heavily on congruent cells than on opposite cells when both cues are provided. Consistent with this idea, congruent cells were more strongly correlated with perceptual decisions during cue integration than were opposite cells (**Figs. 6a** and **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 found that congruent cells showed increased sensitivity during cue combination, whereas opposite cells showed reduced sensitivity. 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 clearly show that slopes were increased under the combined condition for congruent cells and reduced for opposite cells (**Fig. 4c**). This slope increase for congruent cells might have arisen simply because responses were larger under the combined condition than under the single-cue conditions. This seems to be the case in **Figure 2b**, although it was not true generally. For other congruent cells (for example, in **Supplementary Fig. 7**), tuning curve slope was greater under the combined condition despite the fact that combined responses were lower than those elicited by the most effective single cue. Across the population of SCI-congruent cells, there was no significant correlation ($R = 0.15$, $P > 0.09$, Spearman rank correlation) between the slope increase under the combined condition and the ratio of firing rates between the combined and single-cue conditions (**Supplementary Fig. 9** online). Furthermore, the ratio of combined/single-cue firing rates was similar for SCI-opposite cells, yet slopes under the combined condition were consistently reduced for these cells. Thus, improved sensitivity of MSTd neurons under cue combination is not simply a result of greater responses.

A reduction in response variance could also have contributed to the increased sensitivity of congruent cells. However, Fano factor did not differ substantially between congruent and opposite cells (**Fig. 4d-f**).

Hence, the difference in sensitivity between congruent and opposite cells arose primarily from differences in the slope of the response function, and these slope changes were well accounted for by linear weighted summation of single-cue responses. These findings are consistent with predictions of a recent theoretical study²⁸ that proposed 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. However, 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 in our study showed comparable linear summation (**Fig. 5c**), but this produced greater sensitivity for congruent cells and poorer sensitivity for opposite cells. Thus, although neural cue combination may be well described by linear weighted summation, selective decoding and/or correlation of neuronal responses seems necessary to predict the behavioral effects from MSTd responses.

Choice probabilities

We previously reported significant choice probabilities for MSTd neurons under the vestibular condition of the heading discrimination task²⁹, suggesting that vestibular signals in MSTd contribute to perceptual decisions regarding heading. Our current findings show that this relationship held for both congruent and opposite cells (**Supplementary Fig. 8**), indicating that vestibular signals in MSTd are consistently correlated with perceptual decisions independent of the congruency with visual selectivity.

In contrast, choice probabilities in the visual condition depended strongly on congruency. For SCI-congruent cells, the average visual choice probability (0.59) was substantially greater than chance (0.5). When choice probabilities 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, SCI-opposite cells in our study had an average visual choice probability (0.45) significantly less than 0.5, indicating that they tended to fire more strongly when the monkey reported their nonpreferred sign of heading. This suggests that both SCI-congruent and SCI-opposite cells contribute to purely visual judgments of heading, but their activities bear different relationships to behavior.

This difference between SCI-congruent and SCI-opposite cells under the visual condition could be explained by a representation of heading in area MSTd in the form of a place code (or labeled-line code); the activity level of each neuron might then 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. 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 choice probabilities less than 0.5 for SCI-opposite cells. We speculate that the activity of opposite cells might be decoded as evidence in favor of their vestibular, not visual, preference as a result of our protocol for training monkeys to perform the heading task. We initially trained monkeys to discriminate heading solely on the basis of vestibular cues, and we subsequently added the random-dot stimuli and gradually increased motion coherence until thresholds under the combined condition were reduced relative to the vestibular condition. Only then did we introduce visual condition trials. Thus, the monkeys

may have learned to decode MSTd 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 choice probability results is that SCI-opposite cells are selectively correlated with SCI-congruent cells that have the same vestibular preference. Trial-to-trial correlations in responses among neurons are thought to be important—and may be essential—for observing choice probabilities³⁷, such that differences in choice probability across stimulus conditions may be caused by variations in correlated noise among the neurons. If SCI-opposite cells are correlated with SCI-congruent cells that have matching vestibular preferences, this could also explain why SCI-opposite cells show choice probability values below 0.5 under the visual condition.

