HHS Public Access

Author manuscript

Neuron. Author manuscript; available in PMC 2017 April 06.

Published in final edited form as:

Neuron. 2016 April 6; 90(1): 165–176. doi:10.1016/j.neuron.2016.02.012.

Signal, noise, and variation in neural and sensory-motor latency

Joonyeol Lee[1,2], **Mati Joshua**[3], **Javier F. Medina**[4], and **Stephen G. Lisberger**[5]

[1]Center for Neuroscience Imaging Research, Institute for Basic Science (IBS), Suwon 16419, Republic of Korea

[2]Department of Biomedical Engineering, Sungkyunkwan University, Suwon 16419, Republic of Korea

[3]Edmond and Lily Safra Center for Brain Sciences, the Hebrew University, Jerusalem, Israel

[4]Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA

[5]Department of Neurobiology, Duke University School of Medicine, Durham, NC, USA

Abstract

Analysis of the neural code for sensory-motor latency in smooth pursuit eye movements reveals general principles of neural variation and the specific origin of motor latency. The trial-by-trial variation in neural latency in MT comprises: a shared component expressed as neuron-neuron latency correlations; and an independent component that is local to each neuron. The independent component arises heavily from fluctuations in the underlying probability of spiking with an unexpectedly small contribution from the stochastic nature of spiking itself. The shared component causes the latency of single neuron responses in MT to be weakly predictive of the behavioral latency of pursuit. Neural latency deeper in the motor system is more strongly predictive of behavioral latency. A model reproduces both the variance of behavioral latency and the neuron-behavior latency correlations in MT if it includes realistic neural latency variation, neuron-neuron latency correlations in MT, and noisy gain control downstream from MT.

Keywords

reaction time; smooth pursuit eye movements; correlated variation; neuron-behavior correlations; floccular complex; Abducens; area MT

Proofs and correspondence to: Joonyeol Lee, Department of Biomedical Engineering, Sungkyunkwan University, 2066 Seobu-ro, Jangan-gu, Suwon-si, Gyeonggi-do, 16419, Republic of Korea, Voice: +82 (31) 299-4359, joonyeol@skku.edu.

The authors declare no competing financial interests.

Author contributions: Conceptualization, J.L and S.G.L.; Methodology, J.L. and S.G.L.; Investigation, J.L., M.J., and J.F.M.; Formal Analysis, J.L.; Writing – Original Draft, J.L. and S.G.L.; Writing – Review and Editing, J.L. and S.G.L.; Funding Acquisition, S.G.L.

Introduction

To create accurate movements, the nervous system must deal with the fact that neural responses vary from movement-to-movement and, within a single movement, from neuron-to-neuron. Some of the neuron-to-neuron variation is noise and can be eliminated by pooling across groups of nominally identical neurons. But, some of the variation is correlated across neurons and is more challenging to eliminate, at least by simple averaging. The correlated, or shared, variation can propagate through a sensory motor circuit and drive movement-by-movement variation in performance. Our goal is to understand the nature of motor variation, its representation in the nervous system, and how neuron-by-neuron variation is managed.

The system we study, smooth pursuit eye movements, offers a number of advantages that allow us to ask how neural circuits operate in the face of variation. The architecture of the neural circuit for pursuit is known (Figure 1). We can record and quantify the latency and amplitude of neural responses at many sites in the circuit. We already know from a thorough analysis of trial-by-trial variation in pursuit behavior (Osborne et al., 2005) and neural responses (Hohl et al., 2013; Lee and Lisberger, 2013) how pursuit speed and direction are controlled by the representations of target speed and direction in extrastriate area MT. In line with the general theory outlined in the prior paragraph, our analysis revealed two components of variation in the spike frequency of MT neurons. One component is shared, presumably results from common inputs to MT neurons, and propagates through the sensory-motor circuit to create movement-by-movement variation in pursuit direction and speed. A second component is independent in each neuron, has been presumed to arise from the stochastic nature of spiking, and can be eliminated by averaging across large populations of neurons.

Our goal in the present study was to use sensory-motor latency to understand better how variation is processed in the nervous system. Latency varies from trial-to-trial, and also depends on many features of a behavioral task, including the modality of the sensory input, the strength of a sensory stimulus, the contingencies of the behavioral task, and the level of attention to the stimulus. In vision, for example, brighter stimuli cause neural responses with shorter latencies as well as shorter behavioral reaction times (Lisberger and Westbrook, 1985; Miles et al., 1986; Gawne et al., 1996; Warzecha and Egelhaaf, 2000; Bell et al., 2006; Lee et al., 2007; Sundberg et al., 2012). Also, auditory sensory neural responses have short latencies compared to visual sensory responses (Celesia and Puletti, 1971; Clark et al., 1995), as do the motor responses to acoustic versus visual stimuli (Shelton and Kumar, 2010).

Sensory latency defines motor latency to a large degree, but the simple relationships between stimulus and response may belie more complex and interesting internal neural processing. In some paradigms, for example, neural activity ramps to a threshold that appears to trigger a motor response (Hanes and Schall, 1996; Roitman and Shadlen, 2002). The rate of rise of the ramp, and therefore the behavioral reaction time, could vary from movement-to-movement, perhaps as a function of the strength of the evidence favoring a given movement (Gold and Shadlen, 2002; Roitman and Shadlen, 2002), the magnitude of the expected reward (Hikosaka and Watanabe, 2000; Hollerman et al., 1998; Metzger et al., 2006;

Takikawa et al., 2002), or the subject's level of attention (Herrington and Assad, 2009). Even for a simple visual-motor reflex, the ocular following response, latency can be explained by a model that integrates the sensory signal to a threshold, reaching the threshold sooner for stimuli of higher contrast (Miles et al., 1986). In pursuit eye movements, behavioral evidence suggests that latency variation represents sensory errors in estimating the time when the target started to move (Osborne et al., 2005). However, the behavioral evidence is mute on the form of the sensory variation. Does sensory latency vary in lock-step with motor latency? Or does integration of variable sensory amplitudes lead to ramps that reach threshold at varying times?

In the present paper, we measure the variation in the estimates of latency in area MT, as well as in downstream motor areas. In broad strokes, the resulting picture agrees with that for speed and direction. The code for latency contains shared variation that is expressed as neuron-neuron latency correlations. The shared variation in MT, when amplified by a noisy downstream gain control, is sufficient to account for the behavioral variation in latency. However, a surprise arises from analysis of the sources of independent variation in latency in MT. Much of the independent variation is present in the underlying probability of spiking, which we equate loosely with the integrated synaptic inputs to the neuron. An unexpectedly small amount of the independent variation could be attributed to the stochastic nature of spiking. We think that our analysis provides lessons for the processing of neural signals more generally.

Results

Our approach was to study the trial-by-trial relationship between the latency of smooth pursuit eye movements and the latency of neural responses in extrastriate area MT. Our data show that 40–70% of the trial-by-trial variation in behavioral latency can be attributed to correlated variation in the latency of responses to moving stimuli in MT; our modeling suggests potential sources of the remaining latency variation. In the course of analyzing and modeling our data, we also discovered a surprising feature of how noise is managed as signals pass through a neural circuit.

We have reanalyzed, for new purposes, previously published recordings of the electrical activity of single neurons and pairs of single neurons in extrastriate visual area MT (Lee and Lisberger, 2013), as well as single neurons in the floccular complex of the cerebellar cortex, and two groups of brainstem neurons (Joshua et al., 2013; Medina and Lisberger, 2007). During recordings, monkeys tracked a "step-ramp" of target position (Figure 2A) after Rashbass (1961). After a short reaction time of ~100 ms, monkeys initiate smooth eye movements in the direction of target motion. A target step customized for the ramp speed compensates for the latency of pursuit and allows monkeys to produce smooth eye movements without saccades, or with saccades delayed sufficiently beyond the initiation of pursuit to allow us to analyze saccade-free initiation of smooth pursuit. For each recording session, we presented many repetitions of the same small number of step-ramp target motions.

Relationship between pursuit latency and neural latency

We first estimated the strength of the trial-by-trial relationship between neural and behavioral latency (gray traces in Figure 2A). We used an objective method to evaluate pursuit latency for each behavioral trial (see ***Methods***), and sorted all trials for a given neuron into five groups (quintiles) according to pursuit latency. For each quintile, we used a similar objective method to estimate the latency and amplitude of the mean spike density function (see ***Methods***). The example MT neuron analyzed in Figures 2B and C shows a clear progression of neural response latencies in relation to the 5 groups of behavioral response latencies. Regression analysis shows a strong and statistically significant relationship between neural and behavioral latency (Figure 2D, t-test, $p < 0.05$), and a much weaker relationship between neural response amplitude and behavioral latency (Figure 2E).

The strength of the relationship between neural and behavioral latency increases as the signals progress from MT to the final motor pathway (Figure 2F, Supplementary Figure 1A–D). For each neuron, we computed the sensitivity of neural latency to behavioral latency as the slope of the regression line (Figure 2D). Sensitivity averaged 0.18 for neurons from area MT and increased to 0.77 for Purkinje cells in the cerebellar floccular complex. Sensitivity was approximately 1 for the last-order brainstem interneurons that receive monosynaptic inhibition from the floccular complex (FTNs) and for neurons in the Abducens nucleus. In contrast, we found no consistent relationship between neural response amplitude and pursuit latency in any of the 4 areas we studied (Figure 2G).

We conclude that neural latency rather than response amplitude determines behavioral latency. At least for the conditions we used, pursuit latency does not appear to be determined by a ramp-to-threshold mechanism. This conclusion is supported by the observation that mean pursuit latency tracks mean neural latency in MT almost perfectly when we change the form or contrast of the moving visual stimulus (J. Yang and S.G. Lisberger, unpublished observations).

