Cell PRESS
Neuron
Article

Perceptual Learning Reduces Interneuronal Correlations in Macaque Visual Cortex

Yong Gu,[1] Sheng Liu,[1] Christopher R. Fetsch,[1] Yun Yang,[1] Sam Fok,[1] Adhira Sunkara,[1] Gregory C. DeAngelis,[2] and Dora E. Angelaki[1,*]
[1]Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, MO 63110, USA
[2]Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627, USA
*Correspondence: angelaki@pcg.wustl.edu
DOI 10.1016/j.neuron.2011.06.015

SUMMARY

Responses of neurons in early visual cortex change little with training and appear insufficient to account for perceptual learning. Behavioral performance, however, relies on population activity, and the accuracy of a population code is constrained by correlated noise among neurons. We tested whether training changes interneuronal correlations in the dorsal medial superior temporal area, which is involved in multisensory heading perception. Pairs of single units were recorded simultaneously in two groups of subjects: animals trained extensively in a heading discrimination task, and "naive" animals that performed a passive fixation task. Correlated noise was significantly weaker in trained versus naive animals, which might be expected to improve coding efficiency. However, we show that the observed uniform reduction in noise correlations leads to little change in population coding efficiency when all neurons are decoded. Thus, global changes in correlated noise among sensory neurons may be insufficient to account for perceptual learning.

INTRODUCTION

Perceptual learning enhances sensory perception and leads to improved behavioral performance (Goldstone, 1998), but the neural basis of this phenomenon remains incompletely understood. One hypothesis is that responses of sensory neurons are altered by learning to increase the information that is encoded. In this case, one would expect to observe neural correlates of increased sensitivity in early sensory areas. However, previous studies have found little or no change in the tuning properties of single neurons in early visual cortex, and it remains unclear whether these changes could account for perceptual learning (Chowdhury and DeAngelis, 2008; Crist et al., 2001; Ghose et al., 2002; Law and Gold, 2008; Raiguel et al., 2006; Schoups et al., 2001; Yang and Maunsell, 2004; Zohary et al., 1994a). Alternatively, perceptual learning may arise from changes in how sensory information is decoded and interpreted by higher brain areas (Dosher and Lu, 1999; Law and Gold, 2008; Li et al., 2004).

Most neurophysiological studies of perceptual learning focused on the activity of individual neurons; however, behavior arises from population activity. By pooling information from many cells, the noise inherent in responses of single neurons could be reduced, thus improving coding efficiency. Theorists have shown that the information capacity of a population code depends on the correlated noise among neurons (Abbott and Dayan, 1999; Averbeck et al., 2006; Oram et al., 1998; Sompolinsky et al., 2001; Wilke and Eurich, 2002). In general, correlated noise could either decrease or increase the information transmitted by a population of neurons, depending on how correlated noise varies with the similarity of tuning between neurons ("signal correlations"; Averbeck et al., 2006; Oram et al., 1998; Wilke and Eurich, 2002). The impact of correlations could be strong when the relevant neuronal population is large (Bair et al., 2001; Shadlen et al., 1996; Smith and Kohn, 2008; Zohary et al., 1994b).

Whether perceptual learning improves population coding efficiency through changes in the correlated variability among sensory neurons remains unknown. Modest noise correlations have been measured in a number of cortical areas (V1: Bach and Krüger, 1986; Gutnisky and Dragoi, 2008; Poort and Roelfsema, 2009; Reich et al., 2001; Smith and Kohn, 2008) (but see Ecker et al., 2010) (V4: Cohen and Maunsell, 2009; Mitchell et al., 2009) (IT: Gawne et al., 1996; Gawne and Richmond, 1993) (MT: Cohen and Newsome, 2008; Huang and Lisberger, 2009; Zohary et al., 1994b), but how these correlations differ between untrained and trained animals has not, to our knowledge, been tested.

To examine the effect of training on correlated noise, we simultaneously recorded pairs of single neurons in the dorsal medial superior temporal area (MSTd), a multisensory area thought to be involved in heading perception based on optic flow and vestibular signals (Angelaki et al., 2009; Britten, 2008). Correlated noise among pairs of neurons was examined in two groups of animals: one group ("naive") was only trained to fixate; the other group ("trained") also learned to perform a fine heading discrimination task. Noise correlations were significantly weaker in trained than naive animals, whereas tuning curves, response variability, and discrimination thresholds of individual neurons were similar. Importantly, training reduced noise correlations uniformly, regardless of tuning similarity between pairs of neurons. As a result, if all neurons contribute equally to perception, this change in correlated noise is unlikely to account for improvements in perceptual sensitivity with training.

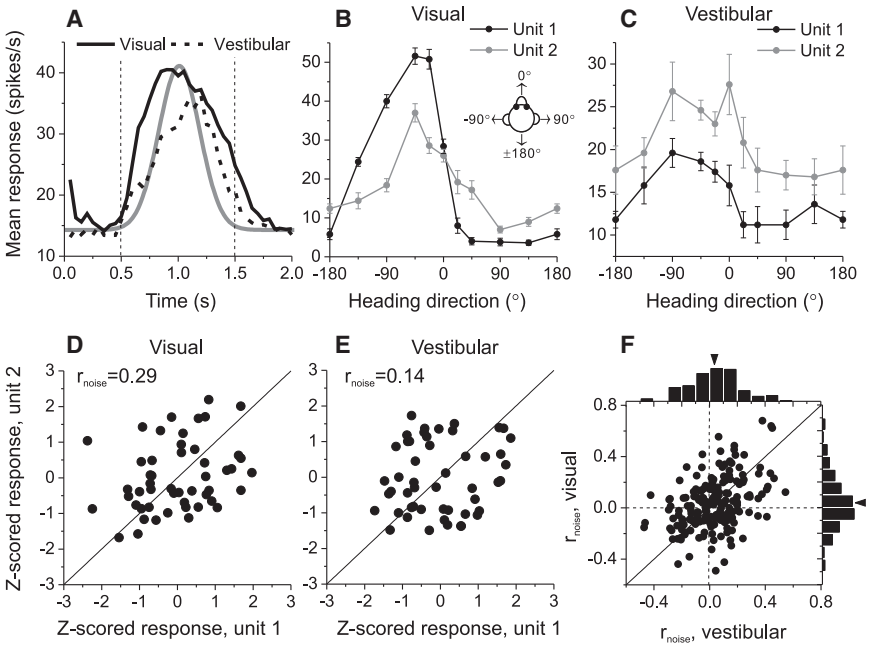

Figure 1. Measuring Noise Correlation (r_{noise}) between Pairs of Single Neurons in Area MSTd

(A) Response time courses for populations of neurons with significant tuning ($p < 0.05$, ANOVA) in the visual (solid curve, n = 231) and vestibular (dashed curve, n = 118) conditions. Gray curve represents the Gaussian velocity profile of the stimulus. Vertical dashed lines bound the time window over which spikes were counted for analysis.

(B and C) Visual and vestibular heading tuning curves, respectively, for a pair of simultaneously recorded MSTd neurons (black and gray curves). Error bars: SEM.

(D and E) Normalized responses from the same two neurons were weakly correlated across trials during visual (D) and vestibular (E) stimulation, with noise correlation values of 0.29 and 0.14, respectively.

(F) Comparison of noise correlations measured during visual and vestibular stimulation (n = 179). Arrow heads indicate mean values.

RESULTS

Monkeys were presented with two types of heading stimuli while maintaining fixation on a head-fixed target: inertial motion delivered by a motion platform in the absence of optic flow (vestibular condition) and optic flow stimuli presented while the animal was stationary (visual condition, see Experimental Procedures for details). Consistent with previous findings (Gu et al., 2006; Takahashi et al., 2007), many MSTd neurons were tuned to heading direction, and their responses mainly followed the Gaussian velocity profile of the stimulus (Figure 1A). We analyzed responses obtained during the middle 1 s of the stimulus period, during which neuronal activity was robust. Tuning curves of two simultaneously recorded cells are shown in Figures 1B and 1C. The similarity of heading tuning between pairs of neurons was quantified as the Pearson correlation coefficient of mean responses across all stimulus directions ("signal correlation", r_{signal}). For this example pair of neurons, $r_{signal} = 0.83$ and 0.79 for the visual and vestibular conditions, respectively.

Noise Correlations in Area MSTd

As in other cortical areas, the spike counts of MSTd neurons in response to an identical stimulus vary from trial to trial, as illustrated in Figure 1D (visual condition) and Figure 1E (vestibular condition). Each datum in these plots represents the spike counts of the two neurons from a single trial. Because heading direction varied across trials, spike counts from individual trials have been z-scored to remove the stimulus effect and allow pooling of data across directions (see Experimental Procedures). "Noise correlation" is then computed as the Pearson correlation coefficient of the normalized trial-by-trial spike counts, and reflects the degree of correlated variability across trials. For this example pair of cells, there was a weak positive correlation, such that when one neuron fired more spikes, the other neuron did as well (visual condition: $r_{noise} = 0.29$, $p = 0.04$, Figure 1D; vestibular condition: $R = 0.14$, $p = 0.3$, Figure 1E).

