Methods

Behavioral Methods. Before recording began, both monkeys were well-trained on the location-scene association task. Each trial started with the animal fixating a central fixation spot for 300 ms. The monkey then saw 4 identical visual targets superimposed on a complex color visual scene for 500 ms. Following a 700 ms delay interval, the fixation spot disappeared cueing the animal to make an eye movement to one of the 4 visual targets. Fixation was required for the duration of the scene and delay periods of the task. For each visual scene, only one of the targets was associated with reward and the rewarded targets were counterbalanced across novel scenes. Each new scene always had a rewarded target location different from that of the other new scenes in the set. If 4 new scenes were presented on a given day, the rewarded target locations for the set of new scenes on the next day was chosen randomly. If fewer than 4 new scenes were presented on a given day, we included the remaining rewarded target direction/s with the set of new scenes presented on the following day. To control for neural activity associated with particular eye movements or particular rewarded target locations, animals were always presented with a random mix of new scenes and 2-4 highly familiar "reference" scenes. Each of the 4 reference scenes was associated with a different one of the four possible rewarded target locations. Because new and reference scenes were always presented in a randomly intermixed fashion, the reference scenes could not be used as a cue for the correct location of any given new scene. Performance on reference scenes was always at or near 100% correct. Over the course of the session animals eventually learned, through trial and error, which target location was associated with each novel scene.

Neither monkey exhibited a spatial response bias during the training or the recording sessions. Instead, they responded by trial and error to each new scene until a correct response was made. We did not give animals explicit correction trials (i.e., immediate repetition of an incorrectly performed trial), though if animals were having a particularly difficult time learning a new scene, we adjusted the relative proportion of scenes presented such that the problematic scene was presented more often.

Electrophysiological Recording. Recordings were done in two adult male rhesus monkeys (Macaca mulatta) using standard electrophysiological recording techniques (single coated tungsten sharp electrodes purchased from Fredrick Haer & Co and on-line template matching spike sorting system made by Alpha Omega Engineering, Israel). The recording chambers were placed stereotaxically using magnetic resonance images of the brains of each monkey. To determine the location of the recording sites, we first measured the distance from the dorsal surface of the brain to the bottom of the brain directly using a thin microelectrode probe. We then used the MRI images from each animal to calculate the distance from the bottom of the brain to the dorsal and ventral limits of the hippocampus. All of our recording sites fell within the hippocampus for both animals. We made no attempt to select neurons based on their firing patterns. Instead, we collected a data set for the first well-isolated neuron encountered in the hippocampus.

Data Analysis

A. Behavioral Learning Curve Analysis. We estimated whether learning had taken place and at which trial it occurred by a two-step process. For each scene presentation,

the data took the form of a sequence of zeros and ones where a zero denoted an incorrect response and a one denoted a correct response. The initial part of this binary series, or response profile, was usually a sequence of incorrect responses (zeros). As the experiment progressed it contained more and more correct responses (ones).

As a first step in our analysis, we classified a scene as learned if for that scene the monkey obtained 7 correct consecutive responses i.e. if the binary series contained a sub-sequence of seven or more consecutive ones. We show that the probability of this occurring by chance is very small, as follows. We wish to find the probability of 7 or more correct responses in n trials with a probability of a correct response of $1/4$ and a probability of an incorrect response of $3/4$. We compute this by summing the probabilities of sequences of length 7 or more with starting points shifted incrementally from trials 1 to $n-6$. Starting at trial 1, the probability of a sub-sequence of 7 ones starting at trial 1 with any combination of zeros and ones in the $n-7$ subsequent trials is $(1/4)^7$. We then compute the probability of a sub-sequence of ones starting at trial 2 with a zero at trial 1 and any combination of zeros and ones occurring after the 7 correct responses. This is $\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^7$. The probabilities of the sub-sequence starting at trials 3 to 8 are the same i.e. $\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^7$. At trial 9, in order to define disjoint events, we excluded the probability that there was a sub-sequence starting at trial 1. Thus, the probability of a sub-sequence starting at trial 9 without 7 consecutive ones occurring at earlier trials is $\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^7 \times \left(1-\left(\frac{1}{4}\right)^7\right)$. At trial 10, we exclude the events of a sub-sequence starting at 1

