Supporting Information

We have proposed a theoretical framework that explains how temporal representations can be learned as a function of reward by local networks of recurrently connected excitatory neurons. In the following supplemental methods, we present mathematical details of our network model and the RDE learning rule, develop and demonstrate additional forms of the activity dependent *H* and *e* functions, provide implementation details of the integrate and fire neuron model described in the main article, discuss the theoretical prediction of increased noise correlations between neurons participating in a temporal representation, and present extensions to our model that demonstrate how network complexity can change quantitative aspects of our temporal representations without impacting the underlying qualitative features. The supporting information also includes 9 supplemental figures that illustrate additional aspect of our model from the main text and from the supplemental methods section.

Supplemental Methods:

Encoding time in a recurrent network. We start from the assumption that recurrent excitatory reverberations within a local network could underlie the temporal responses seen in V1 (Fig. S1). The following analysis, based on a leaky-integrator neuron model, quantifies the network structure required for this form of temporal representation.

The activity level of each neuron *i* in the recurrent layer of this network is represented in terms of an abstract variable, V_i, which approximates firing rate or average post-synaptic depolarization (1) and is roughly analogous to the flux reported by a voltage sensitive dye. Neural activity is driven both by the external feed-forward inputs and lateral excitation from within the recurrent layer.

The dynamics of a single neuron in our network of *N* excitatory cells are described by a first order differential equation:

$$\tau_m \frac{dV_i}{dt} = -V_i + I_{ext,i} + \sum_N L_{ij} V_j \qquad \text{(S1)}$$

where $I_{ext,i}$ is the external feed-forward input to neuron *i* and $L_{i,j}$ is the weight connecting pre-synaptic cell *j* to post-synaptic cell *i*. Following the experimental paradigm used by Shuler and Bear (2), we are primarily interested in the behavior of this system following a brief period of stimulation during which $I_{ext,i} > 0$. Post-stimulus dynamics (input vector $\mathbf{I} = \mathbf{0}$) can be described using matrix notation:

$$\tau_m \frac{d\mathbf{V}}{dt} = (\mathbf{L} - \mathbf{1}) \cdot \mathbf{V} \qquad \text{(S2)}$$

where $\mathbf{V} = [V_1, V_2, \ldots V_N]^{\mathrm{T}}$ is the vector of neural activities, $\mathbf{L}$ is an *N*x*N* matrix of recurrent weights, and $\mathbf{1}$ is the unit matrix.

Activity in an isolated neuron ($L_{i,j} = 0 \ \forall \ i,j$) decays exponentially to rest with an intrinsic time constant, τ_m (10 ms Fig. 2*B, black line*), which is abstractly related to the membrane

capacitance or synaptic time constant of a real neuron. Our goal is to represent time by decreasing the effective activity decay rate so that it corresponds to a desired interval, and to do so by setting the recurrent synaptic weights (Fig. 2*B, gray line*).

The network starts in a resting state ($\mathbf{V} = \mathbf{0}$) and there are two *1xN* input vectors, $\mathbf{I}^{\mu}$ (where $\mu = 1$ for left or 2 for right eye stimulation). These vectors are chosen to be orthogonal to each other so that neurons in the recurrent layer respond monocularly and are normalized to length 1. If the recurrent weight matrix is constructed such that the synaptic strength between each neuron *i* and *j* in the recurrent layer has the form:

$$L_{ij} = \sum_{\mu=1}^{2} \lambda^{\mu} I_i^{\mu} I_j^{\mu} \tag{S3}$$

where

$$\lambda^{\mu} = 1 - \frac{\tau_m}{\tau_d^{\mu}} \tag{S4}$$

then $\mathbf{I}^{\mu}$ are the eigenvectors of $\mathbf{L}$ with eigenvalues λ^{μ}. During the transient stimulation period the network activity grows in the direction of the stimulating vector. After the stimulation period ends, the state of the network is set by, and in the direction of, the input vector and decays with a desired time constant, τ_d, rather than at the intrinsic decay rate:

$$\frac{d\mathbf{V}}{dt} = \frac{1}{\tau_m}(\mathbf{L} - \mathbf{1}) \cdot \mathbf{V} = \frac{1}{\tau_m}(1 - \frac{\tau_m}{\tau_d^{\mu}} - 1) \cdot \mathbf{V} = -\frac{1}{\tau_d^{\mu}} \cdot \mathbf{V} \tag{S5}$$

.