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

An alternative possibility is that SCI-congruent cells exhibit greater interneuronal correlations than do SCI-opposite cells under the combined condition, leading to larger choice probabilities for SCI-congruent cells. We examined this possibility by computing noise correlations between single- and multi-unit activity recorded from the same electrode, as these noise correlations have been reported to be predictive of significant choice probabilities³⁹. We found robust noise correlations between single- and multi-unit responses but did not observe any dependence of these correlations on congruency index for any of the stimulus conditions (Supplementary Fig. 10 online). These data do not exclude a contribution of noise correlations to our findings, but they support the possibility that variations in choice probability are linked to selective decoding of MSTd neurons³⁶. It has not been shown directly, however, that selective decoding can lead to variations in choice probability without corresponding changes in the structure of noise correlations. Indeed, there is evidence to the contrary based on simulations of how populations of middle temporal neurons contribute to motion discrimination³⁷.

SCI-opposite neurons do not contribute to improved sensitivity during cue combination, but these cells may have important roles in parsing retinal image motion into components related to self-motion and object motion. Specifically, differences in activity between populations of SCI-congruent and SCI-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 summary, our findings implicate area MSTd in sensory integration for heading perception and establish a 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 (for example, Bayesian) inference.

METHODS

Motion stimuli. Two rhesus monkeys (*Macaca mulatta*) weighing ~6 kg each were trained using a virtual reality system²⁵. Translation of the monkey in the horizontal plane was accomplished by a motion platform (MOOG 6DOF2000E, Moog). To activate vestibular otolith organs, each inertial motion stimulus followed a smooth trajectory with a Gaussian velocity profile and a peak acceleration of $\sim 1 \text{ m s}^{-2}$.

A projector (Mirage 2000, Christie Digital) was mounted on the motion platform and rear-projected images ($90 \times 90^\circ$ of visual angle) onto a tangent screen. Visual stimuli depicted movement through a three-dimensional 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 $0.15 \text{ cm} \times 0.15 \text{ cm}$ triangle. Stimuli were presented stereoscopically as red and green anaglyphs, viewed through Wratten filters (red no. 29, green no. 61, Kodak). The stimuli contained a variety of depth cues, including horizontal disparity, motion parallax and size information. Motion coherence was manipulated by randomizing the three-dimensional location of a percentage of stars on each display update while the remaining stars moved according to the specified heading. This manipulation degraded optic flow as a heading cue and was used to reduce psychophysical sensitivity under the visual condition such that it matched vestibular sensitivity.

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 $2^\circ \times 2^\circ$ electronic window around the fixation point. At the end of the 2-s 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 paradigm²⁹. Nine logarithmically spaced heading angles were tested for each monkey, including an ambiguous straight-forward direction (monkey A: $\pm 9^\circ, \pm 3.5^\circ, \pm 1.3^\circ, \pm 0.5^\circ$ and 0° ; monkey C: $\pm 16^\circ, \pm 6.4^\circ, \pm 2.5^\circ, \pm 1^\circ$ and 0°). These values were carefully chosen 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. Under 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 under this condition depends heavily on vestibular signals²⁹. Under the visual condition, the motion platform remained stationary while optic flow simulated the same range of headings. Under the combined condition, congruent inertial motion and optic flow were provided²⁵. Each of the 27 unique stimulus conditions (9 headings \times 3 cue conditions) was typically repeated ~ 30 times, for a total of ~ 800 discrimination trials per recording session.