Trial-by-trial correlation between pursuit latency and neural latency

The nervous system relies on single spike trains of many neurons to generate a motor response to a single presentation of a stimulus. To understand how a population of neurons encodes the latency of a behavioral response, we analyze the trial-by-trial variation in neural latencies, and its correlation i) among MT neurons and ii) between MT neurons and pursuit latency. Previously, the magnitude and sign of neuron-behavior amplitude correlations has been informative about the details of neural processing downstream from area MT (Hohl et al., 2013).

To estimate neuron-behavior latency correlations, we used the fact that the regression slope of z-scored data is equal to Pearson's correlation coefficient (Rodgers and Nicewander, 1988). We estimated neural and behavioral latency in each trial (see ***Methods***), binned the trials for each neuron into 5 equal-sized quintiles ordered by response latency (Figure 3A), and plotted the z-scored neural versus behavioral latencies (Figure 3C, D). For z-scoring, we estimated the standard deviation of the latency of neural responses by plotting average latency in each of the quintiles as the x-values in a distribution (Figures 3B, blue symbols)

and fitting a Gaussian with the constraint that the x-value for each point was the mean of one of the quintiles of the Gaussian.

The example neuron illustrated in Figure 3 showed a trial-by-trial correlation of 0.55 between neural and pursuit latency (Figure 3C), and a weaker correlation of 0.13 between neural response magnitude and pursuit latency (Figure 3D). We obtained very similar results when we analyzed data from single trials (gray symbols) instead of the averages for the 5 quintiles (large black symbols), and when we binned the data according to neural latency instead of pursuit latency.

For the four populations of neurons we analyzed, the neuron-behavior latency correlations increased as signals passed through the pursuit circuit. Values averaged 0.15 in area MT, and increased to 0.37 in floccular Purkinje cells and about 0.6 in floccular target neurons (FTNs) and Abducens neurons (Figure 3E, see Supplementary Figure 1E–H for population histograms). In contrast, the trial-by-trial correlation between neural response magnitude and pursuit latency, obtained through a procedure analogous to the latency analysis, was very close to zero in all four areas (Figure 3F, Supplementary Figure 1I–L). This supports our conclusion that neural latency rather than response amplitude determines behavioral latency.

Validation of computational methods for estimating neural latency

Measurement of neural latencies from single trials is challenging, especially for cortical neurons with irregular spiking. To provide an independent test of the veracity of our methods, we created model neurons with artificial spike trains where we controlled the latency in each trial. Analysis of the model spike trains with the same procedures used for our real data confirmed that we were able to recover the latencies used to generate each model spike train.

For each run of the simulation we started with the mean probability of spiking from our population of MT neurons (Figure 4A) and jittered it in time with a standard deviation of 20 ms to create 100 underlying probabilities of spiking with 100 different, known latencies. Then, we generated three spike trains (see ***Methods***) for each trial using three different coefficients of variation (0.4, 0.7, and 1.0). We evaluated the latency of each model spike train and measured the correlation between estimates of latency and the known latency in each model trial. We did so either using data from individual trials, or by binning the trials into quintiles, as before.

Our analysis revealed strong but imperfect correlations between the latency of the underlying probability of spiking and the measured latency (Figure 4B). As expected, the stochastic nature of spiking introduces some independent variation in neural latency. Thus, the correlation between the latency of the underlying probability of spiking and the measured latency of the model spike trains decreased as the CV of the artificial spiking increased. The correlation was larger for the quintiled data, indicating that our analysis improved marginally the estimates of the actual latency. Of course, it is not possible to dissociate how much of the “independent” noise in this model comes from neuron’s spiking statistics (CV) and how much comes from measurement errors. However, we are reassured by the fact that the quintile analysis is able to recover 36% of the variance of the latency of

the underlying probability of spiking when the CV is 1, and much more when CV is 0.4 (Figure 4B).

Several observations demonstrate that our measures of latency were not influenced by response amplitude. First, when we created model spike trains with fluctuations in both response latency and amplitude, measured latency was correlated only with actual latency and measured rate was correlated only with actual rate (Figure 4C, D). The response amplitudes of the model spike trains varied by a scale factor of standard deviation 0.4, and latencies had a standard deviation of 20 ms. Second, the sensitivity of behavioral latency to neural latency was unaffected by restricting the range of neural response amplitudes to ±0.4 or ±0.8 standard deviations of the mean response amplitude for each neuron (Supplementary Figure 2A). Even though the neural latency varied systematically, the neural response amplitude was invariant across the quintiles for both the full data set and the analysis of trials with response amplitudes within 0.8 SD of the mean (Supplementary Figure 2B). Finally, if neural response amplitude had a systematic influence on our estimates of latency, then correlations between behavioral latency and neural response amplitude should have appeared in our data. They did not (Figures 2F, G).

Neuron-neuron latency correlations in MT

Correlated variation of neural responses within a population is an important factor in determining the magnitude of the variation of decoded population output (Zohary et al., 1994; Huang and Lisberger, 2009; Moreno-Bote et al., 2014) as well as the neuron-behavior correlations (Shadlen et al., 1996; Lee and Lisberger, 2013). Indeed, without neuron-neuron correlations, averaging across the large number of neurons in MT should render behavioral variation very small and preclude neuron-behavior correlations (Schoppik et al., 2008). Therefore, our next step was to analyze neuron-neuron latency correlations for 40 pairs of MT neurons that met two criteria: i) >50 repetitions of a stimulus that evoked responses statistically higher than baseline firing rate in both neurons; and ii) mean response latencies of both neurons <100 ms.

Neuron-neuron latency correlations were present in many pairs of neurons. The simplest demonstration of this phenomenon comes from averaging the firing rate of one neuron in a pair within the quintiles defined by the single-trial neural latencies of the other neuron (Figure 5A–D). The average firing rates separate to a greater extent when ordered by each neuron itself (Figure 5A, B), but they still show a correlated progression when ordered by the latencies of the other neuron (Figure 5C, D). Analysis of the z-scored data reveals neuron-neuron latency correlations that were statistically significant (slopes different from zero, F-test, $p<0.05$) in Figures 5F and G for the analyses based on both the quintiles (large black symbols) and the individual trials (gray symbols). The neuron-neuron latency correlations tended to be positive across our whole sample of MT neurons (Figure 5E), and the mean correlation of ~0.1 was significantly different from zero ($p<10^{-5}$, two-tailed t-test). Details of how we estimated neuron-neuron latency correlations appear in the ***Methods***.

Components of variation in neural latencies

Up to this point, we have considered the relationship between neural and behavioral latency, in effect analyzing a decoding question. Now, we turn to analysis of the encoding of latency by using a computational approach to understand the origin of the variation in latencies of the spike trains of MT neurons. We ask whether variation in what we will call the "underlying probability of spiking aligns with the shared component of variation in neural latency, while the stochastic nature of spike timing aligns with the independent component.

First, we estimated the variation in the latency of the underlying probability of spiking. We represented the underlying probability of spiking for each neuron as the average spike density function across 50–100 trials. We then simulated the underlying probability of spiking in individual trials by jittering the mean response in time according to a Gaussian distribution with a standard deviation of σ_{prob}. We created a model spike train for each jittered trace, constrained to have the same coefficient of variation of interspike intervals (CV) as the neuron being modeled. We estimated the neural latency of each simulated spike train, and calculated the standard deviation of latency, σ_{spikes}. We repeated the simulations using a range of values of σ_{prob} and established the relationship between σ_{prob} and σ_{spikes} for each neuron (Figure 6A). We then estimated the actual standard deviation of the latency of the underlying probability of spiking as the value σ_{prob} for the value of σ_{spikes} that matched our original analysis of the actual data.

The analysis in Figure 6 divides latency variation into two components that are due to the underlying probability of spiking (the inputs to a neuron?) and the stochastic timing of spikes in each trial. Almost all MT neurons showed non-zero values of the standard deviation of the latency of the underlying probability of spiking, and many showed standard deviations that were quite large (x-axis of Figure 6B). We take this as evidence that variation in the underlying probability of spiking is a major contributor to the variation in the latency of neural responses in MT. In Figure 6B, the vertical distance from each symbol to the line of slope 1 shows the surprisingly small contribution of the stochastic timing of spikes to variation in neural latency.

Next, we estimated the trial-by-trial MT-pursuit latency correlations between spiking probability and the eye velocity in the initiation of smooth pursuit. To do so, we used a relationship derived in the ***Methods***:

$$R_{prob,eye} \approx R_{spikes,eye} \frac{\sigma\left(L_{spikes}\right)}{\sigma\left(L_{prob}\right)} \quad (1)$$

Here, $R_{prob,eye}$ and $R_{spikes,eye}$ define the MT-pursuit latency correlations for the underlying probability of spiking and the actual spike trains, and σ_{prob} and σ_{spikes} were defined earlier. We can compute $R_{prob,eye}$ because we measured $R_{spikes,eye}$ and σ_{spikes}, and we can read the value of σ_{prob} off of graphs like Figure 6A.

The neuron-behavior latency correlation for the underlying probability of spiking averaged 0.28 across our sample of neurons (Figure 6C, black bars), compared to an average correlation of 0.15 for the latency of spike trains (Figure 6C, gray bars). Thus, the latency of

the underlying probability of spiking of the average MT neuron explains ~8% of pursuit latency variation. This number seems large, considering the number of synapses between area MT and the final behavior and the number of neurons in area MT. However, we also note that the neuron-behavior latency correlation for the underlying probability of spiking would be 1 if all independent variation were attributed to the stochastic nature of spiking. We conclude that trial-by-trial variation in the underlying probability of firing is an important contributor to independent variation in neural latency in MT. Later, we will evaluate quantitatively how much of the latency variation might arise downstream from MT (Schoppik et al., 2008).