We first examined whether correlated noise in MSTd depends on stimulus modality (Figure 1F). Noise correlations computed from visual and vestibular responses were significantly correlated across 179 pairs of neurons ($R = 0.38$, $p << 0.001$, Spearman rank correlation), and their means were not significantly different (vestibular: 0.035 ± 0.014 SEM, visual: 0.039 ± 0.015, $p > 0.8$, paired t test). Thus, to gain statistical power, we recomputed r_{noise} by pooling z-scored responses across stimulus conditions, thereby obtaining a single value of r_{noise} for each pair of neurons.

As observed in other visual areas (Huang and Lisberger, 2009; Smith and Kohn, 2008), noise correlations depended on the distance between two simultaneously recorded MSTd neurons, as illustrated in Figure S1, which shows distributions of r_{noise} for three distance groups: <0.05 mm, 0.05–1 mm, and >1 mm. Average noise correlations were significantly greater than zero for the first two groups (<0.05 mm: 0.042 ± 0.021 SEM, $p = 0.049$, t test; 0.05–1 mm: 0.062 ± 0.024, $p = 0.011$), but not for the group of distant pairs (>1 mm: 0 ± 0.15, $p = 0.9$). Thus, the following analyses were focused on 127 neuronal pairs separated by <1 mm (results were similar for the whole data set).

Comparison of Noise Correlations in Trained and Naive Animals

Our main goal was to examine whether training modifies interneuronal correlations. Five animals were previously trained to perform a heading discrimination task, in which they reported whether their heading was leftward or rightward relative to straight ahead (Gu et al., 2007, 2008a). These monkeys' heading discrimination thresholds (corresponding to 84% correct) were high (>10°) at early stages of training, and gradually decreased to a plateau of only a few degrees (1~3°), as illustrated in

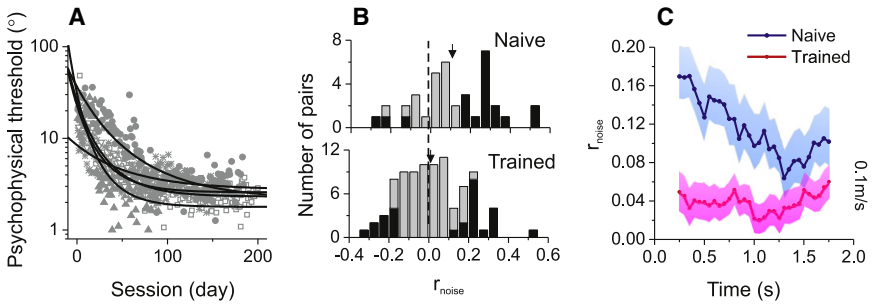

Figure 2. Training Effects on Behavior and Interneuronal Correlations

(A) Vestibular psychophysical thresholds from five monkeys (denoted by different symbol shapes) decreased gradually during training on a heading discrimination task. Solid curves: best fitting exponential functions for each animal. Thresholds are shown for the vestibular condition, not the visual condition, because optic flow stimuli were introduced later in training, and also because visual motion coherence varied across sessions to match visual and vestibular sensitivity (Gu et al., 2008a).

(B) Distributions of noise correlations for naive (top, n = 38) and trained (bottom, n = 89) animals. Black bars indicate r_{noise} values that are significantly different from zero. Arrows: population means.

(C) Average (±SEM) time course of noise correlations in trained (red, n = 89) and naive animals (blue, n = 38).

Figure 2A (Fetsch et al., 2009; Gu et al., 2007, 2008a). We measured noise correlations after these "trained" animals had reached asymptotic performance, and we compared them with correlations measured in three "naive" animals that had never been trained to perform any task other than visual fixation.

Our most conspicuous finding was a difference in mean r_{noise} between trained and naive animals (Figure 2B). Correlations in trained animals were shifted toward zero, as compared with those in naive animals. The mean noise correlation in the trained group (0.023 ± 0.017 SEM, n = 89) was significantly smaller than that for naive animals (0.116 ± 0.031, n = 38, p = 0.006, t test). Note that, for both groups of animals, r_{noise} was measured during an identical passive fixation task (see Experimental Procedures).

Because the stimulus was dynamic (Figure 1A, gray curve), we examined the time course of noise correlation in trained and naive animals by computing r_{noise} in 500 ms sliding windows (with 50 ms steps). As illustrated in Figure 2C, r_{noise} was significantly greater in naive than trained animals throughout the time course of the neural response (p = 0.002, permutation test, see Experimental Procedures). The difference in r_{noise} between naive and trained animals was largest at the beginning of the trial and gradually decreased with time ($R = -0.9$, $p << 0.001$, Spearman rank correlation, Figure S2A). Importantly, these observations held true when correlations were examined for individual animals (Figure S2B). Thus, the overall reduction in correlated noise among MSTd neurons was a robust finding in trained animals.

Effects of Training on Tuning Curves, Variability, and Sensitivity of Single Neurons

It is possible that the difference in correlated noise between naive and trained animals could be an indirect effect of training on the response properties of individual neurons. Moreover, training-related changes in correlated noise might emerge in parallel with changes in the heading sensitivity of single neurons. To address these issues, we examined the effect of training on the time courses of firing rates and response variability. As illustrated in Figures 3A and 3C, the time course of the population-average response to the preferred heading was indistinguishable between trained and naive animals (p = 0.8, permutation test, see Experimental Procedures). There was also no significant effect (p = 0.5, permutation test) of training on the time course of the Fano factor, which measures the ratio of response variance to mean response (Figures 3B and 3D, see also Experimental Procedures and Figure S3). This finding contrasts with a previous report that Fano factor in area V4 was significantly reduced after animals were trained to discriminate orientation (Raiguel et al., 2006). In MSTd, the difference in noise correlation between naive and trained animals does not appear to be linked to changes in firing rates or Fano factors.

We further explored whether training shaped the tuning properties of individual MSTd neurons. For this analysis, we only included neurons with significant heading tuning in the horizontal plane (p < 0.05, one-way ANOVA). To gain statistical power, we exploited a much larger database of single-unit responses from

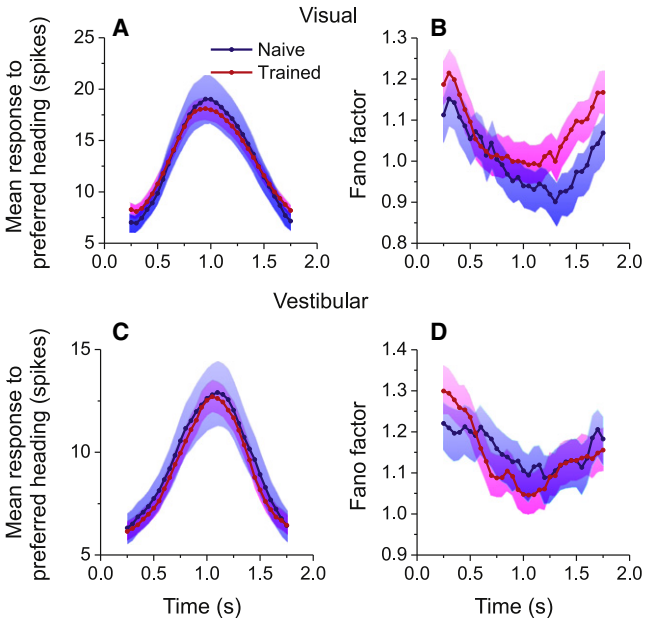

Figure 3. Training Does Not Affect the Time Courses of Mean Responses and Response Variability during Visual (top) and Vestibular (bottom) Stimulation

(A and C) Time course of the average response to stimuli presented at each cell's preferred heading in trained (red, n = 146) and naive animals (blue, n = 64). Error bands: SEM.

(B and D) Time course of Fano factor in trained (red) and naive (blue) animals. Error bands: 95% confidence intervals. Data were derived from the same 127 pairs of neurons as in Figures 2B and 2C.