This result demonstrates how to encode time, defined as the effective activity decay rate engendered by network structure, by constructing an appropriate weight matrix. This analysis assumes two feed forward input vectors, each corresponding to LGN activation patterns resulting from the stimulation of a single eye, but it can be extended to encode times associated with multiple, arbitrary input patterns. The form of equation S4 dictates practical limits to this approach; for large encoded times, small changes in synaptic weights cause a large changes in the network decay rate (Fig. S8) (3, 4). Following the experimental paradigm, the network must learn a weight matrix that encodes the intervals between paired stimuli and rewards; we next address how this can be done.

Learning with Reward Dependent Expression of synaptic plasticity (RDE). We have formulated a learning rule based on the hypothesis that reward can modulate expression of the slow molecular processes which lead to long term plasticity and are initiated by coactivation between neurons. We call these non-specific plasticity precursors 'proto-weights'. A simple equation describing the activity driven dynamics of the proto-weight, L^p, associated with a single synapse between two neurons *i* and *j* has the form:

$$\tau_p \frac{dL_{ij}^p}{dt} = -L_{ij}^p + H(V_i, V_j) \tag{S6}$$

where $H(V_i,V_j)$ describes a Hebbian dependence on the pre- and post-synaptic activity levels (see below). We assume that τ_p, the proto-weight decay time constant, is sufficiently long so that the proto-weights do not decay to resting levels before the time of reward (i.e. $\tau_p > T^\mu$).

Proto-weights do not directly impact the efficacy of synaptic signaling; rather, they determine the magnitude of long-term synaptic plasticity as a function of reward. Changes in synaptic weight are proportional to the magnitude of the proto-weights within individual synapses at the time of reward:

$$\frac{dL_{ij}}{dt} = \eta L_{ij}^p(t) \cdot \left(r_0 - e_V(i,t)\right)\ \delta(t - T^\mu) \tag{S7}$$

where η <<1 is the learning rate, r_0 is a measure of reward magnitude which is inhibited by cortical activation according to the function $e_V(i,t)$, and $\delta(x)$ is the Dirac delta function. A classical Hebbian learning rule is insufficient for our purposes since it is not known during stimulation when or whether reward will follow. Inhibition of reward is required so that learning will stop at the appropriate level even with continued training (see below, Fig. S2).

Activity dependent plasticity function H. Our goal is to create a weight matrix, as in equation S3, that encodes the reward timing. Using the simplest form of Hebbian plasticity, suggested by our analysis, proto-weights increase proportionally to the product of pre- and post-synaptic activity levels:

$$H(V_i,V_j) = V_i(t) \cdot V_j(t). \tag{S8}$$

We demonstrate in the main report that this form of learning works well for monocular input patterns (Fig. 4). A large family of functions, when embedded in the plasticity rule (Eq. S7), could potentially yield the desired network behavior. The exact choice of a Hebbian plasticity function *H* is non-trivial and is not the focus of this work. Currently, experimental observations are insufficient to constrain a specific functional form beyond the simplest forms suggested by our analysis.

Formulating $H(V_i,V_j)$ for binocular inputs. If the network contains binocularly responsive neurons (simulated by non-orthogonal input patterns), the simple Hebbian form of *H* presented in the main body is insufficient to produce well-segregated temporal responses to both input patterns. Since binocular neurons in the network will form synaptic connections to monocularly responsive neurons for both input patterns, as training progresses they will start to 'pull-up' activity levels in the nominally unresponsive populations resulting in non-selective activation and a failure to converge to the correct time constants. Learning can be made to converge, however, by choosing a more general form of *H* that allows for depression of synaptic weights between neurons that are not coactive. One form that meets this criterion is:

$$H(V_i,V_j) = (V_i - \theta_1)(V_j - \theta_1) - \theta_2. \tag{S9}$$

This form of H is similar to the covariance rule (5, 6) and can lead to negative (inhibitory) weights in the linear model.