We estimated the neuron-neuron latency correlation for the underlying probability of spiking using the same computational approach described above. We created a computational model of each pair of neurons, created neuron-neuron correlation in the underlying probability of spiking, and laid down simulated spike trains for each trial in each neuron (see ***Methods***). We analyzed the correlations between the model spike trains for many different values of correlation of the underlying probability of firing. The analysis created graphs of neuron-neuron correlation for spike trains as a function of that for the underlying probability of spiking (e.g. Figure 6D). We fitted a regression line to the measurements, and used the line to estimate the neuron-neuron latency correlation for the underlying probability of spiking given that for the actual spike trains of the two neurons. We excluded 12 pairs of MT neurons where the regression analysis explained less than 40% of the variance in graphs like Figure 6D. Three pairs yielded estimates of latency correlation in the underlying probability of spiking that were slightly greater than 1: we set those values to one.

Neuron-neuron latency correlations in the underlying probability of spiking ranged from -0.5 to 1, as shown in the marginal histogram of Figure 6E. They averaged 0.29 and were about four times larger than the correlations for spike trains. Figure 6E suggests a structure in the neuron-neuron correlations, where pairs of neurons with more similar latencies had higher values of correlation. However, the small number of pairs in our sample made it impossible to specify an equation that would describe the structure, or to verify the statistical significance of the suggested structure.

It is noteworthy that the values of neuron-neuron latency correlation for the underlying probability of spiking are well below one for most pairs of neurons: therefore, there is considerable independent variation in the underlying probability of spiking. We conclude that the independent and shared components of neural latency variation do not map directly onto the variation caused by stochastic, irregular spike timing and shared inputs. Further, the low neuron-neuron latency correlations agree with our finding of relatively small neural-behavior latency correlations for the underlying probability to spiking. They imply that low neuron-neuron latency correlations, rather than downstream noise, probably cause the low values of neuron-behavior latency correlation.

A sensory source for variation in pursuit latency?

The analyses we have completed so far enable a computational analysis to quantify how much of the variation in pursuit latency can arise from correlated variation in neural latency in MT. We simulated a population of 1000 spiking MT neurons that mimicked i) the neuron-

neuron latency correlations between underlying probability of spiking, ii) the mean and variance of firing rate as a function of time, and iii) the latency variation and CV of our sample of MT neurons (see ***Methods***). It was critical to know the distribution of neuron-neuron correlations in *underlying probability of spiking* for this exercise (red curve in marginal histogram to Figure 6E), because there is no other way to create spike trains with correlated latencies other than starting from the underlying probability of firing. We ran the simulation 200 times to represent 200 trials of target motion. For each simulated trial, we used a linear decoder to pool the 1000 spike trains from the neurons in the population and create a population spike density function:

$$\dot{E}'(t) = \sum_{i=1,1000} G(\sigma=10ms) \otimes a_i \vec{S}_i(t) \quad (2)$$

where $\dot{E}'(t)$ is the simulated eye velocity as a function of time, a_i is the decoding weight for the i^{th} model neuron, G denotes a Gaussian function with σ=10 ms, and $\vec{S_i}(t)$ is the spike train. We estimated behavioral latency on each simulated trial using $\dot{E}'(t)$ the same way we had for the pursuit data.

Decoding the model MT population response could account for up to 7 ms of the 10 ms standard deviation of the actual latency of pursuit reported by Osborne et al. (2005). The exact prediction depended on our assumptions about neuron-neuron latency correlations. The standard deviation of the predicted latency was close to zero when we set the neuron-neuron latency correlations to the biologically unreasonable value of zero (Figure 6F, black filled diamond). The standard deviation of predicted latency increased to 7 ms as we increased the “delta-latency constant” of the structure in neuron-neuron latency correlations up to a value of 24 ms (the red curve in Figure 6E). The standard deviation of the predicted latency was 4 ms when we assumed a Gaussian distribution of otherwise unstructured neuron-neuron latency corrections across the MT population with a mean value equal to the mean of 0.29 in our data (Figure 6F, open circle). Doubling or halving the number of neurons in the simulation did not alter these conclusions.

The conclusions from Figure 6F depend to some degree on assumptions about how the MT population is decoded to drive eye velocity. Reasoned changes in the decoder’s weights could increase or decrease the predicted behavioral latency variation. For example, weighting each neuron in proportion to its neural latency variation or its mean neuron-neuron latency correlation with its neighbors increased the predicted variation of behavior latency. Weightings in inverse proportion of those parameters decreased the predicted variation in behavioral latency. The magnitude of these effects was fairly small, but could be enhanced by widening the range of decoding weights. At an extreme, driving pursuit only with neurons that lack neuron-neuron latency correlations with their neighbors would create the situation where MT contributes nothing to the variation in behavioral latency. We will return to this issue in the *Discussion*, but for now we suggest that the correlated latency variation in MT probably accounts for 40–70%, but not all, of the variation in pursuit latency.

Sensory latency variation and downstream multiplicative noise can account fully for the behavioral latency variation and neuron-behavior latency correlations

As a final step, we evaluated whether a source of variation downstream from MT could increase the variation in the pursuit latency predicted by our model. We needed to ask this question carefully. While downstream noise definitely would increase the predicted variation in pursuit latency, it would have the undesirable effect of reducing the magnitude of MT-pursuit latency correlations (Schoppik et al., 2008). In Figure 7, we search for a form of downstream noise that i) amplifies the latency variation emerging from MT to predict a standard deviation of pursuit latency of 10 ms and ii) predicts MT-pursuit latency correlations of ~0.28 for the underlying probability of firing.

We start with the model MT population that assumed a Gaussian structure of neuron-neuron latency correlations with a mean of 0.29, and equal decoding weights for all model neurons. Recall that this model predicted a standard deviation of pursuit latency of 4 ms (Figure 6F). We then explored the effects of altering the decoded signals with different kinds of noise. As before, we ran the simulation 200 times to simulate 200 behavioral trials, computed the mean population spike density function as a template, and then estimated the latency for each trial by shifting and scaling the template.

We created downstream noise using Gaussian functions G(μ,σ), with mean and standard deviation equal to μ and σ. The full model allowed both parameters of G(μ,σ) to vary and defined the trajectory of pursuit as:

$$P(t)=\eta \cdot \frac{\sum_i MT_i(t)}{N} \quad (3)$$

where $P(t)$ is the pursuit output, η is a random variable drawn from the distribution $G(\mu,\sigma)$, $MT_i(t)$ is the time-varying spike density function of the i^{th} model MT neuron and $N=1000$ model MT neurons. We allowed μ to vary from 1 to 2.5 and σ to vary from 0 to 0.5 and ran the model for the full distribution of both parameters. We computed the standard deviation of the decoded latency and the mean model neuron-behavior latency correlations for each of the pairs of parameter values.

A model with multiplicative downstream noise accounted for all the statistics of our data. In Figures 7B and C, the gray scale indicates how much the performance of the model deviated from that of the actual data as a function of μ and σ. Darker colors show parameters that matched the data better. The red circles indicate a set of parameters that matched the standard deviation of latency and the MT-pursuit correlations: μ=1.4 and σ=0.22. These parameters implement a noisy downstream gain control (Figure 7D).

Other downstream noise models performed less well. When we constrained μ to be 1.0 so that the model implemented multiplicative noise without a gain adjustment, the parameters that reproduced the standard deviation of pursuit latency underestimated the measured values of MT-pursuit latency correlations (Figure 7D). The model failed in exactly the same way when we changed it to implement additive noise:

$$P(t)=\eta+\frac{\sum_i MT_i(t)}{N} \quad (4)$$

We conclude that a specific model can reproduce the full statistics of pursuit latency, including its variation and the magnitude of MT-pursuit latency correlations. Latency variation would arise from correlated latency variation in MT and be amplified by a specific form of downstream noise. The downstream noise could be a second, independent source of variation in pursuit latency, or might be driven by the correlated variation in MT. In the latter scenario, the entire variation in pursuit latency would arise from MT.

Discussion

We have made two steps forward in understanding how a sensory code is created and converted into action. First, our analysis of neural responses at multiple levels of the pursuit circuit has provided the first complete account of the evolution of response latency through a sensory-motor circuit. Second, detailed analysis of the statistics of MT neuron responses has revealed potentially important new insights into neural codes for sensation and action. Our analysis suggests a change in the paradigm for understanding the origins of variation in neural responses and how that variation is (and is not) reduced through downstream decoding.

Pitfalls in measuring latency from spike trains

The challenges of estimating neural latency from noisy spike timing in single behavioral trials forced us to use approximations to estimate the trial-by-trial relationship between neural and behavioral latency. We partitioned the trials from a given neural recording session into 5 quintiles according to their latencies, and estimated the latency of each spike train iteratively using an objective, Bayesian approach. Comparison with the results of more standard data analysis approaches showed that our approach reduced the effects of measurement noise by 10–20%. Also, because our correlation analysis relies on computing a regression slope for z-scored data, it should be resilient against contamination by independent noise.

Our approach avoids the artifacts caused by tradeoffs between latency and magnitude. These would have contaminated our analysis if, for example, we had estimated response magnitude separately in fixed analysis windows (Lee et al., 2010). Other analysis methods have potential problems. The time when a neural response goes more than two standard deviations above the resting response depends on the rate of rise of the neural response. The time when the neural response reaches a fixed percentage of the peak response depends on the time and size of the peak. By scaling carefully-chosen templates and shifting them in time, we think we have obtained excellent estimates of the latency and amplitude of neural responses or eye velocity. Our approach measures latency independent of response amplitude, and it reproduces the known latency well in models of neural spike trains.

Independent versus shared variation in neural responses

Prior analyses of the amplitude of neural responses and their effect on behavioral output have broken the trial-by-trial variation of neural responses into two components. A “shared” component seems to be controlled by inputs that are distributed widely to a population of sensory neurons. It may arise from top-down influences such as general cortical excitability or attention (Goris et al., 2014) or from variation that arises early in sensory processing (Chaisanguanthum et al., 2014). The shared component is expressed as trial-by-trial correlations in the responses of pairs of neurons to repetitions of a given sensory stimulus. The shared component of variation dominos through the system and plays a major role in determining the magnitude of behavioral variation. An “independent” component is different in different neurons and therefore can be averaged away and does not have an impact on behavioral variation. At least tacitly, the field seems to attribute the shared component of variation to fluctuations in the integrated synaptic drive to a neuron, which we call “underlying probability of spiking”, and the independent component to the stochastic nature of spiking (Chaisanguanthum et al., 2014; Goris et al., 2014).