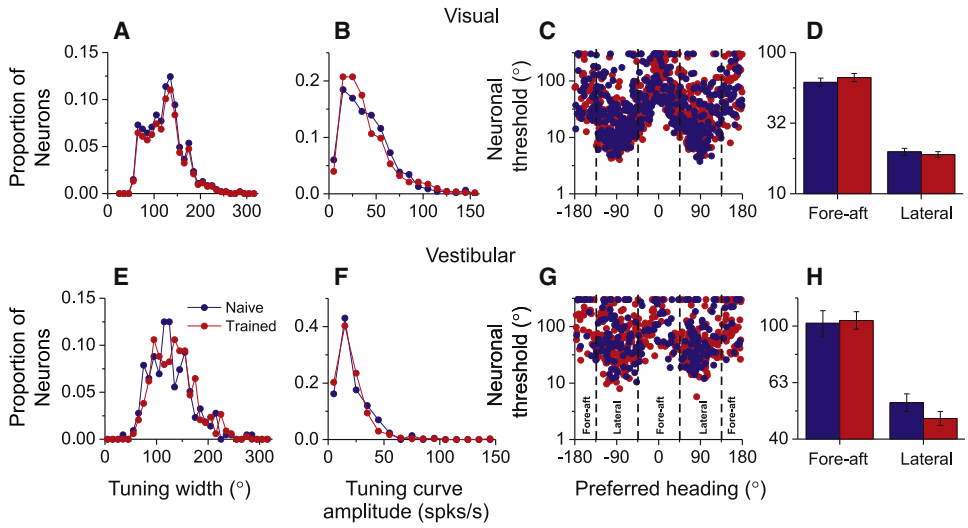

Figure 4. Effects of Training on Heading Tuning in MSTd during Visual (top) and Vestibular (bottom) Stimulation

(A and E) Distributions of tuning width (full width at half-maximum response) for naive (blue) and trained (red) animals.

(B and F) Distributions of tuning curve amplitude (peak to trough modulation).

(C and G) Neurons preferring lateral headings are more sensitive to heading variations around straight ahead than neurons preferring fore-aft motion, with little difference between naive and trained animals.

(D and H) Comparison of average neuronal sensitivity between fore-aft and lateral neurons (see G). Data were culled from a large database of MSTd neurons recorded with a single electrode. Only neurons with significant heading tuning ($p < 0.05$, ANOVA) were included (visual: n = 992; vestibular: n = 556). Error bars: SEM.

naive and trained animals, recorded with a single electrode (vestibular: n = 556; visual: n = 992). As shown in Figures 4A and 4E, distributions of tuning width (full width at half-height) were very similar for naive and trained animals. There was no significant difference in median tuning width for the visual condition (naive: 124.5° versus trained: 126°, $p = 0.21$, Wilcoxon rank- sum test). The difference in median tuning width was significant for the vestibular condition (naive: 121° versus trained: 131°, $p = 0.045$). However, this effect was weak and, notably, training slightly increased tuning width in the vestibular condition, an effect opposite to that expected if training increases discriminability (e.g., Yang and Maunsell, 2004). Similarly, as shown in Figures 4B and 4F, training did not have any significant effect on the distribution of tuning curve amplitudes in either the visual condition (naive: 35.4 spks/s versus trained: 31.8 spks/s, $p = 0.24$, Wilcoxon rank-sum test) or the vestibular condition (naive: 17.4 spks/s versus trained: 17.2 spks/s, $p = 0.36$). Thus, training animals to perform a fine heading discrimination task did not significantly shape the heading tuning of individual MSTd neurons.

However, it remains possible that training only shaped the tuning of a subset of neurons that were most informative for heading discrimination around the straight-forward reference used in training (e.g., Raiguel et al., 2006; Schoups et al., 2001). If so, then effects might only be seen for neurons most sensitive to heading variations around straight forward, and may have been missed in the above analysis. To examine this further, we interpolated tuning curves and used Fisher information analysis (Gu et al., 2010, see Experimental Procedures) to compute the sensitivity of each neuron for discriminating heading around straight forward. As shown in Figures 4C and 4G, the most sensitive neurons (lowest thresholds) are generally those that prefer lateral headings, such that their tuning curves have a steep slope around straight-ahead. For quantitative analysis, neurons were divided into two groups by heading preference: "fore-aft" neurons with heading preferences within 45° of forward (0°) or backward (±180°) motion, and "lateral" neurons with heading preferences within 45° of leftward (−90°) or rightward (90°) movements. Consistent with previous findings (Gu et al., 2008a, 2010), lateral neurons were significantly more sensitive than fore-aft neurons for heading discrimination around straight ahead ($p << 0.001$, Factorial ANOVA, Figures 4D and 4H). However, neuronal sensitivity was not significantly different between naive and trained animals ($p > 0.5$, factorial ANOVA) for either group of neurons, with no significant interaction effect ($p > 0.3$). In summary, whereas heading discrimination training clearly reduced correlated noise among MSTd neurons, we find no clear evidence that training altered the basic tuning properties or sensitivity of individual neurons, including those neurons that are most informative for performing the task. This result also generalizes to neuronal discrimination of heading about any arbitrary reference (Figure S4).

Training Effects on the Noise-Signal Correlation Structure

It is well established that r_{noise} is related to r_{signal} (Cohen and Maunsell, 2009; Cohen and Newsome, 2008; Gutnisky and Dragoi, 2008; Huang and Lisberger, 2009; Kohn and Smith, 2005; Smith and Kohn, 2008; Zohary et al., 1994b), so it is important to evaluate whether training alters this relationship. Figures 5A and 5B show the relationship between r_{noise} and r_{signal}, with

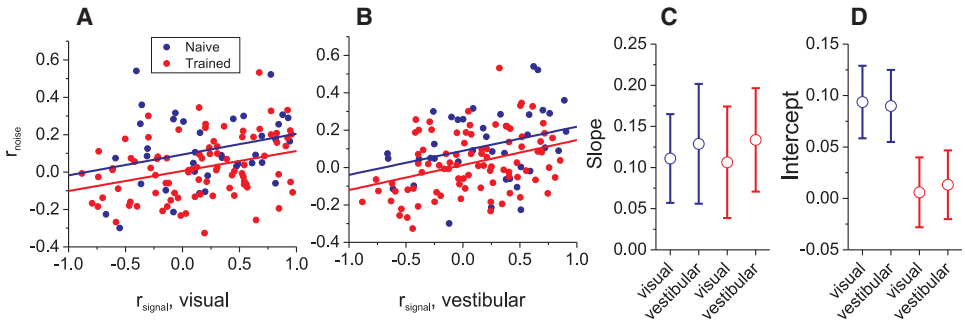

Figure 5. Relationship between Noise Correlation (r_{noise}) and Signal Correlation (r_{signal}) in MSTd

(A and B) Noise correlations depend significantly on r_{signal} computed from visual (A) or vestibular (B) tuning curves. Lines represent regression fits (ANCOVA). Red: data from trained animals (n = 89); blue: data from naive animals (n = 38).

(C and D) Regression slopes (C) and intercepts (D) obtained from the fits in (A) and (B). Error bars: 95% confidence intervals.

each datum corresponding to a pair of MSTd neurons. This relationship was quantified using general linear models (analysis of covariance, ANCOVA), with r_{signal} in each stimulus condition (visual or vestibular) as a continuous variable and training group (trained or naive) as a categorical factor. There was a significant positive correlation between r_{noise} and r_{signal} in both stimulus conditions (vestibular: $p = 0.0001$; visual: $p = 0.0004$, ANCOVA), reflecting the fact that noise correlations tended to be positive for pairs of neurons with similar tuning ($r_{signal} > 0$) and near zero or negative for pairs with opposite tuning ($r_{signal} < 0$).

Importantly, the slope of the relationship between r_{noise} and r_{signal} (Figures 5A and 5B) was not significantly affected by training (vestibular: $p = 0.9$; visual: $p = 0.9$, ANCOVA interaction effect), as also indicated by overlap of the 95% confidence intervals around the regression slopes (Figure 5C, nearly identical slopes were obtained by Type II regression). In contrast, training had a significant main effect on r_{noise} (vestibular: $p = 0.02$; visual; $p = 0.008$ ANCOVA), and the 95% confidence intervals around the regression intercepts were non-overlapping for naive and trained animals (Figure 5D). Thus, training reduced noise correlations uniformly across all signal correlations, such that the dependency of r_{noise} on r_{signal} remained unchanged.

Multisensory MSTd neurons can have matched visual and vestibular heading preferences ("congruent" cells) or mismatched preferences ("opposite" cells) (Gu et al., 2006, 2008a). Thus, we also tested whether r_{noise} depends on congruency. Specifically, the two units in each pair could be (1) both congruent, (2) both opposite, or (3) a mixture of congruent and opposite cells. As illustrated in Figure S5, r_{noise} was not substantially affected by congruency. Next, we incorporate these results into an information analysis to investigate how the fidelity of population activity changes between naive and trained animals.

Computation of Covariance Matrix

Although neurons were recorded pair-wise, our goal is to determine whether population activity in MSTd can account for the effect of training on behavioral sensitivity. For this purpose, we need to estimate the covariance matrix that characterizes correlations among the MSTd population in naive and trained animals. This was done by assigning each value of the covariance matrix according to the measured noise and signal correlation structures in our data set. Because r_{noise} depended on r_{signal} in both the vestibular and visual conditions (Figures 5A and 5B), both relationships were taken into account when constructing the covariance matrices. For simplicity, all neurons in the simulations discussed below were assumed to have congruent visual and vestibular heading preferences. Results were similar when variable congruency was introduced into the simulation, consistent with the observation that noise correlations were not strongly influenced by congruency (Figure S5).