and 2, so the probability is $\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^7 \times \left(1-\left(\frac{1}{4}\right)^7-\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^7\right)$. This pattern continues until we reach the last possible position for 7 consecutive ones at trial $n-6$. We computed recursively the total probability of 7 or more ones appearing anywhere in the sequence as the sum of these probabilities. Assuming the number of trials ranged from 25 to 60, we found that the probability of 7 or more correct responses occurring consecutively varied from 0.0009 to a maximum of 0.0025. Since these values are small, we can conclude that getting 7 consecutive correct responses is highly unlikely to occur by chance. Only cells corresponding to scenes reaching this behavioral criterion were further analyzed.

The second step was to identify the trial where learning occurred. We analyzed the data series using a state-space approach (Fahrmeir and Tutz, 1994; Kitagawa and Gersh, 1996) requiring a forward filter (Brown et al., 1998; Smith and Brown, 2003) and a fixed interval smoothing algorithm (Mendel, 1995), which enabled us to compute the probability, p_k, of a correct response as a function of trial number, k. Using the logistic transform to ensure probability values remained between 0 and 1, we wrote the model

$$p_k = \frac{\exp(-1.0986 + x_k)}{1+\exp(-1.0986 + x_k)} \tag{A.1}$$

where the constant -1.0986 ensured that when $x_0 = 0$ at the initial trial, $p_0 = 0.25$, the probability of a correct response by chance. Here, x_k was an unobserved state variable modeled by the random walk

$$x_k = x_{k-1} + \varepsilon_k \tag{A.2}$$

where ε_k are independent Gaussian random variables with mean 0 and variance σ_ε^2. The parameter σ_ε^2 governed the algorithm's learning rate; increasing σ_ε^2 increased the learning rate and vice versa. This particular formulation allowed us to specify the probability of a chance correct answer and allowed us to vary smoothly the probability of a correct response away from chance through the magnitude of the variance of the state variable. We relate an animal's observed responses to the probability of a correct response using the local Bernoulli model

$$p(dN(k) \mid \text{history}) = p_k^{dN(k)}(1-p_k)^{1-dN(k)} \tag{A.3}$$

where $dN(k)$ is one or zero depending whether there is a correct or incorrect response, respectively, at trial k.

We estimated the state variable across the trials using a forward filter and then smoothing the result with fixed interval smoothing. As an initial condition, we set $x_0 = 0$, ensuring that $p_0 = 0.25$. Details of the derivation of forward filter algorithm for the state variable in the context of an arbitrary point process model of neural spiking dynamics are given in Smith and Brown (2003). To smooth the results, we used a fixed interval smoothing algorithm (Mendel, 1995).

We made our choice of value of the parameter σ_ε^2 based on analysis of two randomly selected series of trial responses. For our analysis of the two data series, we chose initial values for σ_ε^2 in the range between 0.01 and 0.5. The smoothing algorithm fits were robust to initial values of $\sigma_\varepsilon^2 > 0.16$. We chose $\sigma_\varepsilon^2 = 0.36$. At this stage, we fixed the parameter and analyzed all the data sets.

We computed the probability of a correct response from the smoothed state variable using the standard change of variables formula from probability theory. Namely, approximating the probability density of the state at time k as the Gaussian density with mean $\hat{x}_k$ and variance, $\hat{\sigma}_k^2$, it follows from Eq. A.1 that the probability density, $f_k(u)$, for the probability of correct answer at time k is defined by

$$f_k(u \mid \mu, \hat{x}_k, \hat{\sigma}_k^2) = \frac{1}{\sqrt{2\pi\hat{\sigma}_k^2}} \frac{1}{u(1-u)} \exp\left(-\frac{1}{2} \left[\frac{\log\left[\frac{u}{(1-u)\exp(-1.0986)} \right] - \hat{x}_k}{\hat{\sigma}_k} \right]^2 \right). \qquad \text{(A.4)}$$

The confidence limits for the probability are computed by numerical integration of Eq. A.4. In our analyses we computed 95% confidence intervals. We considered the animal to have learned the task when the lower limit of the 95% confidence interval for the probability of correct response was greater than 0.25.