We implemented H as in equation S9 with our deterministic neuron model and 10% binocular overlap for each input pattern. The values of θ_k were set heuristically with parameters θ_1=2.5, and θ_2=0 to allow for convergence with the given amount of overlap. This form of the learning rule allows training of two selective responses with correct time constants in a network with non-orthogonal overlap (Figs. S3 and S4), and similar results (not shown) can be achieved with non-zero values of θ_2 that prevent plasticity in the absence of input. A straightforward implementation creates negative weights in the final connectivity matrix, and some neurons would have both negative and positive efferent synaptic connections. Although the assumption that single neurons can have both excitatory and inhibitory influence on other neurons is not biologically realistic, a more physiologically accurate implementation of the model, including both excitatory and inhibitory inter-neurons, could likely accomplish similar results.

The difficulty of choosing H is similar to the difficulty of choosing an appropriate learning rule for associative memory that is both computationally desirable and is consistent with experimental observations (6-8) and a quantitative description of complex experimental results may require a more biophysically realistic plasticity model (9). Our demonstration, using a variant of the well know covariance plasticity rule, that RDE can produce distinct temporal representations in the presence of binocular indicates that the basic model framework is robust enough to accommodate different components while still producing the desired temporal representations.

If RDE is applied with binocular inputs, individual binocularly responsive neurons can participate in multiple temporal representations corresponding to the reward intervals associated with each eye's stimulation. These binocularly responsive neurons display different apparent decay time constants depending on which input pattern is presented (Figs. S4, S5). Generally, the ability of a single neuron to participate in the computation of several different temporal representations is evidence that the dynamics are being driven by a network property and not due to changes in the intrinsic response characteristics of a single cell. Experimental evidence of similar characteristics in biological networks would constitute additional support for a key assumption of this model.

Reward inhibition function e_V. Synaptic weights, with a direct implementation of the learning rule described above, will continue to potentiate with ongoing training even after the correct temporal representation is achieved. Since the desired behavior is for the network to stop learning when evoked activity persists until the time of reward, RDE assumes that ongoing activity in the network can quash the expressive action of the reward signal (similar to mechanisms used in reinforcement learning, see (10, 11)). A simple form for the inhibition function is:

$$e_V(t) = \beta \overline{V}(t) \qquad \text{(S10)}$$

Where the spatial average of network activity is defined as:

$$\overline{V}(t) = \frac{1}{N}\sum_{i=1}^{N} V_i(t) \qquad \text{(S11)}$$

and β is a scaling factor that sets the activity level where reward is fully inhibited. This formulation assumes that the cumulative network activity works to inhibit reward globally (as suggested for the dopaminergic system, 12) and we show that it is sufficient to train our network.

Formulating a local form of e_V. Alternatively, a local inhibition rule for neuron *i* can be written as:

$$e_V(i,t) = \beta V_i(t). \qquad \text{(S12)}$$

The global option could represent an actual inhibition of the reward signal by mean cortical activity, whereas the local option could represent inactivation of a single neuron's response to the reward signal, possibly through receptor sensitivity to membrane voltage (13). We have found that both local and global forms of reward inhibition can produce very similar results (compare Figs. 4 and S6) with both spiking and non-spiking neuron models.

Conductance based integrate and fire neuron model. The equation governing the dynamics of the membrane potential of a single neuron *i*, v_i, is:

$$C\frac{dv_i}{dt} = g_L(E_L - v_i) + g_{E,i}(E_E - v_i) + g_{I,i}(E_I - v_i) \qquad \text{(S13)}$$

where *C* is the membrane capacitance, and E_x are the reversal potentials for ionic currents associated with leakage, excitation, and inhibition and g_L is the leak conductance. $g_{E,i}$ and $g_{I,i}$ are excitatory and inhibitory conductances, respectively. These conductances are equal to the product of synaptic activation, *s*, and synaptic weights.