We constructed covariance matrices with two different correlation structures (see Experimental Procedures): (1) r_{noise} depended on r_{signal} with regression slopes and intercept specified according to data from naive animals: $r_{noise} = 0.12 \times r_{signal,\ vestibular} + 0.091 \times r_{signal,\ visual} + 0.072$, and (2) r_{noise} depended on r_{signal} with slopes and intercept derived from trained animals: $r_{noise} = 0.12 \times r_{signal,\ vestibular} + 0.091 \times r_{signal,\ visual} + 0.005$. Note that the slopes were common across the two correlation structures, since no significant difference in slopes was found (Figure 5C). We then used these covariance matrices to compute the precision with which a population of MSTd neurons in naive or trained animals could discriminate heading, as described below. Importantly, noise correlations did not depend on whether trained monkeys performed a passive fixation task or the heading discrimination task ($p = 0.3$, t test), as shown in Figure S6 for a subset of neuronal pairs recorded in both tasks. Thus, we are justified in predicting heading discrimination performance from population activity measured during the fixation task for both trained and naive animals.

Effect of Training on Population Coding Efficiency

We computed population discrimination thresholds from the inverse of Fisher information (I_f), an upper bound on information capacity that can be extracted by any unbiased estimator (Abbott and Dayan, 1999; Seung and Sompolinsky, 1993). Predicted thresholds from I_f define the performance that an ideal observer could achieve, based on MSTd population activity, in a fine heading discrimination task. For a simulated population of neurons with independent noise, predicted thresholds decreased steadily with population size (Figure 6A, dashed black curve). As expected from previous findings (Bair et al., 2001;

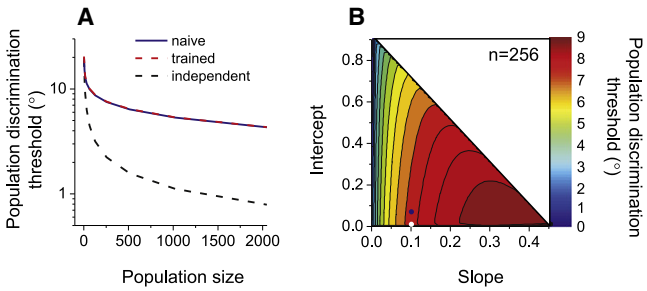

Figure 6. Impact of Noise Correlations on Population Coding Efficiency

(A) Population heading discrimination thresholds as a function of population size. Each simulated population contained neurons with wrapped-Gaussian tuning curves (bandwidth = 135°) and uniformly distributed heading preferences. Blue, dashed-red and dashed-black curves denote three correlation structures that correspond to naive, trained, and independent (r_{noise} = 0) neuronal cell pairs, respectively.

(B) Contour plot illustrating population (n = 256) discrimination thresholds (color coded) as a function of the slope and intercept of the relationship between r_{noise} and r_{signal}. The white region corresponds to a parameter range in which r_{noise} could exceed the allowable range of [−1 1]. Blue and white symbols denote the parameters measured in naive and trained animals, respectively.

Cohen and Maunsell, 2009; Shadlen et al., 1996; Smith and Kohn, 2008; Zohary et al., 1994b), correlated noise similar to that seen in our naive animals degraded population coding efficiency (Figure 6A, blue curve). For a simulated population of 2000 neurons, the predicted heading discrimination threshold was ~5-fold larger compared with the case of independent noise. Surprisingly, the uniform reduction in r_{noise} that we observed in trained animals (Figure 5) had little effect on predicted discrimination thresholds, as compared with naive animals (Figure 6A, red curve).

Why doesn't the reduction in mean noise correlation seen in trained animals improve the sensitivity of the population code? We simulated performance of a population of neurons using many covariance matrices that were constructed by systematically varying both the slope and intercept of the relationship between r_{noise} and r_{signal}. As shown in Figure 6B, predicted thresholds were very sensitive to changes in the slope of the relationship between r_{noise} and r_{signal}. In contrast, changes in the intercept of the r_{noise} versus r_{signal} relationship had weak effects on predicted thresholds. Counterintuitively, a uniform increase in r_{noise} (across all values of r_{signal}) produced a mild decrease in population thresholds, improving performance slightly (barely visible in Figure 6A, see also Abbott and Dayan, 1999; Wilke and Eurich, 2002). These simulations suggest that a uniform reduction of noise correlations in trained animals is expected to have little impact on discrimination performance.

This conclusion is based on the assumption that all neurons contribute to discrimination performance. We can infer from the simulations of Figure 6B that a change in noise correlation produces different effects for neurons with positive and negative signal correlations. To illustrate this, consider a population consisting of a single pair of neurons, having r_{signal} that could range from −1 (opposite heading preferences) to +1 (matched preferences). As illustrated in Figure 7A, reducing the noise correlation between this pair of neurons results in a lower population threshold (red curve below blue curve) when the pair of neurons has positive r_{signal}. In contrast, reducing noise correlation increases the predicted threshold for negative r_{signal} (see also Figure S7A). This simple prediction was confirmed when decoding responses of pairs of MSTd neurons. For each pair of neurons, we compute a discrimination threshold under the assumption of correlated noise, as well as the assumption of independent noise. As shown in Figure 7B, pairs of neurons with positive r_{signal} yield discrimination thresholds that increase with r_{noise}, whereas pairs with negative r_{signal} have discrimination thresholds that decrease with r_{noise} ($R = 0.49$, $p << 0.001$, Spearman rank correlation). Thus, in a population of neurons with an even mixture of positive and negative signal correlations, the opposite effects of correlated noise will counteract each other.

With this intuition in hand, we consider larger pool sizes (e.g., n = 256 in Figure 7C). If the direction preferences of neurons in the population are broadly distributed, roughly equal numbers of cell pairs will have positive and negative r_{signal} (Figure 7C, left inset) and population thresholds for naive and trained animals will be similar. If we narrow the distribution of direction preferences to generate more cell pairs with positive r_{signal}, the weaker noise correlations in trained animals substantially enhance coding efficiency (Figure 7C, middle and right insets, see also Figure S7B). The more similar the heading tuning among neurons in the population, the greater the benefit of reducing noise correlations. At best, however, the predicted population discrimination threshold for trained animals is ~8% lower than for naive animals (Figure 7C, right inset, see also Figure S7B). Clearly, the effect of interneuronal correlations on population coding depends critically on the structure of the correlations, which involves both the relationship between r_{noise} and r_{signal} and the distribution of tuning similarity among neurons.

Possible r_{signal} Distributions in Area MSTd

Might heading be decoded from a subpopulation of MSTd neurons with similar tuning properties (positive r_{signal}), such that the uniform reduction of r_{noise} in trained animals might improve discrimination performance? Although we cannot firmly exclude this possibility, two observations suggest that it is unlikely. First, electrical microstimulation of multiunit clusters with either leftward or rightward heading preferences can bias choices during a heading discrimination task (Britten and van Wezel, 1998, 2002; Gu et al., 2008b). Second, significant choice probabilities, which may reflect the contribution of single cortical neurons to behavior (Britten et al., 1996; Gu et al., 2007; Purushothaman and Bradley, 2005) (but also see Nienborg and Cumming, 2010), were reported for MSTd neurons preferring both rightward and leftward headings (Gu et al., 2007, 2008a). Thus, we further examined the dependence of choice probability and noise correlation on heading preference.

Compared with neurons with lateral heading preferences, neurons with a preference for fore-aft movement show significantly smaller choice probabilities ($p = 0.019$, t test, Figures 8A and 8B). This result is consistent with the notion that neurons with direction preferences deviated away from straight ahead

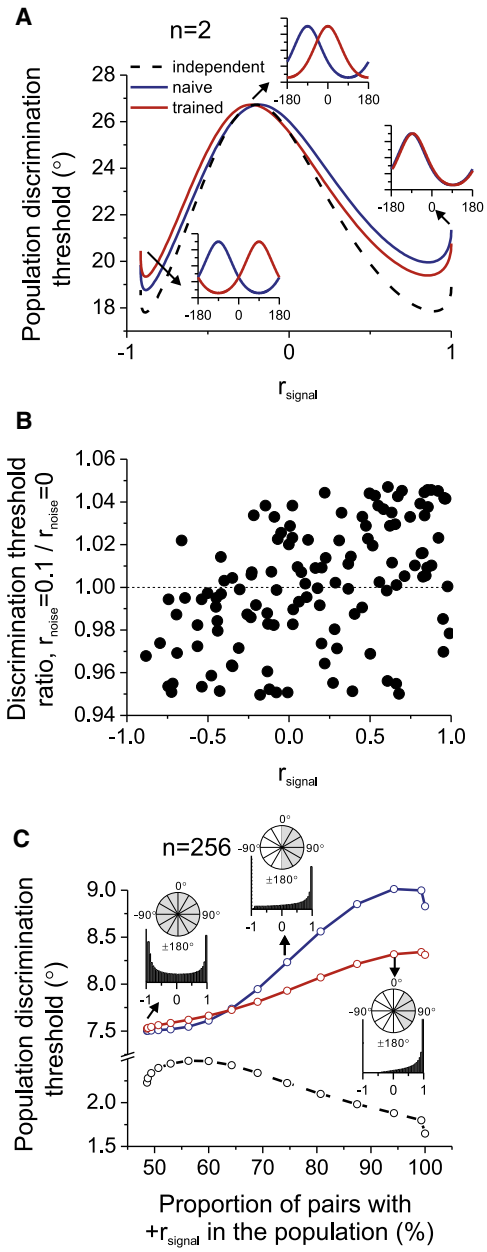

Figure 7. Reduced Noise Correlations Improve Coding Efficiency for Neurons with Similar Tuning and Reduce Coding Efficiency for Neurons with Dissimilar Tuning

(A) Heading discrimination thresholds of a pair of neurons with various signal correlations. One neuron has a fixed heading preference of −90°, while the other cell's heading preference varies from −90° (right inset) to 0° (middle inset) to 90° (left inset).