Example. Figure S1 below shows the results for our method applied to simulated data. Circles on the x-axis indicate a trial with a correct response. We show the smoothed estimate (red line) of probability of correct response with 95% confidence bounds (dashed lines red lines). We define the trial where learning has occurred as the one for which the lower limit of the 95% confidence interval first exceeds 0.25. Thus, the animal learns at trial 16, consistent with the beginning of the string of correct responses.

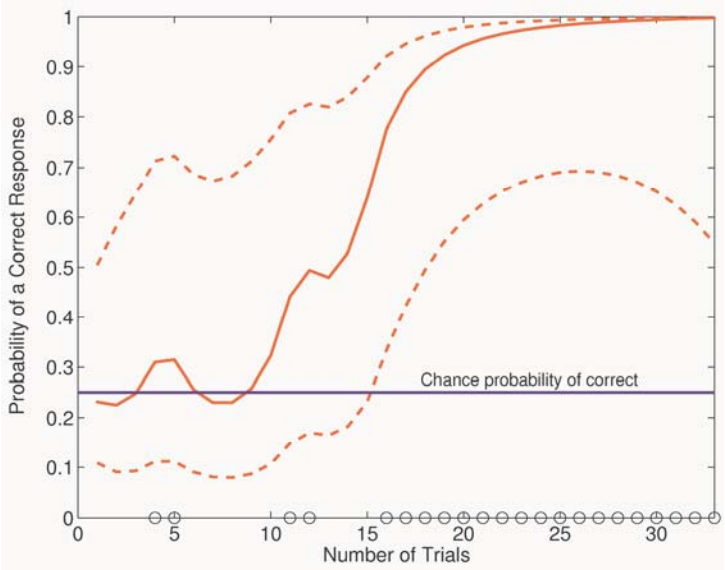

Figure S1. Example of estimated of learning curve (red line) applied to simulated data (positive responses are shown as circles on the x-axis). We show the corresponding 95% confidence limits (dashed lines) and the probability of a correct response occurring by chance (horizontal blue line). We assume the animal learned when the lower confidence limit crossed the chance probability line i.e. the animal learns here at trial 16.

B. Neural Activity. To track the neural activity for each cell-scene combination, we fit a dynamic firing rate function, based on the construction of an adaptive point process filter algorithm (Brown et al., 2001; Frank et al., 2002). We initially used only the raw firing rate data, but we found that these data were very noisy and, due to the inherent variability of neural spiking, could be inaccurate estimates of the true firing rate. For each set of trials corresponding to a given scene, we constructed a cardinal spline curve defining $\lambda(t|\theta)$, the firing rate, as a function of a set of control points θ at time *t*. This spline

model, which had been used previously to track receptive field plasticity in rat hippocampal and entorhinal cortex neurons (Brown et al., 2001; Frank et al., 2002), allowed a flexible form of the rate function to be estimated for each neuron. The control points for the model each had a coordinate defined by a time into the trial and a corresponding firing rate, and were spaced at 50 ms intervals along the time axis from t_1 = -50 ms (50 ms before the fixation time for the trial) to t_{83} = 4050 ms (a time 50 ms beyond the time at the end of the longest trial). The time locations and associated magnitudes for the control points of the spline were thus defined as the pairs: $\{(t_j, \theta_j)\}_{j=1}^{83}$. $\lambda(t \mid \theta)$ is then given by:

$$\lambda(t \mid \theta) = [v(t)^3 \quad v(t)^2 \quad v(t) \quad 1] \begin{bmatrix} -0.5 & 1.5 & -1.5 & 0.5 \\ 1 & -2.5 & 2 & -0.5 \\ -0.5 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_{j-1} \\ \theta_j \\ \theta_{j+1} \\ \theta_{j+2} \end{bmatrix} \tag{B.1}$$

for $t \in (t_j, t_{j+1}]$. where $v(t) = \dfrac{t - t_j}{t_{j+1} - t_j}$. In addition, as neural firing rates cannot be negative, we defined $\lambda(t \mid \theta) = \max(\lambda(t \mid \theta), 0)$.