Synaptic activation jumps by a percentage ρ, saturates at 1, and decays with time constant τ_s.

$$\frac{ds}{dt} = -\frac{1}{\tau_s}s + \rho(1-s)\cdot \sum_{\substack{pre-\\ spikes}} \delta\left(t - t_{pre}\right) \qquad \text{(S14)}$$

The resting membrane voltage was set to -60 mV, and reversal potentials for excitatory, inhibitory, and leak ionic species were -5, -75, and -60 mV respectively. Spiking occurred when membrane voltage reached a threshold value of -55 mV after which *v* was reset to -61 mV and held for a 2 ms absolute refractory period. The leak conductance was 0.01 nS and membrane capacitance was set to give a membrane time constant of 20 ms. As in previous models (14, 15), synaptic activation decays with a slow time constant to account for NMDA activation dynamics (80 ms used here) and ρ = 1/7. At each time point, a windowed estimate of firing rate was updated according to:

$$\frac{dR_i}{dt} = \frac{1}{\tau_w}\left(\sum_{t_i}\delta(t-t_i) - R_i\right) \qquad \text{(S15)}$$

where t_i are the times of spikes, τ_w is the estimate time window (100 ms), and R_i is the rate estimate. R_i was used as a proxy for the activity variable (V from the continuous model) when calculating changes in the proto-weights, according to equation S6. While the spiking neuron model could have incorporated a plasticity rule that more explicitly related weight changes to spike timing without the R variable, we chose this form to minimize training differences between the linear and non-linear models.

The recurrent layer was stimulated with feed-forward inputs consisting of Poisson spike trains with time-varying intensity (to mimic LGN activity, see (16, 17)) for 400 ms. Left and right eye stimulation patterns were orthogonal, each stimulating 50 neurons in the network. The β parameter was set so that reward would be completely inhibited when the average firing rate was around 20 Hz, and the offsets between the end of stimulation and reward delivery were 500 ms for the left eye, 1000 ms for the right. Proto-weights, for both neuron models, decayed with a time constant of 5 sec.

Dynamical equations for both neuron models were integrated numerically in custom code written in MATLAB (linear model) and c++ (spiking model).

Theoretical prediction of increased noise correlations in trained network. Since time is stored in excitatory connections between neurons that respond to the same stimulus, we expect that noise correlations between neurons within responsive populations to be higher in trained networks than in naïve networks. To verify this, a naïve network of spiking neurons and one trained with a reward time of 1400 ms were both allowed to run randomly, without stimulation, for 100 seconds starting from 50 random seeds and calculated correlations as follows.

Our integrate and fire model produces a binary spike train for each neuron i, defined as $x_i^k(t) = 1$ if neuron i fires a spike at time t on trial k, and 0 otherwise. The spike count over a single run for a single neuron, where $1 \le t \le T$, is:

$$N_k^i = \sum_{t=1}^{T} x_i^k(t). \qquad \text{(S16)}$$

Recurrent layer neurons were divided into two groups based on their indici where group 1 (g_1) included neurons i=1-50 and group 2 (g_2) neurons i=51-100. The spike trains for the two groups were combined to form two new random processes;

$$x_{g_1}^k(t) = \sum_{i=1}^{50} x_i^k,\; x_{g2}^k(t) = \sum_{i=51}^{100} x_i^k \qquad \text{(S17a,b)}$$

and spike counts were calculated for these two new processes such that

$$N^k_{g_1,g_2} = \sum_{t=1}^{T} x^k_{g_1,g_2}(t). \tag{S18}$$

The Pearson correlation coefficient, *r*, reported in figure S7 was calculated (18) for the two combined processes according to:

$$r = \frac{E\left[N_{g_1} N_{g_2}\right] - E[N_{g_1}]E[N_{g_2}]}{\sigma_{Ng_1}\sigma_{Ng_2}}, \tag{S19}$$

where *E* is the expectation over *M* trials.