Electrophysiology. Extracellular single-unit recording was carried out as described previously^{25,29}. Area MSTd was located by structural magnetic resonance imaging and mapping of physiological response properties, such as direction selectivity for visual motion and visual receptive fields encompassing a large proportion of the contralateral 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° , $\pm 22.5^\circ$, $\pm 45^\circ$, $\pm 90^\circ$, $\pm 135^\circ$ 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 four or five repetitions were collected for each stimulus. Only MSTd neurons with significant tuning under 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 Gaussian³⁰. The psychophysical threshold for each stimulus condition was taken as the s.d. parameter of the Gaussian fit.

Predicted thresholds for the combined condition, assuming optimal (maximum likelihood) cue integration, were computed as⁵

$$\sigma_{\text{prediction}} = \sqrt{\frac{\sigma_{\text{vestibular}}^2 \times \sigma_{\text{visual}}^2}{\sigma_{\text{vestibular}}^2 + \sigma_{\text{visual}}^2}}$$

where $\sigma_{\text{vestibular}}$ and σ_{visual} represent psychophysical thresholds under 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 opposite-directed headings (for example, $+1^\circ$ versus -1°) based solely on the firing rate of the recorded neuron and a presumed antineuron with opposite tuning^{29,31}. Neurometric functions were constructed from these ROC values and fitted 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 monkey's choice at the end of each trial (that is, 'preferred' versus 'null' choices). ROC values were calculated from these two distributions whenever there were at least three choices in each group, and this yielded a choice probability for each heading direction. We combined data across headings (after z-score normalization) to compute a grand choice probability for each cue condition²⁹. The statistical significance of choice probabilities (relative to the chance level of 0.5) was determined using permutation tests (1,000 permutations).

For opposite cells, the definition of preferred and null choices is different for the vestibular and visual conditions. In computing choice probabilities, we defined preferred and null choices according to the tuning of the neuron under each particular stimulus condition. Thus, if the opposite neuron in Figure 2d consistently responded more strongly when the monkey reported rightward movement, it had a choice probability greater than 0.5 for the vestibular condition and a choice probability less than 0.5 for the visual condition.

To quantify the congruency between visual and vestibular tuning functions measured during discrimination, we calculated a congruency index. 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. Congruency index (CI) was defined as the product of these two correlation coefficients:

$$\text{Congruency index} = R_{\text{vestibular}} \times R_{\text{visual}}$$

Congruency index ranges from -1 to 1 , with values near 1 indicating that visual and vestibular tuning functions have a consistent slope (Fig. 2b) and values near -1 indicating opposite slopes (Fig. 2d). Congruency index reflects both the congruency of tuning and the steepness of the slopes of the tuning curves around straight ahead. Congruency index was considered to be significantly different from 0 when both of the constituent R values were significant ($P < 0.05$). We denoted neurons having values of congruency index significantly different from 0 as SCI-congruent (congruency index > 0) or SCI-opposite (congruency index < 0). We also examined a global measure of

visual-vestibular congruency (see Supplementary Methods) and obtained similar results using this measure (Supplementary Figs. 11 and 12 online).

We used a linear weighted summation model to predict responses during cue combination from responses to each single-cue condition using the equation

$$R_{\text{prediction}} = w_{\text{vestibular}} \times R_{\text{vestibular}} + w_{\text{visual}} \times R_{\text{visual}}$$

where $R_{\text{vestibular}}$ and R_{visual} are responses from the single-cue conditions, and $w_{\text{vestibular}}$ 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 under the combined condition. Weights were constrained to lie between -20 and $+20$. The correlation coefficient (R) from a linear regression fit, which ranged from -1 to 1 , was used to assess goodness of fit. We also evaluated three variants of the linear model (described in Supplementary Methods and Supplementary Fig. 6).

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS

We thank A. Turner and E. White for monkey care and training, M.S. Banks for helpful advice, and L. Snyder, A. Pouget and R. Moreno Bote for comments on an earlier version of the manuscript. This work was supported by US National Institutes of Health grants EY017866 (to D.E.A.) and EY016178 and an EJLB Foundation grant (to G.C.D.).