Our analysis suggests that a somewhat different framework holds, at least for neural and behavioral latency. In particular, we show that independent variation is present in the underlying probability of spiking. It cannot be attributed entirely, or even largely, to noise added by irregular spiking. Computational approaches allowed us to estimate the neuron-neuron and neuron-behavior latency correlations for the underlying probability of spiking, revealing values much lower than we would expect if the independent latency noise arose only from the stochastic nature of spiking. On average, the underlying probability of spiking for each MT neuron can explain only about 8% of pursuit latency variance. Similarly, neuron-neuron latency correlations for underlying spiking probabilities averaged only 0.29. If independent variation arose only from the stochastic nature of spiking, then we would expect higher values of correlation. The low value of neuron-neuron latency correlations for the underlying probability of spiking is particularly telling because it does not depend on assumptions about noise added downstream from MT.

The presence of independent latency variation in the underlying probability of spiking raises two questions: i) how does it get there, and ii) is this also true for the amplitude of firing rate? The latter question will have to be evaluated by future analyses. In regards to the former question, we suspect that limited convergence from processing stage to processing stage in the visual cortex is responsible for the appearance of independent variation in the underlying probability of spiking. Consider two extremes. If each MT neuron received its synaptic inputs from only one V1 neuron, then the independent variation of the V1 neuron would contribute substantial variation in the underlying probability of spiking. If each MT neuron received its synaptic inputs from many V1 neurons, then the independent variation in V1 spiking would average away, and we would not have found evidence for independent variation in the underlying probability of spiking. We suggest that the amount of convergence in the inputs to MT neurons is small enough so that independent variation is not averaged away. Alternatively, local cortical circuits could be responsible for the appearance of independent variation in the underlying probability of spiking.

The origin of behavioral response latency

Our analysis shows that neural latency determines behavioral latency, at least in the system we study. Three features of neural signal processing probably contribute to the increase in neuron-behavior latency correlations as signals pass from MT to the cerebellum to the brainstem. First, the cerebellum and brainstem almost certainly are downstream from any additional noise sources, such as the noisy gain control we suggest, so the trial-by-trial relationship between neural and behavioral latency should be stronger (Schoppik et al., 2008). Second, the independent noise contributed by the stochastic nature of spiking is considerably lower in the cerebellum compared to MT, and still lower in the brainstem oculomotor circuits (Ramachandran et al., 2006). Lower independent noise will lead to higher neuron-behavior latency correlations. Third, anatomical convergence may increase as signals pass into the motor system, reducing the amount of independent variation in the underlying probability of spiking and allowing neurons to signal the impending movement with higher fidelity.

Much of the current thinking about the creation of neural latencies is based on examples of integrating neural activity to a threshold. For example, Miles et al. (1986) explained the effect of stimulus contrast and temporal frequency on the latency of the ocular following response with a model that integrated different visual response amplitudes to create ramps to threshold at different rates. In complex tasks such as the countermanding task for saccades (Hanes and Schall, 1996), a gap saccade task (Dorris and Munoz, 1998), or a perceptual decision task (Roitman and Shadlen, 2002), neurons in the superior colliculus, and the parietal and frontal cortices show a ramp increase in neural activity prior to saccades. Saccades occur when the activity crosses a threshold. Our recordings do not fit with these formulations because we did not see any consistent relationship between neural response amplitudes and behavioral latencies. The regime used by pursuit could be different because our task minimizes cognitive influences and studies a highly reactive sensory-motor behavior.

Analysis of variation in pursuit behavior led to the hypothesis that sensory noise causes motor variation in pursuit (Osborne et al., 2005). Recordings from MT provided neural support for that hypothesis for pursuit direction and speed (Hohl et al., 2013; Lee and Lisberger, 2013). Until the present paper, however, the source of latency variation was a matter of speculation and the tools for measuring latency with the requisite precision did not exist. Now, we have shown that MT can account for 40–70% of the variation in pursuit latency. A noisy downstream gain element can account for the rest of the behavioral variation while remaining consistent with our data on neuron-neuron and neuron-behavior latency correlations. Previous studies have shown that FEF_{SEM} is involved in gain modulation of pursuit amplitude (Tanaka and Lisberger, 2001), and could be the gain control area for sensory estimates target speed and direction (Lee et al., 2013; Yang et al., 2012); perhaps FEF_{SEM} also provides a noisy gain control for pursuit latency. Further, the latency variation contributed by the noisy-gain control could be highly correlated with the latency variation in MT, in which case MT could be the sole contributor to pursuit latency variation. Or, latency variation may be independent in MT and FEF_{SEM}, in which case each would contribute about equally.

Our conclusions about the contribution of MT latencies to pursuit latency depend on the properties of the decoder that converts neural responses in MT into pursuit eye movements. We assume that pursuit latency arises as a side effect of pooling the population response in MT to estimate target direction and speed. Thus, we think of latency as an emergent property of a system that is designed to do something else, rather than as a parameter that the decoder is trying to optimize in any way. Our earlier papers imply that the decoder for pursuit is weighting each neuron according to its preferred speed and/or direction (Hohl et al., 2013). Thus, it seems unlikely that the decoder selects weights that would minimize the impact of MT latency variation on pursuit latency variation. Namely, we doubt that it weights neurons that lack neuron-neuron latency correlations with their neighbors more heavily. Thus, we think that the decoder used in our computational analysis probably is appropriate to simulate what actually happens in the nervous system.

Conclusions

Analysis of the neural code for behavioral latency has revealed that neural latency drives behavioral latency, at least in the reduced behavioral paradigms we use. The ramp-to-threshold regime does not seem to apply to the initiation of pursuit eye movements. Analysis of the correlations among the latencies of neural spiking, the underlying probability of spiking, and pursuit behavior reveals that neuron-neuron correlations in a sensory area make a large contribution to trial-to-trial variation in behavioral latency. Further, independent variation appears to have a strong presence in the underlying probability of spiking, and the stochastic nature of spiking makes an unexpectedly small contribution. These findings are based on analysis of a simple cortical sensory-motor behavior, but they may generalize to many neural systems.

Methods

We mined previously published data from 6 monkeys. The methods for acquiring the data and the experimental design were described before (Joshua et al., 2013; Lee and Lisberger, 2013; Medina and Lisberger, 2007). The data recorded during step-ramp pursuit behavior include 135 neurons recorded in extrastriate area MT, 40 Purkinje cells recorded in the floccular complex, 29 floccular target neurons recorded in the vestibular nucleus, and 40 Abducens neurons.

To analyze neural data, we combined responses to identical stimuli by aligning the spike trains on the onset of stimulus motion and summing the spikes from all trials to create peri-stimulus time histograms (PSTH). We measured neural response amplitudes after creating spike density functions by filtering PSTH’s with a Gaussian having a standard deviation of 10 ms, and normalizing so that the spike density function had the units of spikes/s. We computed averages of eye velocity as a function of time on a millisecond time scale.

Estimation of pursuit latency and relationship to neural latency

We used an objective procedure to quantify pursuit latency and gain for each trial. We computed the average horizontal and vertical eye velocity across trials and used the averages from 20 ms before to 100 ms after the initiation of pursuit as templates. We shifted and

scaled the templates to obtain the best least-squares fit to the eye velocities from each individual trial. The free parameters of the fits were time shift (Δt), horizontal scale (G_h), and vertical scale (G_v) The best fitting value of Δt defined pursuit latency in each trial. We included a trial in our analysis only if the best-fitting adjustment of the template accounted for more than 90% of the variance of the actual eye velocities. We proceeded to further analysis only if 50 or more trials passed this screen.

We determined the slope of the relationship between neural latency and pursuit latency by dividing the trials into five equal-sized groups (quintiles) according to the latency of pursuit in each trial. We computed the mean eye velocity and the mean spike density function for each quintile of trials. Then, we created templates by averaging the eye velocity and spike density functions for all the trials in the interval from 20 ms before to 100 ms after the response onset. We found the five values of time shift (Δt) and neural scaling factor (G) that provided the best least-squares fit of the spike density template to the average spike density functions for each quintile. We performed the same analysis using the eye velocity template to fit each quintile's average of eye velocity. The values of Δt and G for the spike density functions and the eye velocity averages defined the neural latency and response magnitude for each of the quintiles. They allowed us to compute the slopes of the relationships between pursuit latency and neural latency or response amplitude.

Calculation of neuron-behavior latency correlations

We estimated trial-by-trial correlations between neural latency and pursuit latency, by modifying a method from Bollimunta et al. (2007) to improve the estimates of neural latency through several iterations of Bayesian inference.

1. We created a spike density function for each individual trial by convolving each spike train with a Gaussian filter (σ = 10 ms), and we averaged across all trials to create a template.
2. We shifted and scaled the template to find the best fit, in the least-square sense, to the spike density function for each individual trial. This yielded a value of Δt and G for each trial, and defined latency and response gain.
3. We sorted the trials according to neural latency, created five equal-size groups (quintiles), and averaged across the spike density functions within each quintile to create five templates.
4. We made five new estimates of neural latency for each trial, one using each of the 5 templates, and defined the trial's neural latency according to the Δt for the template that provided the best fit.
5. We repeated the sorting, template creation, and latency estimation procedure (steps 4–5) to further reduce estimation errors for neural latency.
6. We shifted and scaled the original template to estimate the latencies of the average spike density for each of the five final quintiles.