Using the same notation as in the experimental literature, the spike train correlation between two processes, x_j and x_k, is defined as:

$$C_{jk}(\tau) = \frac{1}{M}\sum_{i=1}^{M}\sum_{t=1}^{T} x^i_j(t)x^i_k(t+\tau), \tag{S20}$$

and the shift predictor between these same processes is:

$$S_{jk}(\tau) = T\sum_{t=1}^{T} P_j(t)P_k(t+\tau), \tag{S21}$$

where the P_k is the post-stimulus time histogram (PSTH) for process k:

$$P_k(t) = \frac{1}{M}\sum_{i=1}^{M} x^i_k(t). \tag{S22}$$

Our cross-correlogram (CC) is defined as:

$$CC(\tau) = \frac{C_{jk}(\tau) - S_{jk}(\tau)}{\sqrt{\nu_j \nu_k}} \tag{S23}$$

where $\nu_{j,k}$ are the average spike frequencies of processes *j* and *k*. An exponential fit of *CC* demonstrates the approximate duration of correlation effects. The normalization by frequency results in a variance that is approximately equal for the naïve and trained networks.

For both the naïve and trained networks, single neuron firing statistics are similar with a CV (1) of 1.18 for the naïve network and 1.04 for the trained. Figure S7 demonstrates that the cross-correlogram (CC) between two groups of 50 neurons changes as a function of training. For the untrained network, where recurrent connections are weak, the CC is flat. In the trained network, however, the CC shows a prominent peak. Correspondingly, the Pearson correlation coefficient between these two groups increases from zero in the naïve network to 0.309 in the trained network. This result is a robust prediction that can be used to test the model experimentally, although it requires a large amount of spontaneous data recorded simultaneously

from multiple neurons. Unfortunately, the amount of spontaneous data available from the original experiment (2) is insufficient to test this prediction.

Temporal representations with complex network structures. The central objective of this work is to describe how recurrent excitatory synapses can serve as the neural substrate of learned temporal representations. Our model, which was purposely designed with a minimal set of assumptions to enable both analysis and clear exposition of the underlying principles, is able to account for the key qualitative features of the experimental data. Successively closer approximations to the experimental data can be achieved by including successively detailed components of the biological network in our model. Such additions, while outside the main thrust of this paper, can account for additional experimental details. Here we demonstrate, using an ad hoc approach, that populations of inhibitory can decrease evoked firing rates and create an additional form of observed temporal representation. The purpose is not to rigorously develop extensions to our model, but rather to demonstrate how network complexity can change its quantitative predictions.

A network of 100 integrate and fire neurons was trained to respond to LE and RE stimulus using RDE (all neuron and training parameters as above). First, a population of 20 inhibitory neurons with fast synaptic time constants (τ_s = 5 ms) was added to the network. Random synaptic connections were made between the recurrent layer and inhibitory neurons (15% probability of connection), and between the inhibitory neurons and recurrent layer (30% probability of connections). Synaptic weights were set so that high rates of activity in the recurrent layer would evoke activity in the inhibitory neurons. As shown in Fig. S6, the presence of inhibitory feedback can decrease evoked firing rates in the recurrent layer.

Next, two new populations of neurons were added to the trained network. The first consisted of 100 neurons that fire spontaneously with an average rate of about 4 Hz. The second is a layer of 100 inhibitory inter-neurons connecting each neuron in the recurrent layer with one of the spontaneous neurons. As shown in Fig. S5 inhibition driven by the recurrent layer results in a sustained decrease in firing rate corresponding to reward times encoded in the network by RDE.

References

1. Dayan P & Abbott LF (2001) *Theoretical neuroscience: computational modeling of neural systems* (MIT Press, Cambridge, MA).
2. Shuler MG & Bear MF (2006) Reward timing in the primary visual cortex. *Science* 311:1606-1609.
3. Seung HS (1996) How the brain keeps the eyes still. *Proc Natl Acad Sci U S A* 93:13339-13344.
4. Tegner J, Compte A, & Wang XJ (2002) The dynamical stability of reverberatory neural circuits. *Biol Cybern* 87:471-481.
5. Linsker R (1986) From basic network principles to neural architecture: emergence of spatial-opponent cells. *Proc Natl Acad Sci U S A* 83:7508-7512.