AUTHOR CONTRIBUTIONS

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

Published online at <http://www.nature.com/natureneuroscience/>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>

- Jacobs, R.A. Optimal integration of texture and motion cues to depth. *Vision Res.* **39**, 3621–3629 (1999).
- Hillis, J.M., Watt, S.J., Landy, M.S. & Banks, M.S. Slant from texture and disparity cues: optimal cue combination. *J. Vis.* **4**, 967–992 (2004).
- Knill, D.C. & Saunders, J.A. Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Res.* **43**, 2539–2558 (2003).
- van Beers, R.J., Sittig, A.C. & Gon, J.J. Integration of proprioceptive and visual position-information: an experimentally supported model. *J. Neurophysiol.* **81**, 1355–1364 (1999).
- Ernst, M.O. & Banks, M.S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429–433 (2002).
- Alais, D. & Burr, D. The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* **14**, 257–262 (2004).
- Molholm, S., Ritter, W., Javitt, D.C. & Foxe, J.J. Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. *Cereb. Cortex* **14**, 452–465 (2004).
- Yuille, A.L. & Bulthoff, H.H. Bayesian decision theory and psychophysics. in *Perception as Bayesian Inference* (eds., Knill, D. & Richards, W.) 123–161 (Cambridge University Press, Cambridge, 1996).
- Mamassian, P., Landy, M.S. & Maloney, L.T. Bayesian modeling of visual perception. in *Probabilistic Models of the Brain: Perception and Neural Function* (eds., Rao, R.P.N., Olshausen, B.A. & Lewicki, M.S.) 13–36 (MIT Press, Cambridge, Massachusetts, 2002).
- Knill, D.C. & Pouget, A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* **27**, 712–719 (2004).
- Stein, B.E. & Meredith, M.A. *The Merging of the Senses* (MIT Press, Cambridge, Massachusetts, 1993).
- Felleman, D.J. & Van Essen, D.C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).
- Ghazanfar, A.A. & Schroeder, C.E. Is neocortex essentially multisensory? *Trends Cogn. Sci.* **10**, 278–285 (2006).
- Wallace, M.T., Ramachandran, R. & Stein, B.E. A revised view of sensory cortical parcellation. *Proc. Natl. Acad. Sci. USA* **101**, 2167–2172 (2004).
- Stanford, T.R., Quessy, S. & Stein, B.E. Evaluating the operations underlying multisensory integration in the cat superior colliculus. *J. Neurosci.* **25**, 6499–6508 (2005).
- Sugihara, T., Diltz, M.D., Averbach, B.B. & Romanski, L.M. Integration of auditory and visual communication information in the primate ventrolateral prefrontal cortex. *J. Neurosci.* **26**, 11138–11147 (2006).
- Avillac, M., Ben Hamed, S. & Duhamel, J.R. Multisensory integration in the ventral intraparietal area of the macaque monkey. *J. Neurosci.* **27**, 1922–1932 (2007).