7. We used the average latencies of the quintiles of trials to anchor each of the quintiles of a Gaussian function. The standard deviation of the best fitting Gaussian provided an estimate of the standard deviation of the measured neural latency.

The value of σ for the Gaussian derived in step 8 allows us to estimate neuron-behavior correlations for latency by taking advantage of the fact that Pearson's correlation coefficient is equal to the regression slope of data that has been converted to z-scores (Rodgers and Nicewander, 1988).

$$Z(L_{spikes}) = R \cdot Z(L_{eye}) \quad (5)$$

where $Z(x)$ is the z-score of x, and L_{spikes} and L_{eye} are the latencies of the neural response and pursuit, respectively. We can rewrite Equation (5) in way that allows us to estimate R from our data:

$$\frac{L_{spikes}}{\sigma(L_{spikes})} = R\frac{L_{eye}}{\sigma(L_{eye})} + k \quad (6)$$

where $\sigma(L_x)$ is the standard deviation of L_x. We do the analysis with trials grouped according to either the pursuit latency or the neural latency as the independent variable. If we define β_{n-b} and β_{b-n} as the regression slopes when neural latency or pursuit latency are the independent variables, then algebraic manipulation of Equation (6) leads to:

$$R = \frac{\sigma(L_{eye})}{\sigma(L_{spikes})}\beta_{b-n} = \frac{\sigma(L_{spikes})}{\sigma(L_{eye})}\beta_{n-b} \quad (7)$$

We estimated the regression coefficient twice, using neural latency and pursuit latency as the independent variables for grouping trials. In principal, the two estimates should be the same. In practice, they are close with each other, but not identical. We use their geometric average as an estimate of the trial-by-trial correlation between given neuron's latency and pursuit latency (Rodgers and Nicewander, 1988):

$$R = \pm\sqrt{\beta_{b-n} \cdot \beta_{n-b}} \quad (8)$$

We used an analogous procedure to estimate the correlation between the latencies of responses of 40 pairs of simultaneously recorded neurons, estimating the latency of the responses of each neuron, rather than of one neuron's neural responses and pursuit.

Underlying probability of spiking

We derived a relationship that allows us to estimate the latency correlation between underlying spiking probability and pursuit from parameters that we can measure. Variation in neural latency has two components: one in the underlying probability of spiking and one in stochastic nature of spike timing. The latter will be independent across neurons, while the former may have a component that is shared across neurons. We define these two components of variation as:

$$Var(L_{spikes}) = Var(L_{prob}) + Var(L_{ind}) \quad (9)$$

where L_{spikes}, L_{prob}, and L_{ind} are the neural latencies of spike trains, the underlying probability of spiking, and the stochastic nature of spike timing. We define the MT-pursuit latency correlation we measure from spikes as:

$$R = R_{spikes,eye} = \frac{\mathrm{cov}(L_{spikes}, L_{eye})}{\sqrt{\mathrm{var}(L_{spikes})\mathrm{var}(L_{eye})}} = \frac{\mathrm{cov}(L_{prob} + L_{ind}, L_{eye})}{\sqrt{\mathrm{var}(L_{spikes})\mathrm{var}(L_{eye})}} \quad (10)$$

We can simplify Equation (10) because L_{ind} is independent from L_{eye}:

$$R_{spikes,eye} \approx \frac{\mathrm{cov}(L_{prob}, L_{eye})}{\sqrt{\mathrm{var}(L_{spikes})\mathrm{var}(L_{eye})}} \quad (11)$$

$$R_{prob,eye} = \frac{\mathrm{cov}(L_{prob}, L_{eye})}{\sqrt{\mathrm{var}(L_{prob})\mathrm{var}(L_{eye})}} \approx R_{spikes,eye} \sqrt{\frac{\mathrm{var}(L_{spikes})\mathrm{var}(L_{eye})}{\mathrm{var}(L_{prob})\mathrm{var}(L_{eye})}} \quad (12)$$

Finally, we can rewrite the right side of Equation (12) as:

$$R_{prob,eye} \approx R_{spikes,eye} \frac{\sigma(L_{spikes})}{\sigma(L_{prob})} \quad (13)$$

Equation (13) estimates the MT-pursuit correlation for the underlying probability of spiking from the values of parameters we can measure. Next steps appear at the relevant site in the ***Results***.

Computer simulations

Our computer simulations created model MT neurons with realistic statistics based on our data. We used the time-varying mean spike density function of our population of MT neurons, or of individual neurons, to define the underlying probability of spiking. For simulations of large populations, we defined each model neuron's mean latency by drawing from the distribution of latencies in our data. We also drew the standard deviation of latency for each neuron's firing probability from a Gamma distribution fitted to the distribution of latency variation for the underlying probability of spiking in our data:

$$\sigma(L_p) \sim \Gamma(k_\sigma, \theta_\sigma) \quad (14)$$

where $\sigma(L_p)$ is the standard deviation of neuron p's firing probability latency, $k_\sigma = 0.73$ is the shape parameter of the gamma distribution, and $\theta_\sigma = 12.3$ is the scale parameter. We then created 100 or 200 simulated trials for each neuron by drawing a latency from Gaussian distributions having the standard deviation assigned to each neuron by Equation (14) and shifting the underlying probability of spiking in time by that amount. We created different

structures of neuron-neuron correlations by varying the correlation matrix used in a Cholesky decomposition:

1. To create model populations that lacked neuron-neuron latency correlations, we used a matrix with all entries as zeros.
2. To create model populations with uniform neuron-neuron latency correlations that matched the mean and standard deviation in our data, we used a matrix with entries chosen randomly from a Gaussian distribution with a mean of 0.29 and standard deviation of 0.24.
3. To create model populations with structured neuron-neuron correlations that depended on the difference in latency between the two neurons of a pair, we defined entries in the matrix according to the exponential:

$$r_{i,j}=r_{\max}e^{-\frac{\left(L_i-L_j\right)^2}{\tau^2}} \quad (15)$$

where $r_{i,j}$ is the correlation between neuron *i* and *j*, L_i and L_j are the mean neural latencies of neurons *i* and *j*, r_{max} is the maximum size of the correlations, and τ is the "delta-latency" constant of the exponential function. Based on our data, we set r_{max} to be one.

At this point, we had defined the latency, amplitude, and trajectory of the underlying probability of firing either for 100 trials of a single model neuron, or for 200 trials of 1,000 model neurons. We modeled the spike timing for each trial as a renewal gamma process whose parameters for the inter-spike interval distribution were selected to match the coefficient of variation (CV) of a given neuron. Then, using the underlying probability of spiking estimated from the given neuron as the rate function, we made spike trains by rescaling the time of the gamma process with a given time-dependent trajectory of the underlying probability of spiking (Brown et al., 2002; Koyama and Shinomoto, 2005). The final step in the creation of model neural responses was to measure the CV of the model spike trains and compute the time-varying mean spike density function exactly as we had in our analysis of the actual data, to verify that our model neurons reproduced the statistics of the actual data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Jeff Beck for helpful comments, and K. MacLeod, E. Montgomery, S. Tokiyama, S. Ruffner, D. Kleinhesselink, D. Wolfgang-Kimball, D. Floyd, S. Happel and K. McGary for technical assistance. Research supported by the Howard Hughes Medical Institute and NIH grant EY03878.

References

Beck JM, Ma WJ, Pitkow X, Latham PE, Pouget A. Not noisy, just wrong: the role of suboptimal inference in behavioral variability. Neuron. 2012; 74:30–39. [PubMed: 22500627]

Bell AH, Meredith MA, Van Opstal AJ, Munoz DP. Stimulus intensity modifies saccadic reaction time and visual response latency in the superior colliculus. Exp Brain Res. 2006; 174:53–59. [PubMed: 16528494]

Bollimunta A, Knuth KH, Ding M. Trial-by-trial estimation of amplitude and latency variability in neuronal spike trains. J Neurosci Methods. 2007; 160:163–170. [PubMed: 17000007]

Brown EN, Barbieri R, Ventura V, Kass RE, Frank LM. The time-rescaling theorem and its application to neural spike train data analysis. Neural Comput. 2002; 14:325–346. [PubMed: 11802915]

Celesia GG, Puletti F. Auditory input to the human cortex during states of drowsiness and surgical anesthesia. Electroen Clin Neuro. 1971; 31:603–609.

Chaisanguanthum KS, Joshua M, Medina JF, Bialek W, Lisberger SG. The neural code for motor control in the cerebellum and oculomotor brainstem. eNeuro. 2014; 1 ENEURO.0004–14.2014.

Clark VP, Fan S, Hillyard SA. Identification of early visual evoked potential generators by retinotopic and topographic analyses. Hum Brain Mapp. 1995; 2:170–187.

Dorris MC, Munoz DP. Saccadic probability influences motor preparation signals and time to saccadic initiation. J Neurosci. 1998; 18:7015–7026. [PubMed: 9712670]

Gawne TJ, Kjaer TW, Richmond BJ. Latency: another potential code for feature binding in striate cortex. J Neurophysiol. 1996; 76:1356–1360. [PubMed: 8871243]

Gold JI, Shadlen MN. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. Neuron. 2002; 36:299–308. [PubMed: 12383783]

Goris RLT, Movshon JA, Simoncelli EP. Partitioning neuronal variability. Nat Neurosci. 2014; 17:858–865. [PubMed: 24777419]

Hanes DP, Schall JD. Neural control of voluntary movement initiation. Science. 1996; 274:427–430. [PubMed: 8832893]

Herrington TM, Assad JA. Neural activity in the middle temporal area and al intraparietal area during endogenously cued shifts of attention. J Neurosci. 2009; 29:14160–14176. [PubMed: 19906965]

Hikosaka K, Watanabe M. Delay activity of orbital and al prefrontal neurons of the monkey varying with different rewards. Cereb Cortex. 2000; 10:263–271. [PubMed: 10731221]

Hohl SS, Chaisanguanthum KS, Lisberger SG. Sensory population decoding for visually guided movements. Neuron. 2013; 79:167–179. [PubMed: 23849202]

Hollerman JR, Tremblay L, Tremblay L, Schultz W. Influence of reward expectation on behavior-related neuronal activity in primate striatum. J Neurophysiol. 1998; 80:947–963. [PubMed: 9705481]

Huang X, Lisberger SG. 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. 2009; 101:3012–3030. [PubMed: 19321645]

Joshua M, Medina JF, Lisberger SG. Diversity of neural responses in the brainstem during smooth pursuit eye movements constrains the circuit mechanisms of neural integration. J Neurosci. 2013; 33:6633–6647. [PubMed: 23575860]

Joshua M, Lisberger SG. A framework for using signal, noise, and variation to determine whether the brain controls movement synergies or single muscles. J Neurophysiol. 2014; 111:733–745. [PubMed: 24259547]

Koyama S, Shinomoto S. Empirical Bayes interpretations of random point events. J Phys a: Math Gen. 2005; 38:L531–L537.