(B) For each pair of MSTd neurons (each datum), we computed the ratio of discrimination thresholds for r_{noise} = 0.1 and r_{noise} = 0. A ratio of 1 (dashed line) indicates that correlated noise did not affect sensitivity.

(C) Predicted discrimination thresholds for a population of 256 neurons with a variable distribution of heading preferences. From left to right, the range of heading preferences narrowed from a uniform distribution ([−180° 180°]) to only rightward headings near 90°. This generated varying distributions of signal correlations, as illustrated for three cases with proportions of positive r_{signal} values equal to 49% (heading preference range: [−180° 180°]), 75% ([0° 180°]), and 94% ([30° 150°]), as denoted by gray shading in insets. Blue, red, and black curves represent correlation structures corresponding to naive, trained, and independent, respectively.

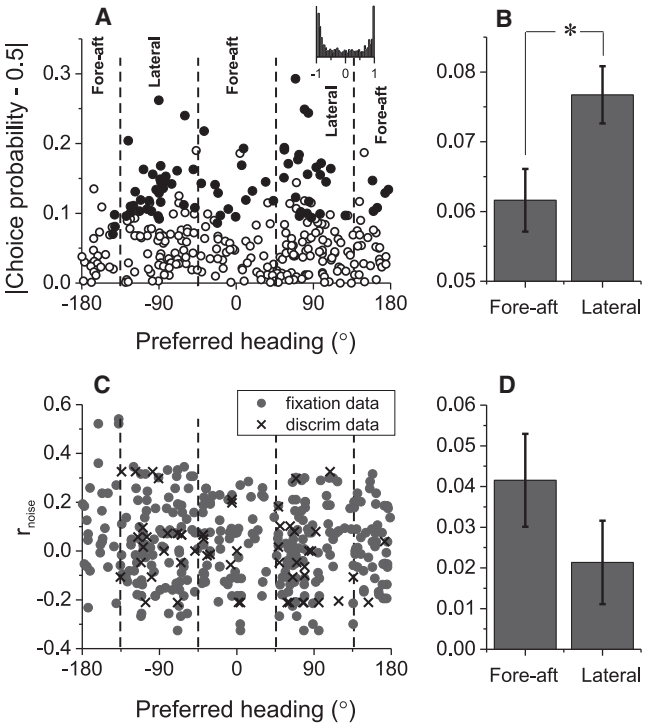

Figure 8. Relationships between Choice Probability, Noise Correlation, and Heading Preferences in MSTd

(A) Choice probability tends to be more deviated away from the chance level of 0.5 for neurons with lateral heading preferences. Filled symbols denote choice probabilities significantly different from 0.5 ($p < 0.05$, permutation test). Dashed lines denote category boundaries for lateral and fore-aft cells. Inset: distribution of expected signal correlations when heading preferences are drawn randomly from cells with significant choice probabilities.

(B) Mean ± SEM of the choice probability data from (A), sorted into groups for lateral and fore-aft neurons (*$p < 0.05$). Data were collected from previous experiments conducted with a single electrode (n = 311), and pooled across vestibular and visual conditions.

(C) Noise correlations did not depend significantly on heading preference. Pairs of cells denoted by gray circles and crosses were recorded during fixation (n = 328) and discrimination tasks (n = 55), respectively. For each pair, the noise correlation is plotted twice, at the preferred heading of each neuron.

(D) Mean ± SEM of the noise correlation data from (C).

are more sensitive to small heading variations and thus contribute more to perception (Gu et al., 2007; Purushothaman and Bradley, 2005). Importantly, there was no significant difference in average choice probability between neurons preferring leftward and rightward headings ($p = 0.11$, t test), suggesting that the population of neurons that contributes to heading perception includes cells with both positive and negative signal correlations (inset in Figure 8A).

Interestingly, a similar dependence on heading preference was not observed for noise correlations in trained animals. As shown in Figures 8C and 8D, there was no significant

dependence of noise correlation on the heading preferences of MSTd neurons ($p = 0.2$, t test). Indeed, the average noise correlation for lateral neurons is a bit smaller than that for the fore-aft neurons. This finding suggests that the variation in choice probability with heading preference (Figures 8A and 8B) is not driven just by correlated noise, but also depends on other factors such as how the signals are read out by decision circuitry.

DISCUSSION

By recording simultaneously from pairs of neurons in macaque area MSTd, we have shown that interneuronal correlations are weaker, on average, in animals trained to perform a fine heading discrimination task as compared with animals experienced only in visual fixation tasks. Although we did not record from the same animals before and after training, the difference in correlated noise between naive and trained subjects was highly significant and consistent across animals within each group.

Our findings suggest that changes in the average strength of noise correlations are not sufficient to account for the effect of training on discrimination performance. The difference in r_{noise} between naive and trained animals was uniform and independent of tuning similarity. If all neurons are decoded uniformly, the increased information capacity of neuronal pools with similar tuning is counteracted by the decreased information capacity of neuronal pools with dissimilar tuning curves. Thus, the effect of correlated noise on discrimination performance is conditional on both the relationship between r_{noise} and r_{signal} and on the demands of the task which may recruit different neuronal pools into play.

Properties of Noise Correlations in MSTd

Compared with noise correlations observed in area MT (Bair et al., 2001; Cohen and Newsome, 2008; Huang and Lisberger, 2009; Zohary et al., 1994b), the average noise correlation in our MSTd sample (distance <1 mm) was substantially weaker (trained animals: 0.023; naive animals: 0.116). The average correlation values we have seen in trained animals are similar to those reported in a recent study of macaque primary visual cortex (Ecker et al., 2010).

We found that noise correlations in MSTd are independent of the sensory stimulus modality (visual or vestibular), but depend on distance such that nearby neurons tend to have stronger correlations than more distant pairs (Huang and Lisberger, 2009; Lee et al., 1998; Smith and Kohn, 2008). Correlations in MSTd also depend strongly on tuning similarity, such that neurons with similar tuning curves tend to have greater correlated noise. In addition, we observed that noise correlations decrease in the presence of a stimulus as compared with prestimulus baseline activity. This result is consistent with previous studies showing that noise correlations decreased following stimulus onset (Smith and Kohn, 2008) and increased with stimulus intensity (e.g., contrast) (Huang and Lisberger, 2009; Kohn and Smith, 2005).

Possible Explanations for the Effect of Training on Correlated Noise

Before accepting the conclusion that correlated noise in MSTd was reduced as a consequence of perceptual learning, we consider some alternatives. One possibility is that naive monkeys undergo larger fluctuations in behavioral state (e.g., arousal, attention) than trained animals, and this might cause slow fluctuations in neuronal responses that can inflate noise correlations (Bair et al., 2001; Ecker et al., 2010; Lampl et al., 1999). To address this issue, we removed slow fluctuations in neural responses by renormalizing the data before computing noise correlations (see Experimental Procedures, Zohary et al., 1994b). This operation had little effect on our measurements, for both naive and trained animals (Figure S8). This suggests that slow fluctuations in response driven by variations in behavioral state do not account for the greater noise correlations seen in naive animals.

Another possibility is that naive animals fixate the visual target less reliably and make more frequent microsaccades that could induce correlations among neural responses (e.g., Bair and O'Keefe, 1998). However, we found that naive animals fixate as accurately as trained animals (Figure S8A). Indeed, naive monkeys as a group made significantly fewer microsaccades than trained animals (Figure S8B). Hence, the reduction of correlated noise in trained animals is unlikely to be explained by differences in eye movements between the two groups of animals.