18. Wuerger, S.M., Hofbauer, M. & Meyer, G.F. The integration of auditory and visual motion signals at threshold. *Percept. Psychophys.* **65**, 1188–1196 (2003).
19. Tanaka, K. *et al.* Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *J. Neurosci.* **6**, 134–144 (1986).
20. Tanaka, K., Fukada, Y. & Saito, H.A. 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.* **62**, 642–656 (1989).
21. Duffy, C.J. & Wurtz, R.H. Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *J. Neurophysiol.* **65**, 1329–1345 (1991).
22. Duffy, C.J. & Wurtz, R.H. Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *J. Neurosci.* **15**, 5192–5208 (1995).
23. Duffy, C.J. MST neurons respond to optic flow and translational movement. *J. Neurophysiol.* **80**, 1816–1827 (1998).
24. Page, W.K. & Duffy, C.J. Heading representation in MST: sensory interactions and population encoding. *J. Neurophysiol.* **89**, 1994–2013 (2003).
25. Gu, Y., Watkins, P.V., Angelaki, D.E. & DeAngelis, G.C. Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *J. Neurosci.* **26**, 73–85 (2006).
26. Takahashi, K. *et al.* Multimodal coding of three-dimensional rotation and translation in area MSTd: comparison of visual and vestibular selectivity. *J. Neurosci.* **27**, 9742–9756 (2007).
27. Bremmer, F., Kubischik, M., Pekel, M., Lappe, M. & Hoffmann, K.P. Linear vestibular self-motion signals in monkey medial superior temporal area. *Ann. NY Acad. Sci.* **871**, 272–281 (1999).
28. Ma, W.J., Beck, J.M., Latham, P.E. & Pouget, A. Bayesian inference with probabilistic population codes. *Nat. Neurosci.* **9**, 1432–1438 (2006).
29. Gu, Y., DeAngelis, G.C. & Angelaki, D.E. A functional link between area MSTd and heading perception based on vestibular signals. *Nat. Neurosci.* **10**, 1038–1047 (2007).
30. Wichmann, F.A. & Hill, N.J. The psychometric function. I. Fitting, sampling and goodness of fit. *Percept. Psychophys.* **63**, 1293–1313 (2001).
31. Britten, K.H., Shadlen, M.N., Newsome, W.T. & Movshon, J.A. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* **12**, 4745–4765 (1992).
32. Green, D.M. & Swets, J.A. *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).
33. Parker, A.J. & Newsome, W.T. Sense and the single neuron: probing the physiology of perception. *Annu. Rev. Neurosci.* **21**, 227–277 (1998).
34. Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S. & Movshon, J.A. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* **13**, 87–100 (1996).
35. Krug, K. A common neuronal code for perceptual processes in visual cortex? Comparing choice and attentional correlates in V5/MT. *Phil. Trans. R. Soc. Lond. B* **359**, 929–941 (2004).
36. Purushothaman, G. & Bradley, D.C. Neural population code for fine perceptual decisions in area MT. *Nat. Neurosci.* **8**, 99–106 (2005).
37. Shadlen, M.N., Britten, K.H., Newsome, W.T. & Movshon, J.A. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J. Neurosci.* **16**, 1486–1510 (1996).
38. Nienborg, H. & Cumming, B.G. Psychophysically measured task strategy for disparity discrimination is reflected in V2 neurons. *Nat. Neurosci.* **10**, 1608–1614 (2007).
39. Nienborg, H. & Cumming, B.G. Macaque V2 neurons, but not V1 neurons, show choice-related activity. *J. Neurosci.* **26**, 9567–9578 (2006).
40. Dodd, J.V., Krug, K., Cumming, B.G. & Parker, A.J. Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *J. Neurosci.* **21**, 4809–4821 (2001).
41. Celebrini, S. & Newsome, W.T. Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. *J. Neurosci.* **14**, 4109–4124 (1994).
42. Uka, T. & DeAngelis, G.C. Contribution of area MT to stereoscopic depth perception: choice-related response modulations reflect task strategy. *Neuron* **42**, 297–310 (2004).
43. Parker, A.J., Krug, K. & Cumming, B.G. Neuronal activity and its links with the perception of multi-stable figures. *Phil. Trans. R. Soc. Lond. B* **357**, 1053–1062 (2002).
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.* **69**, 128–142 (1993).
45. Van Essen, D.C., Maunsell, J.H. & Bixby, J.L. The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographic organization. *J. Comp. Neurol.* **199**, 293–326 (1981).
46. Komatsu, H. & Wurtz, R.H. Relation of cortical areas MT and MST to pursuit eye movements. III. Interaction with full-field visual stimulation. *J. Neurophysiol.* **60**, 621–644 (1988).
47. Komatsu, H. & Wurtz, R.H. Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. *J. Neurophysiol.* **60**, 580–603 (1988).
48. Desimone, R. & Ungerleider, L.G. Multiple visual areas in the caudal superior temporal sulcus of the macaque. *J. Comp. Neurol.* **248**, 164–189 (1986).