Lee J, Kim H, Lee C. Trial-to-trial variability of spike responses of V1 and saccadic response time. J Neurophysiol. 2010; 104(5):2556–2572. [PubMed: 20810695]

Lee J, Lisberger SG. Gamma synchrony predicts neuron-neuron correlations and correlations with motor behavior in extrastriate visual area MT. J Neurosci. 2013; 33:19677–19688. [PubMed: 24336731]

Lee J, Williford T, Maunsell JHR. Spatial attention and the latency of neuronal responses in macaque area V4. J Neurosci. 2007; 27:9632–9637. [PubMed: 17804623]

Lee J, Yang J, Lisberger SG. Control of the gain of visual-motor transmission occurs in visual coordinates for smooth pursuit eye movements. J Neurosci. 2013; 33:9420–9430. [PubMed: 23719810]

Lisberger SG, Westbrook LE. Properties of visual inputs that initiate horizontal smooth pursuit eye movements in monkeys. J Neurosci. 1985; 5:1662–1673. [PubMed: 4009252]

Medina JF, Lisberger SG. Variation, signal, and noise in cerebellar sensory-motor processing for smooth-pursuit eye movements. J Neurosci. 2007; 27:6832–6842. [PubMed: 17581971]

Metzger RR, Greene NT, Porter KK, Groh JM. Effects of reward and behavioral context on neural activity in the primate inferior colliculus. J Neurosci. 2006; 26:7468–7476. [PubMed: 16837595]

Miles FA, Kawano K, Optican LM. Short-latency ocular following responses of monkey. I. Dependence on temporospatial properties of visual input. J Neurophysiol. 1986; 56:1321–1354. [PubMed: 3794772]

Osborne LC, Lisberger SG, Bialek W. A sensory source for motor variation. Nature. 2005; 437:412–416. [PubMed: 16163357]

Ramachandran R, Lisberger SG. Transformation of vestibular signals into motor commands in the vestibuloocular reflex pathways of monkeys. J Neurophysiol. 2006; 96:1061–1074. [PubMed: 16760348]

Rashbass C. The relationship between saccadic and smooth tracking eye movements. J Physiol. 1961; 159:326–338. [PubMed: 14490422]

Rodgers JL, Nicewander WA. Thirteen ways to look at the correlation coefficient. Am Stat. 1988; 42:59–66.

Roitman JD, Shadlen MN. Response of neurons in the al intraparietal area during a combined visual discrimination reaction time task. J Neurosci. 2002; 22:9475–9489. [PubMed: 12417672]

Shelton J, Kumar GP. Comparison between auditory and visual simple reaction times. Neuroscience & Medicine. 2010

Sundberg KA, Mitchell JF, Gawne TJ, Reynolds JH. Attention influences single unit and local field potential response latencies in visual cortical area V4. J Neurosci. 2012; 32:16040–16050. [PubMed: 23136440]

Takikawa Y, Kawagoe R, Itoh H, Nakahara H, Hikosaka O. Modulation of saccadic eye movements by predicted reward outcome. Exp Brain Res. 2002; 142:284–291. [PubMed: 11807582]

Tanaka M, Lisberger SG. Regulation of the gain of visually guided smooth-pursuit eye movements by frontal cortex. Nature. 2001; 409:191–194. [PubMed: 11196642]

Warzecha A, Egelhaaf M. Response latency of a motion-sensitive neuron in the fly visual system: dependence on stimulus parameters and physiological conditions. Vision Res. 2000; 40:2973–2983. [PubMed: 11000395]

Yang J, Lee J, Lisberger SG. The interaction of bayesian priors and sensory data and its neural circuit implementation in visually guided movement. J Neurosci. 2012; 32:17632–17645. [PubMed: 23223286]

Zohary E, Shadlen MN, Newsome WT. Correlated neuronal discharge rate and its implications for psychophysical performance. Nature. 1994; 370:140–143. [PubMed: 8022482]

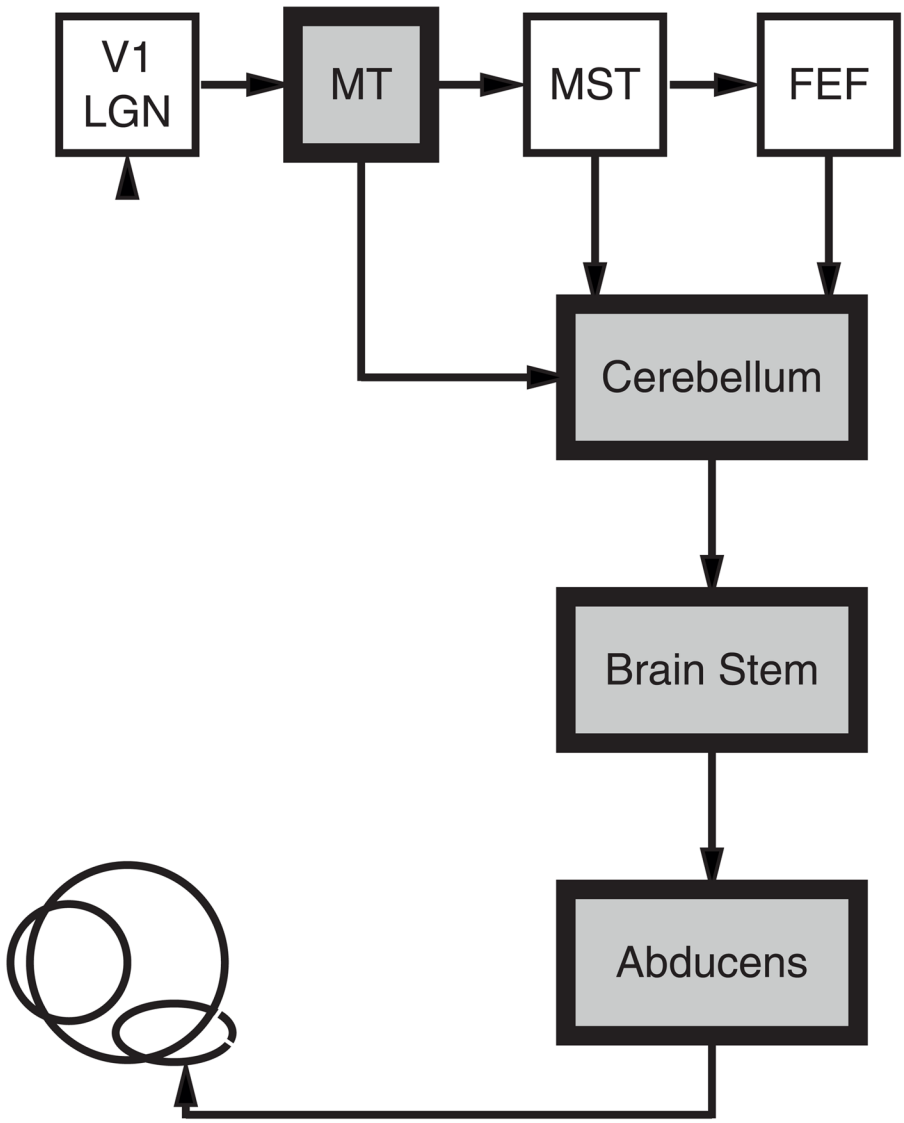

Figure 1. Neural circuit for smooth pursuit eye movement

Areas marked by gray shading indicate the sites where we recorded neural responses. Abbreviations are: V1, primary visual cortex; LGN, lateral geniculate nucleus, MT, middle temporal visual area; MST, medial superior temporal visual area; FEF, smooth eye movement region of the frontal eye fields.