Two recent studies have indicated that attention directed toward the receptive field could reduce correlated noise among pairs of neurons in area V4 (Cohen and Maunsell, 2009; Mitchell et al., 2009). Although both naive and trained monkeys only performed a passive fixation task in our study, trained animals might have paid more attention to the heading stimuli due to their relevance in the discrimination task. We cannot exclude this possibility, but three aspects of our results are inconsistent with an explanation based on attention. First, attention typically increases neuronal activity (Desimone and Duncan, 1995; Kastner and Ungerleider, 2000; Reynolds and Chelazzi, 2004; Reynolds and Heeger, 2009; Treue and Maunsell, 1999), but our analysis shows that mean responses were not significantly different between naive and trained animals (Figure 3). Second, the reduction in noise correlation with increased attention was also accompanied by decreased neuronal variability (Fano factor, Cohen and Maunsell, 2009; Mitchell et al., 2009). However, we did not find a significant difference in Fano factor between naive and trained animals. Finally, there was no difference in noise correlation between the fixation and discrimination tasks for a subset of pairs of neurons that were recorded during both tasks (Figure S6). This result is consistent with an earlier study in which noise correlations in area MT were found to be similar during a motion discrimination task and a visual fixation task (Zohary et al., 1994b).

Any fluctuation in common inputs could cause correlated variability among target neurons. It is thus possible that training decreases the shared, common input to area MSTd, likely on a long timescale during learning (Chowdhury and DeAngelis, 2008). The effect of training on neural circuitry may have occurred at two levels. First, training may have altered the feed-forward sensory input to MSTd from other cortical and subcortical areas, without changing the average tuning properties of single neurons (Jenkins et al., 1990; Recanzone et al., 1993; Weinberger, 1993). Second, training may have altered

feedback connections to MSTd, including feedback from decision circuitry. Our results are consistent with recent findings that perceptual learning does not substantially alter sensory cortical representations, but rather sculpts the decoding of sensory signals by decision circuitry (Dosher and Lu, 1999; Law and Gold, 2008). If training alters the read out of heading signals from MSTd, this, in turn, may modify the shared feedback to MSTd neurons from downstream circuitry. It is currently not possible to discern which of these training-related changes contributes most to the reduction in correlated noise that we have observed.

Although our data suggest that learning does not alter the sensory representation of heading in a manner that could account for the improvement in behavioral sensitivity with training, it is important to note that we cannot rule out the possibility that training altered the heading tuning and sensitivity of neurons in other brain areas that may also be involved in heading perception, such as area VIP (Zhang and Britten, 2011). In addition, although we assume that noise correlations in MSTd were altered by perceptual learning, we cannot exclude the possibility that some other aspect of training, such as learning the operational rules of the task, may have driven the changes in correlated noise that we have observed. Finally, it is unclear whether the effect of training on correlated noise is specific to tasks for which area MSTd is thought to provide critical input. If we had trained animals to perform a task that was irrelevant to self-motion perception, such as a somatosensory or auditory discrimination task, we presumably would not expect to see changes in correlated noise in MSTd. However, this possibility remains to be tested.

Consequences for Population Coding Efficiency

Despite a robust effect of training on the average noise correlation in MSTd, our simulations show that an optimal, unbiased decoding of all neurons does not predict a substantial change in performance due to learning. Indeed, theorists have shown that correlated noise may or may not harm population coding (Abbott and Dayan, 1999; Averbeck et al., 2006; Wilke and Eurich, 2002). In general, positively correlated noise between neurons with similar tuning (or more generally, any situation in which both neurons fire more strongly under one stimulus/task condition than another) harms the signal to noise ratio of the population code because it cannot be removed by pooling across neurons (Bair et al., 2001; Shadlen et al., 1996; Zohary et al., 1994b). Reducing shared noise among neurons in such cases is thus expected to improve population sensitivity.

Indeed, the effect of attention on the fidelity of population codes appears to follow this logic (Cohen and Maunsell, 2009). In a typical spatial attention task, most neurons with receptive fields at the attended location will increase their response. Because attention has a consistent polarity of effect on the responses of nearby neurons, stronger attention will tend to increase the responses of both neurons in a pair. Hence, most pairs of nearby neurons will have positive signal correlations with respect to the effect of attention. As a result, a reduction in correlated noise due to attention can improve the signal-to-noise ratio of the population code. However, in other contexts for which signals are decoded from populations that include neurons with dissimilar tuning properties, increasing correlated noise can improve the signal-to-noise ratio of a population code (Figure 7A), as differences in tuning effectively cancel more of the noise in a population response (Abbott and Dayan, 1999; Averbeck et al., 2006; Poort and Roelfsema, 2009; Wilke and Eurich, 2002). Reducing correlated noise in the latter case can harm the coding efficiency of the population. In our heading discrimination task, it is likely that responses are decoded from neurons with a broad range of heading preferences (Gu et al., 2008b, 2010); in this context, reducing correlated noise uniformly across neurons with all signal correlations (Figures 5A and 5B) does not improve the fidelity of the neural code (Figure 6A). Thus, the impact of correlated noise on population coding depends on (1) the structure of noise correlations and their dependence on signal correlation, and (2) the composition of neuronal pools upon which decoding is based.

We conclude that the effects of training on heading discrimination are not likely to be driven by the reduction in correlated noise that we have observed in area MSTd. Combined with previous observations that perceptual learning has little or no effect on basic tuning properties of single neurons in visual cortex (Chowdhury and DeAngelis, 2008; Crist et al., 2001; Ghose et al., 2002; Law and Gold, 2008; Raiguel et al., 2006; Schoups et al., 2001; Yang and Maunsell, 2004; Zohary et al., 1994a), our results suggest that changes in sensory representations are not necessarily involved in accounting for the improvements in behavioral sensitivity that accompany perceptual learning (at least for some sensory systems and tasks; see also Bejjanki et al., 2011). Rather, our findings support the idea that perceptual learning may primarily alter the routing and/or weighting of sensory inputs to decision circuitry, an idea that has recently received experimental support (Chowdhury and DeAngelis, 2008; Law and Gold, 2008, 2009).

EXPERIMENTAL PROCEDURES

Subjects

Physiological experiments were performed in 8 male rhesus monkeys (Macaca mulatta) weighing 4–8 kg. Animals were chronically implanted with a plastic head-restraint ring that was firmly anchored to the apparatus to minimize head movement. All monkeys were implanted with scleral coils for measuring eye movements in a magnetic field (Robinson, 1963). Animals were trained using standard operant conditioning to fixate visual targets for fluid reward. All animal surgeries and experimental procedures were approved by the Institutional Animal Care and Use Committee at Washington University and were in accordance with NIH guidelines.

Motion Stimuli

Neurons were tested with two types of motion stimuli using a custom-built virtual reality system (Gu et al., 2006, 2007, 2008b). In the "vestibular" stimulus condition, monkeys were passively translated by a motion platform (Moog 6DOF2000E; East Aurora, NY) along a smooth trajectory (Gaussian velocity profile with peak-acceleration of $\sim$1 m/s^2 and duration of 2 s, Figure 1A). In the "visual" stimulus condition, optic flow was provided by rear-projecting images onto a tangent screen in front of the monkey using a 3-chip DLP projector (Christie Digital Mirage 2000) that was mounted on the motion platform. Visual stimuli (90 × 90°) depicted movement through a 3D cloud of stars that occupied a virtual space 100 cm wide, 100 cm tall, and 50 cm deep. The stimulus contained multiple depth cues, including horizontal disparity, motion parallax, and size information.

Experimental Protocol and Task

Animals were trained to maintain visual fixation on a head-fixed target at the center of the screen. Eye position was required to stay within a 2° × 2° electronic window throughout each trial in order to receive a water/juice reward. The majority of the data presented here were recorded while passively fixating animals experienced a range of different heading directions that spanned the horizontal and/or vertical plane (Gu et al., 2006; Takahashi et al., 2007). Specifically, headings relative to straight ahead were 0, ±22.5, ±45, ±90, ±135°, ±180°. Different heading directions and stimulus types (visual or vestibular) were interleaved randomly within a single block of trials. Each distinct stimulus was typically repeated five times (minimum of three repetitions for inclusion).

In each trial, a fixation point first appeared at the center of the screen. After fixation was established for 100–200 ms, the motion stimulus began and lasted for 2 s. In the vestibular condition, the motion platform always began its movement from a common central position. The animal was rewarded if they maintained visual fixation throughout the duration of the stimulus. At the end of the trial (or when fixation was broken), the fixation point disappeared and the motion platform moved back to the original central position during a 2 s intertrial interval. In the visual condition, the random-dot field appeared on the display after fixation was established, and again moved for 2 s. The dots then disappeared and the animal was rewarded for maintaining fixation, followed again by a 2 s intertribal interval.

Three animals were trained only to perform the passive fixation task, whereas five animals had been extensively trained to perform a heading discrimination task (Fetsch et al., 2009; Gu et al., 2007, 2008a), in which they were asked to report whether their perceived heading was leftward or rightward relative to straight ahead by making a saccade to one of two choice targets. For a subpopulation of neurons in these trained animals, responses were obtained while the animals performed both the fixation task and the heading discrimination task.