The adaptive algorithm specifies how the control magnitidues θ should be changed at each moment in time to reflect the current state of the system. For this problem we used the same algorithm in our previous work (Frank et al., 2002), which states that

$$\theta_t = \theta_{t-1} + \varepsilon \frac{d\lambda}{d\theta} [dN(t) - \lambda(t \mid \theta_{t-1})], \tag{B.2}$$

where $\lambda(\theta_{t-1})$ is the value of rate function at time *t-1*, θ_t is a vector of values of the control points at time *t*, $dN(t)$ is a *1* if there is a spike at time *t* and *0* if there is no spike time *t*, □, the updating interval, is *1* msec and ε is the algorithm learning rate. Eq. B.2 updates θ_t and the value of the rate function at *t* is computed as $\lambda(\theta_t)$. At each moment in time, this algorithm updates only the four control points closest to the current time point based on an error term computes as the difference observed spiking $dN(t)$ with the predicted probability of spiking $\lambda(\theta_{t-1})$□. The algorithm combines the results of that comparison with the estimates from the previous time step θ_{t-1}, and the relative contribution of the new spiking information (the error) and the previous estimate is determined by the learning rate parameter ε. For these analyzes, we chose a learning rate of 5 based on a careful examination of the data and on our previous simulation results. That learning rate allows for a single spike to change the firing rate function by as much as 5 Hz, and therefore weights heavily the new spiking information as compared to the previous estimate.

The initial magnitudes for the control points in the fixation (0 – 300 ms), scene (300 – 800 ms), delay (800-1500 ms), and response periods (1500 ms – end) were set to be the mean rates for the first trial for their respective periods. That ensured that the adaptive estimate would begin in a reasonable place and that it could therefore immediately track changes in the firing rate.

C. Spike train analysis. Using the distribution of the interspike intervals for each individual neuron, we calculated the proportion of spikes associated with interspike

interval of less than 10ms. This measure evaluates the proportion of spikes belonging to bursts (Buzsaki, Penttonen et al., 1996; Frank, Brown et al., 2001).

D. Selectivity Index. Using the absolute value of the averaged responses to each condition (minus baseline), we calculated a selectivity index (SI) before vs. after learning. SI provides a measure of the depth of selectivity and takes account of the responses to all the different scenes presented. It is defined as

$$SI = (n - \sum_{i=1}^{n} (\lambda_i / \lambda_{\max})) / (n-1), \qquad \text{(C.1)}$$

where n is the total number of scenes, λ_i is the firing rate of the neuron to the i^{th} scene and $\lambda_{\max}$ is the neuron's maximum firing rate to one of the scenes. Thus, if a neuron responds to only one scene and not to any other scene, the SI would be 1. If the neuron responded identically to all scenes, the SI would be 0.

References

Brown, E.N., Frank, L.M., Tang, D., Quirk, M.C., and Wilson, M.A. (1998). A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal lace cells. *J. Neurosci.*, 18, 7411-7425.

Brown, E.N., Nguyen, D.P., Frank, L.M., Wilson, M.A. and Solo, V. (2001). An analysis of neural receptive field plasticity by point process adaptive filtering. *PNAS*, 98:12261-66.

Buzsaki, G., Penttonen, M., Nadasdy, Z. and Bragin, A. (1996). Pattern and inhibition-dependent variation of pyramidal cell dendrites by fast spikes in the hippocampus in vivo. *PNAS*, 93:9921-9925.

Fahrmeir, L., and Tutz, D. (1994). Dynamic-stochastic models for time-dependent ordered paired-comparison systems. *J. Am. Stat. Assoc.*, 89, 1438-1449.

Frank, L.M., Brown, E.N. and Wilson, M.A. (2001). A comparison of firing properties of putative excitatory and inhibitory neurons from CA1 and the entorhinal cortex. *J. Neurophysiol.*, 86:2029:2040.

Frank, L.M., Eden, U.T., Solo, V., Wilson, M.A. and Brown, E.N. (2002). Contrasting patterns of receptive field plasticity in the hippocampus and the entorhinal cortex: an adaptive filtering approach. *J. Neurosci*., 22:3617-30.

Kitagawa, G., and Gersh, W. (1996). *Smoothness priors analysis of time series.* New York: Springer-Verlag.

Mendel, J.M. (1995). *Lessons in estimation theory for signal processing, Communication, and control.* New Jersey: Prentice Hall.

Smith, A.C., and Brown, E.N. (2003). Estimating a state-space model from point process observations. *Neural Computation*, 15, 1-27.