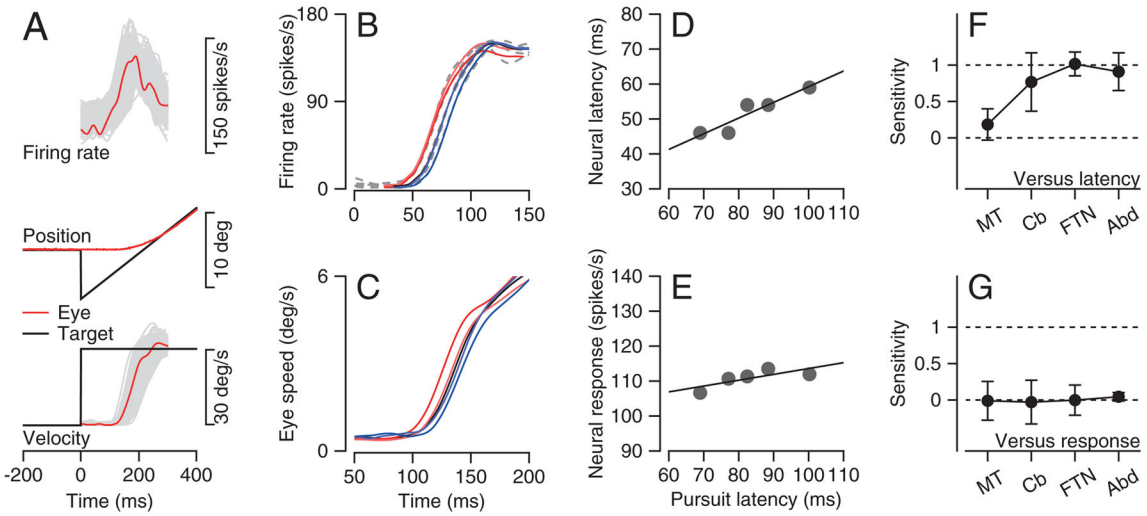

Figure 2. Sensitivity of pursuit latency to neural latency and response magnitude at 4 levels of the pursuit circuit

A. Neural responses during step-ramp target motion. From top to bottom, the traces show firing rate, eye and target position, and eye and target velocity. Black, red, and gray traces show target motion, mean responses, and trial-by-trial variation for a single recording session. **B, C**: Average spike density functions (B) and eye speed trajectories (C) for an example MT neuron. The five traces show averages for 5 quintiles of trials divided according to pursuit latency. Colors identify the data from the same groups in the eye velocity and firing rate traces. **D, E**: Regression analysis of neural latency (D) and response amplitude (E) versus pursuit latency. Symbols show averages for the 5 quintiles of trials sorted by latency, and lines show the results of linear regression. **F, G**: Mean and standard deviation of pursuit latency's sensitivity to neural latency (**F**) and response amplitude (**G**) in 4 recording sites: n=135, 40, 29, and 40 for recordings in MT, floccular complex, FTNs, and Abducens neurons.

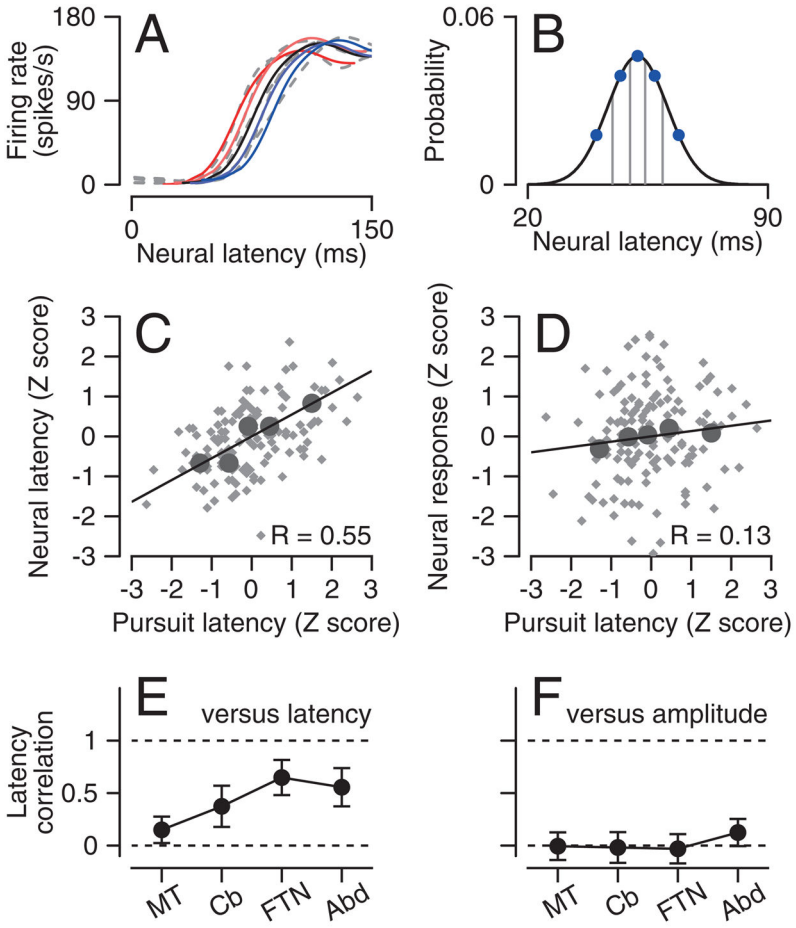

Figure 3. Neuron-behavior correlations for neural response latency or amplitude

A: Colored solid lines show the scaled and translated templates that provided the best fits to the dashed gray lines, which show averages of the five quintiles of responses with different pursuit latencies. **B**: Curve shows the Gaussian estimates of the distribution of latency for the neuron and blue symbols show the data used to derive the curve. **C, D**: Scatter plots one neuron showing z-scores of neural latency (**C**) and response amplitude (**D**) versus z-scores of pursuit latency. Gray and black symbols show estimates for all individual trials and averages for each of the 5 quintiles. Lines show the results of regression analysis. **E, F**: Population neuron-behavior correlations at each recording site for neural latency (**E**) and neural response amplitude (**F**). Error bars are standard deviations across neurons.

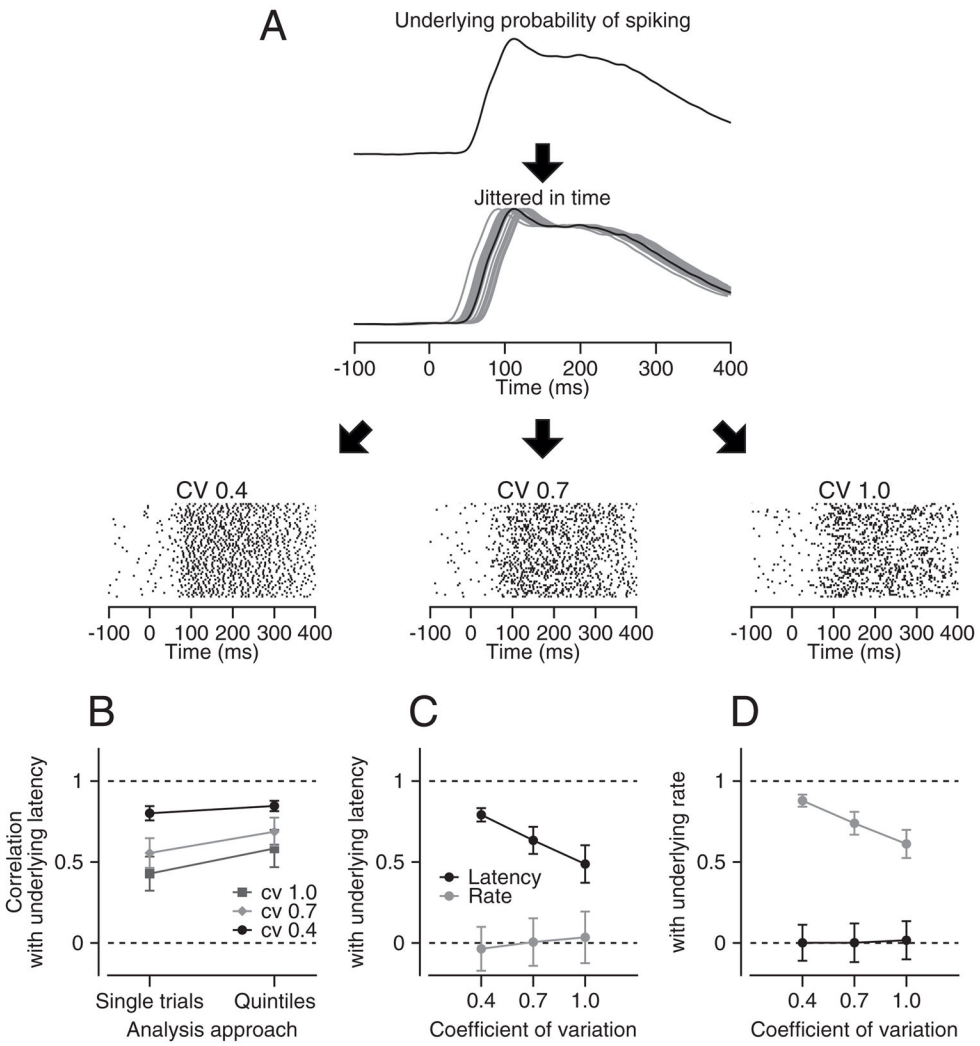

Figure 4. Procedure for evaluating accuracy of neural latency estimates

A: schematic showing how we simulated spike trains from by the underlying probability of firing. The rasters show simulations of 100 trials for 3 coefficients of variation (CV). **B**: Comparison of analysis done on single trials versus on trials divided into quintiles. **C, D**: Analysis of how well the analysis procedure could separate variation in latency versus amplitude of the underlying probability of firing. The y-axes plot the correlation between the actual latency (C) or the underlying probability of firing (D) with the latency or rate measured from the spike trains. Error bars are standard deviations from 100 repeats.