Electrophysiological Recordings

We conducted extracellular recordings of action potentials from single neurons in area MSTd. For most recordings, 2 to 4 tungsten electrodes (Frederick Haer, Bowdoinham, ME; tip diameter 3 μm, impedance 1–2 MΩ at 1 kHz) were used to record multiple single neurons simultaneously. In some cases (57 pairs), two to four electrodes were placed inside multiple guide tubes separated by 0.8–25 mm (different hemispheres). In other cases (55 pairs), multiple electrodes were placed inside a single guide tube. The distance between two simultaneously recorded neurons was estimated from both the horizontal and vertical (depth) coordinates (shank diameter = 75 μm).

Data from another 67 cell pairs were obtained from previous recordings with a single electrode (Fetsch et al., 2007; Gu et al., 2006, 2007; Takahashi et al., 2007), for which a second cell was isolated offline using spike sorting software (Spike2, Cambridge Electronics Design). Only pairs of neurons from a single electrode that showed clearly separate clusters in the first three principle components of the spike waveform were included in the sample. Since the exact distance between neurons recorded from a single electrode was unknown, we arbitrarily assigned it to be 50 μm. Although noise correlations were slightly greater for pairs of neurons recorded from a single electrode (0.042 ± 0.02) than for pairs recorded from different electrodes (0.033 ± 0.015), this difference was modest and not significant ($p > 0.7$, t test). Thus, data collected with single and multiple electrodes were pooled for analysis, yielding 179 cell pairs from a total of 270 neurons (maximum of 5 pairs in an experiment).

Area MSTd was located ~15 mm lateral to the midline and ~2–6 mm posterior to the interaural plane, and was identified using both MRI scans and neurophysiological response properties (see Gu et al., 2006 for details). MSTd neurons had large receptive fields that typically occupied a quadrant or a hemifield on the display screen and were often centered in the contralateral visual field but could extend well into the ipsilateral field. Once the electrodes were targeted to MSTd, we recorded from any neuron that was spontaneously active or could be activated by patches of flickering dots.

Data Analysis

Noise and Signal Correlations

Noise correlation (r_{noise}) was computed as the Pearson correlation coefficient (ranging between −1 and 1) of the trial-by-trial responses from a pair of neurons driven by the same stimulus (Bair et al., 2001; Zohary et al., 1994b). The response in each trial was taken as the number of spikes during the middle 1 s of the stimulus period (Gu et al., 2006). For each heading direction, responses were z-scored by subtracting the mean response and dividing by the standard deviation. This operation removed the effect of heading on the responses, such that the measured noise correlation reflected trial-to-trial variability. To avoid artificial correlations caused by outliers, we removed data points with z-scores larger than 3 (Zohary et al., 1994b). We then pooled data across headings to compute r_{noise}; the corresponding p value was used to assess the significance of correlation for each pair of neurons.

Because there was no significant difference in r_{noise} between visual and vestibular stimulus conditions (Figure 1F), we pooled responses across conditions to gain statistical power. To remove slow fluctuations in responsiveness that could result from changes in cognitive state over time (e.g., arousal), we renormalized the z-scored responses in blocks of 20 trials, as described by Zohary et al. (1994b). This additional normalization had no significant effect on r_{noise} ($p > 0.3$, paired t test; $R = 0.9$, $p << 0.001$, Spearman rank correlation, $n = 127$, Figure S8). More importantly, the effect of renormalization on noise correlations was similar in naive and trained animals ($p = 0.7$, interaction effect, $p = 0.9$, group effect, ANCOVA, Figure S8). This suggests that the greater noise correlations in naive animals were not the result of larger slow fluctuations in neural response (Ecker et al., 2010), such as might arise if naive animals experienced greater fluctuations in arousal during the session.

Signal correlation (r_{signal}) was computed as the Pearson correlation coefficient (ranging between −1 and 1) between the tuning curves from two simultaneously recorded neurons. Tuning curves for each stimulus condition were constructed by computing the mean response (average firing rate during the middle 1 s of the stimulus duration) across trials for each heading direction.

Permutation Test

Permutation tests were applied to test for significant differences between trained and naive animals with respect to: the difference in time courses of noise correlation (Figure 2C), mean response (Figures 3A and 3C), and Fano factor (Figures 3B and 3D). We first computed the sum of squared differences between two time courses:

$$x^2 = \sum_{i=1}^{n} (T_{trained,i} - T_{naive,i})^2 \quad (1)$$

using a 500 ms sliding window moved in 50 ms steps, for a total of 31 data points. We then created permuted naive and trained groups by randomly drawing data from the original groups, pooled together. Within each cell, all of the responses were preserved (no shuffling across trials). We computed a new x^2 value for each permutation ($x^2_{permuted}$), and this process was repeated 10,000 times. A p value was computed as the proportion of $x^2_{permuted} > x^2$. A difference between the two groups of animals was considered significant if $p < 0.05$.

Fano Factor

Fano factor, or the variance/mean ratio, was computed from log-log scatter plots of the variance of the spike count against the mean spike count, and this was done for each 500 ms time window used to compute time courses. The data were fit by minimizing the orthogonal distance to the fitted line (type II regression). The slope was generally close to 1 and was thus forced to be 1 for convenience, such that variance scaled linearly with mean spike count. The Fano factor was then computed as $10^{\wedge intercept}$ (see Figure S3).

Population Coding

Fisher information (I_F) provides an upper limit on the precision with which an unbiased estimator can discriminate between small variations in a variable (x) around a reference value (x_{ref}) (Pouget et al., 1998; Seung and Sompolinsky, 1993). We computed the smallest deviation in heading around straight ahead (threshold, Δx) that could be reliably discriminated (at 84% correct) by an ideal observer:

$$\Delta x = \frac{\sqrt{2}}{\sqrt{I_f}} \quad (2)$$

where I_F was computed according to (Abbott and Dayan, 1999):

$$I_F(x_{ref}) = f'(x_{ref})^T Q^{-1}(x_{ref}) f'(x_{ref}) + 0.5 Tr\left[Q'(x_{ref}) Q^{-1}(x_{ref}) Q'(x_{ref}) Q^{-1}(x_{ref})\right] \quad (3)$$

Here, f′ denotes the derivative of a matrix of tuning curves; superscript T denotes the matrix transpose, Tr represents the trace operation, and superscript −1 indicates the matrix inverse. The reference heading was straight ahead in our simulations (x_{ref} = 0°). Q represents the covariance matrix of neural responses, which was given by

$$Q_{i,j}(x_{ref}) = r_{i,j}\sqrt{f_i(x_{ref})f_j(x_{ref})} \quad (4)$$

where $r_{i,j}$ denotes the noise correlation between the i^{th} and j^{th} neurons. When i = j, $r_{i,j}$ was set to 1. When i ≠ j, $r_{i,j}$ was assigned according to a linear relationship between noise and signal correlation:

$$r_{i,j} = a_{vestibular} \times r_{signal,vestibular,i,j} + a_{visual} \times r_{signal,visual,i,j} + b \quad (5)$$

We minimized the orthogonal distance between the fit plane and the raw data using type II regression.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Experimental Procedures and eight figures and can be found online at doi:10.1016/j.neuron.2011.06.015.

ACKNOWLEDGMENTS

This work was supported by grants from National Institutes of Health (EY019087 to D.E.A., and EY016178 to G.C.D.).

Accepted: June 8, 2011
Published: August 24, 2011

REFERENCES

Abbott, L.F., and Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. Neural Comput. *11*, 91–101.

Angelaki, D.E., Gu, Y., and DeAngelis, G.C. (2009). Multisensory integration: psychophysics, neurophysiology, and computation. Curr. Opin. Neurobiol. *19*, 452–458.

Averbeck, B.B., Latham, P.E., and Pouget, A. (2006). Neural correlations, population coding and computation. Nat. Rev. Neurosci. *7*, 358–366.

Bach, M., and Krüger, J. (1986). Correlated neuronal variability in monkey visual cortex revealed by a multi-microelectrode. Exp. Brain Res. *61*, 451–456.

Bair, W., and O'Keefe, L.P. (1998). The influence of fixational eye movements on the response of neurons in area MT of the macaque. Vis. Neurosci. *15*, 779–786.

Bair, W., Zohary, E., and Newsome, W.T. (2001). Correlated firing in macaque visual area MT: time scales and relationship to behavior. J. Neurosci. *21*, 1676–1697.

Bejjanki, V.R., Beck, J.M., Lu, Z.L., and Pouget, A. (2011). Perceptual learning as improved probabilistic inference in early sensory areas. Nat. Neurosci. *14*, 642–648.

Britten, K.H. (2008). Mechanisms of self-motion perception. Annu. Rev. Neurosci. *31*, 389–410.

Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S., and Movshon, J.A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. Vis. Neurosci. *13*, 87–100.

Britten, K.H., and van Wezel, R.J. (1998). Electrical microstimulation of cortical area MST biases heading perception in monkeys. Nat. Neurosci. *1*, 59–63.

Britten, K.H., and Van Wezel, R.J. (2002). Area MST and heading perception in macaque monkeys. Cereb. Cortex *12*, 692–701.

Chowdhury, S.A., and DeAngelis, G.C. (2008). Fine discrimination training alters the causal contribution of macaque area MT to depth perception. Neuron *60*, 367–377.

Cohen, M.R., and Newsome, W.T. (2008). Context-dependent changes in functional circuitry in visual area MT. Neuron *60*, 162–173.

Cohen, M.R., and Maunsell, J.H. (2009). Attention improves performance primarily by reducing interneuronal correlations. Nat. Neurosci. *12*, 1594–1600.

Crist, R.E., Li, W., and Gilbert, C.D. (2001). Learning to see: experience and attention in primary visual cortex. Nat. Neurosci. *4*, 519–525.

Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. Annu. Rev. Neurosci. *18*, 193–222.

Dosher, B.A., and Lu, Z.L. (1999). Mechanisms of perceptual learning. Vision Res. *39*, 3197–3221.

Ecker, A.S., Berens, P., Keliris, G.A., Bethge, M., Logothetis, N.K., and Tolias, A.S. (2010). Decorrelated neuronal firing in cortical microcircuits. Science *327*, 584–587.

Fetsch, C.R., Wang, S., Gu, Y., Deangelis, G.C., and Angelaki, D.E. (2007). Spatial reference frames of visual, vestibular, and multimodal heading signals in the dorsal subdivision of the medial superior temporal area. J. Neurosci. *27*, 700–712.

Fetsch, C.R., Turner, A.H., DeAngelis, G.C., and Angelaki, D.E. (2009). Dynamic reweighting of visual and vestibular cues during self-motion perception. J. Neurosci. *29*, 15601–15612.

Gawne, T.J., and Richmond, B.J. (1993). How independent are the messages carried by adjacent inferior temporal cortical neurons? J. Neurosci. *13*, 2758–2771.

Gawne, T.J., Kjaer, T.W., Hertz, J.A., and Richmond, B.J. (1996). Adjacent visual cortical complex cells share about 20% of their stimulus-related information. Cereb. Cortex *6*, 482–489.

Ghose, G.M., Yang, T., and Maunsell, J.H. (2002). Physiological correlates of perceptual learning in monkey V1 and V2. J. Neurophysiol. *87*, 1867–1888.

Goldstone, R.L. (1998). Perceptual learning. Annu. Rev. Psychol. *49*, 585–612.

Gu, Y., Watkins, P.V., Angelaki, D.E., and DeAngelis, G.C. (2006). Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. J. Neurosci. *26*, 73–85.

Gu, Y., DeAngelis, G.C., and Angelaki, D.E. (2007). A functional link between area MSTd and heading perception based on vestibular signals. Nat. Neurosci. *10*, 1038–1047.

Gu, Y., Angelaki, D.E., and Deangelis, G.C. (2008a). Neural correlates of multisensory cue integration in macaque MSTd. Nat. Neurosci. *11*, 1201–1210.

Gu, Y., DeAngelis, G.C., and Angelaki, D.E. (2008b). Effects of microstimulation in area MSTd on heading perception based on visual and vestibular cues. Society for Neuroscience abstract 460.9.

Gu, Y., Fetsch, C.R., Adeyemo, B., Deangelis, G.C., and Angelaki, D.E. (2010). Decoding of MSTd population activity accounts for variations in the precision of heading perception. Neuron *66*, 596–609.

Gutnisky, D.A., and Dragoi, V. (2008). Adaptive coding of visual information in neural populations. Nature *452*, 220–224.

Huang, X., and Lisberger, S.G. (2009). Noise correlations in cortical area MT and their potential impact on trial-by-trial variation in the direction and speed of smooth-pursuit eye movements. J. Neurophysiol. *101*, 3012–3030.

Jenkins, W.M., Merzenich, M.M., Ochs, M.T., Allard, T., and Guíc-Robles, E. (1990). Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. J. Neurophysiol. *63*, 82–104.

Kastner, S., and Ungerleider, L.G. (2000). Mechanisms of visual attention in the human cortex. Annu. Rev. Neurosci. *23*, 315–341.

Kohn, A., and Smith, M.A. (2005). Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. J. Neurosci. *25*, 3661–3673.

Lampl, I., Reichova, I., and Ferster, D. (1999). Synchronous membrane potential fluctuations in neurons of the cat visual cortex. Neuron *22*, 361–374.

Law, C.T., and Gold, J.I. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. Nat. Neurosci. *11*, 505–513.

Law, C.T., and Gold, J.I. (2009). Reinforcement learning can account for associative and perceptual learning on a visual-decision task. Nat. Neurosci. *12*, 655–663.

Lee, D., Port, N.L., Kruse, W., and Georgopoulos, A.P. (1998). Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. J. Neurosci. *18*, 1161–1170.

Li, W., Piëch, V., and Gilbert, C.D. (2004). Perceptual learning and top-down influences in primary visual cortex. Nat. Neurosci. *7*, 651–657.

Mitchell, J.F., Sundberg, K.A., and Reynolds, J.H. (2009). Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. Neuron *63*, 879–888.

Nienborg, H., and Cumming, B. (2010). Correlations between the activity of sensory neurons and behavior: how much do they tell us about a neuron's causality? Curr. Opin. Neurobiol. *20*, 376–381.

Oram, M.W., Földiák, P., Perrett, D.I., and Sengpiel, F. (1998). The "Ideal Homunculus": decoding neural population signals. Trends Neurosci. *21*, 259–265.

Poort, J., and Roelfsema, P.R. (2009). Noise correlations have little influence on the coding of selective attention in area V1. Cereb. Cortex *19*, 543–553.

Pouget, A., Zhang, K., Deneve, S., and Latham, P.E. (1998). Statistically efficient estimation using population coding. Neural Comput. *10*, 373–401.

Purushothaman, G., and Bradley, D.C. (2005). Neural population code for fine perceptual decisions in area MT. Nat. Neurosci. *8*, 99–106.

Raiguel, S., Vogels, R., Mysore, S.G., and Orban, G.A. (2006). Learning to see the difference specifically alters the most informative V4 neurons. J. Neurosci. *26*, 6589–6602.

Recanzone, G.H., Schreiner, C.E., and Merzenich, M.M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. J. Neurosci. *13*, 87–103.

Reich, D.S., Mechler, F., and Victor, J.D. (2001). Independent and redundant information in nearby cortical neurons. Science *294*, 2566–2568.

Reynolds, J.H., and Chelazzi, L. (2004). Attentional modulation of visual processing. Annu. Rev. Neurosci. *27*, 611–647.

Reynolds, J.H., and Heeger, D.J. (2009). The normalization model of attention. Neuron *61*, 168–185.

Robinson, D.A. (1963). A method of measuring eye movement using a scleral search coil in a magnetic field. IEEE Trans. Biomed. Eng. *10*, 137–145.

Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. Nature *412*, 549–553.

Seung, H.S., and Sompolinsky, H. (1993). Simple models for reading neuronal population codes. Proc. Natl. Acad. Sci. USA *90*, 10749–10753.

Shadlen, M.N., Britten, K.H., Newsome, W.T., and Movshon, J.A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. J. Neurosci. *16*, 1486–1510.

Smith, M.A., and Kohn, A. (2008). Spatial and temporal scales of neuronal correlation in primary visual cortex. J. Neurosci. *28*, 12591–12603.

Sompolinsky, H., Yoon, H., Kang, K., and Shamir, M. (2001). Population coding in neuronal systems with correlated noise. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. *64*, 051904.

Takahashi, K., Gu, Y., May, P.J., Newlands, S.D., DeAngelis, G.C., and Angelaki, D.E. (2007). Multimodal coding of three-dimensional rotation and translation in area MSTd: comparison of visual and vestibular selectivity. J. Neurosci. *27*, 9742–9756.

Treue, S., and Maunsell, J.H. (1999). Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. J. Neurosci. *19*, 7591–7602.

Weinberger, N.M. (1993). Learning-induced changes of auditory receptive fields. Curr. Opin. Neurobiol. *3*, 570–577.

Wilke, S.D., and Eurich, C.W. (2002). Representational accuracy of stochastic neural populations. Neural Comput. *14*, 155–189.

Yang, T., and Maunsell, J.H. (2004). The effect of perceptual learning on neuronal responses in monkey visual area V4. J. Neurosci. *24*, 1617–1626.

Zhang, T., and Britten, K.H. (2011). Parietal area VIP causally influences heading perception during pursuit eye movements. J. Neurosci. *31*, 2569–2575.

Zohary, E., Celebrini, S., Britten, K.H., and Newsome, W.T. (1994a). Neuronal plasticity that underlies improvement in perceptual performance. Science *263*, 1289–1292.

Zohary, E., Shadlen, M.N., and Newsome, W.T. (1994b). Correlated neuronal discharge rate and its implications for psychophysical performance. Nature *370*, 140–143.