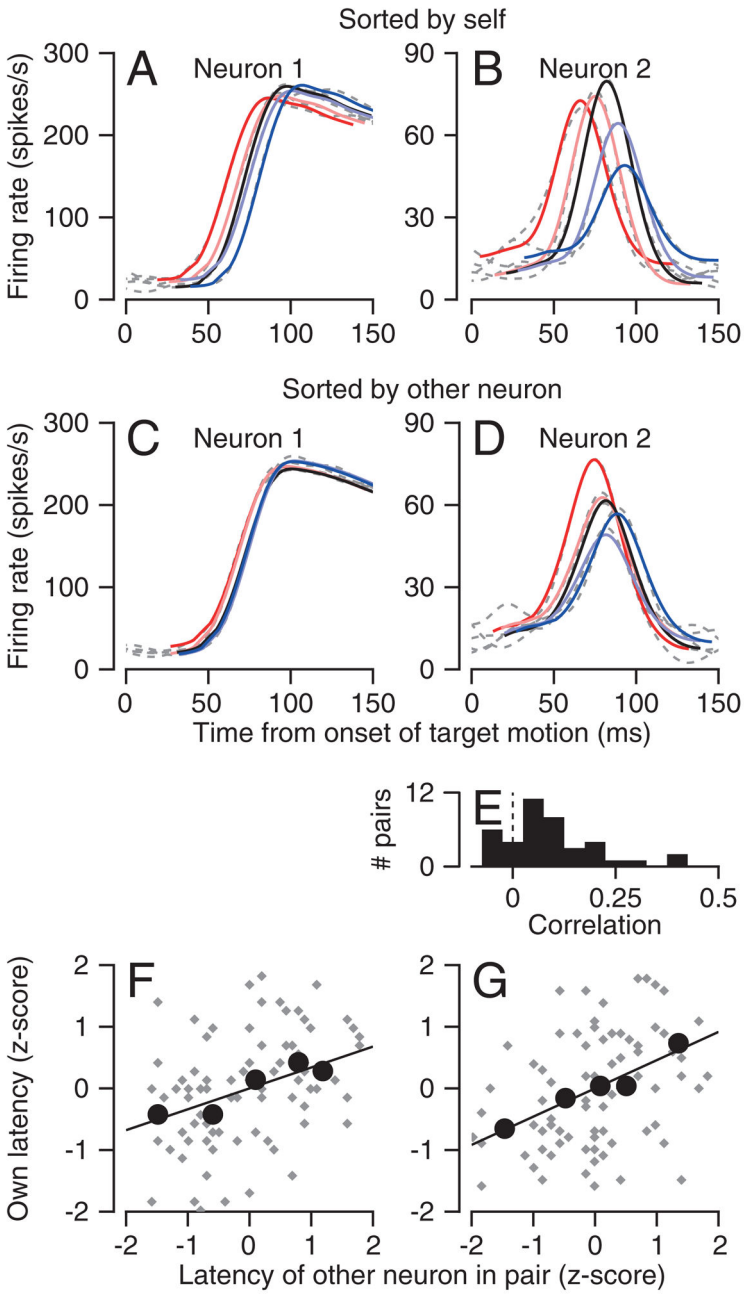

Figure 5. Neuron-neuron correlation for latency in area MT

A, B: Quintiles of spike density functions of the quintiles of two neurons recorded at the same time. **C, D**: Spike density functions of the same neurons, sorted according to neural response latency of the other neuron in the pair. In **A–D**, the colors and dashed gray traces show the scaled and translated templates and the averages of spike density. **E**: Distribution of neuron-neuron latency correlations in MT. **F, G**: Scatter plots showing the z-score of the latency of one neuron as a function of the z-score of the latency of the other neuron in the pair. Gray and black symbols show data for single trials and averages across the 5 quintiles.

Lines show the result of linear regression. The two graphs plot analysis of a single data set performed separately on the basis of the latency of each neuron.

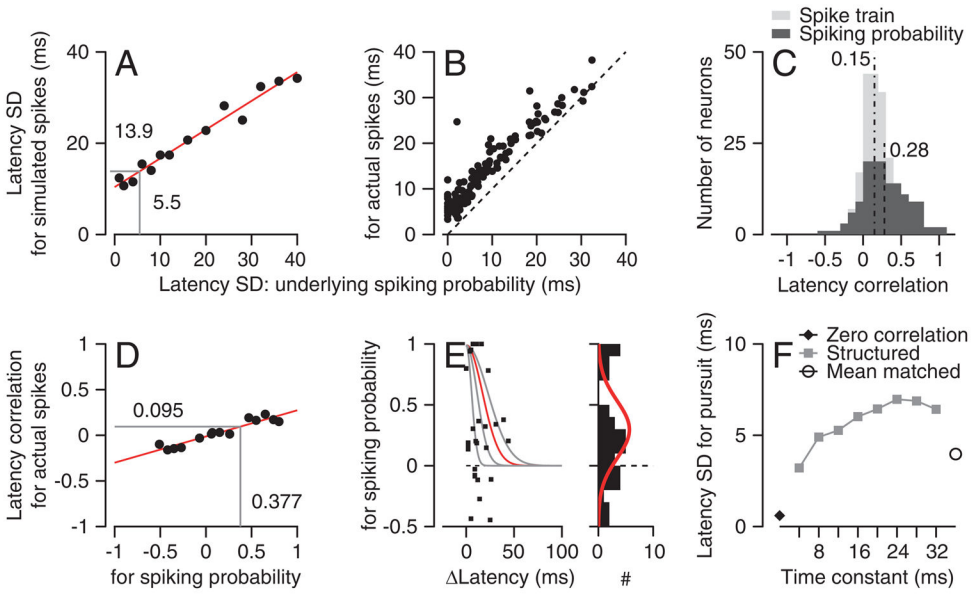

Figure 6. Effect of neuron-neuron latency correlations on the variation of behavioral latency

A: Relationship between the standard deviation of latency for simulated spikes and the underlying probability of spiking in a model MT neuron. **B**: Symbols show the relationship between standard deviation of latency measured from actual spike trains and for the underlying probability of spiking for a sample of MT neurons. Dashed line shows equality line. **C**: Distribution of MT-pursuit latency correlations. Black and gray histograms show values for actual spike trains and for the underlying probability of spiking. Vertical dashed lines show the population means. **D**: Relationship between the neuron-neuron latency correlation for simulated spikes and the underlying probability of spiking in a pair of model MT neurons. **E**: Each symbol shows neuron-neuron correlations in the underlying probability of spiking for a pair of neurons as a function of the latency difference between the two neurons. The gray and red exponential functions show potential descriptions of the data with different “delta-latency” constants. The marginal histogram summarizes the distribution of neuron-neuron latency correlations in the underlying probability of spiking. The red curve shows a Gaussian fit used to create model populations. **F**: Latency variation of pursuit under different assumptions about neuron-neuron latency correlations in a model population response. Filled diamond, neuron-neuron latency correlations were absent; gray symbols, structured correlations with different “delta-latency” time constants; open symbol, uniform neuron-neuron latency correlations with the same mean value as the data.

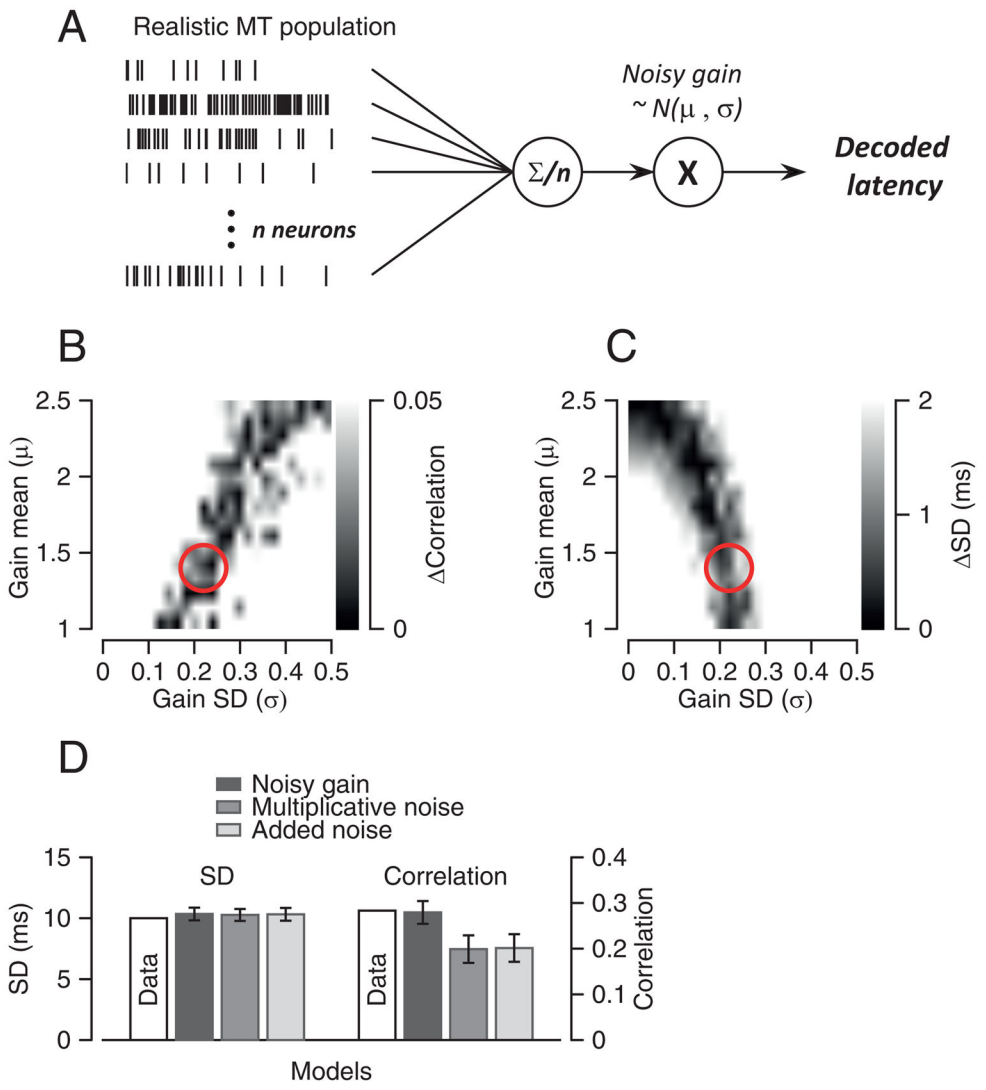

Figure 7. Model that uses downstream noise to account for statistics of pursuit latency

A: Schematic diagram of a model that uses a realistic model MT population response, an averaging population decoder, and noisy gain downstream from decoding. **B, C**: Gray scale representation of the difference between simulated and actual MT-pursuit correlations (B) and latency standard deviation (C) as a function of the value of downstream gain (y-axis) and the value of the downstream noise SD (x-axis). Red circles show the parameters that provide the best prediction of the actual data for both parameters. **D**: Comparisons among models. Error bars are standard deviations obtained from running each simulation 100 times.