6. Sejnowski TJ (1977) Storing covariance with nonlinearly interacting neurons. *J Math Biol* 4:303-321.
7. Chechik G, Meilijson I, & Ruppin E (2001) Effective neuronal learning with ineffective Hebbian learning rules. *Neural Comput* 13:817-840.
8. Dayan P & Willshaw DJ (1991) Optimising synaptic learning rules in linear associative memories. *Biol Cybern* 65:253-265.
9. Shouval HZ, Bear MF, & Cooper LN (2002) A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. *Proc Natl Acad Sci U S A* 99:10831-10836.
10. Rescorla RA & Wagner AR (1972) *Classical Conditioning II: current research and theory*, eds Black AH & Prokasy WF (Appleton-Century-Crofts, New York, NY), pp 64-69.
11. Sutton RS & Barto AG (1998) *Reinforcement learning: an introduction.* (MIT Press, Cambridge, MA).
12. Schultz W, Tremblay L, & Hollerman JR (1998) Reward prediction in primate basal ganglia and frontal cortex. *Neuropharmacology* 37:421-429.
13. Ben-Chaim Y*, et al.* (2006) Movement of 'gating charge' is coupled to ligand binding in a G-protein-coupled receptor. *Nature* 444:106-109.
14. Compte A, Brunel N, Goldman-Rakic PS, & Wang XJ (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex* 10:910-923.
15. Lisman JE, Fellous JM, & Wang XJ (1998) A role for NMDA-receptor channels in working memory. *Nat Neurosci* 1:273-275.
16. Mastronarde DN (1987) Two classes of single-input X-cells in cat lateral geniculate nucleus. II. Retinal inputs and the generation of receptive-field properties. *J Neurophysiol* 57:381-413.
17. Ruksenas O, Bulatov A, & Heggelund P (2007) Dynamics of spatial resolution of single units in the lateral geniculate nucleus of cat during brief visual stimulation. *J Neurophysiol* 97:1445-1456.
18. Bair W, Zohary E, & Newsome WT (2001) Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J Neurosci* 21:1676-1697.

Supplementary Figure Legends

Figure S1 Example responses of single V1 neurons to visual stimulation. Plot of the average evoked response of a neuron recorded in a naïve animal (dotted green) and the after training responses of neurons trained with short (red) and long (blue) reward times. In the naïve animal, neurons respond briefly during the period of stimulation (green bar). During training, left eye and right eye stimulations are *paired* with rewards delivered after a short (ST) or long (LT) delay period, respectively (dashed vertical lines). After training, neuronal responses evoked by a given stimulus can persist until the reward time paired with that stimulus. The plots show the difference between the dominant and non-dominant eye responses per neuron smoothed with a Gaussian kernel (SD 50 ms); for detailed methods and plots of average population responses see Shuler and Bear (2006).

Figure S2 Dependence of τ_d on λ. The value of λ (Eq. S4) required in order to set the correct weight matrix to encode a particular time (Eq. S3) limits the effective interval range that our approach can learn robustly. This plot shows the dependence of the network time constant (in ms) on the eigenvalue of the weight matrix for two different values of the intrinsic neuron decay time constant. The steep rise of the curve in the temporal region of interest for our training task (1-2 seconds) means that recurrent synaptic weights must be learned to a high degree of precision. This makes robust learning inherently difficult. Note that the relatively faster dynamics associated with the shorter τ_m (red line) result in a steeper curve than with the longer value of τ_m (blue line). An analogous plot can also be generated numerically for the spiking neuron model.

Figure S3 Training in the rate based model. This plot shows the response of the neurons responsive to left (blue lines) or right eye (red lines) stimulation during each epoch of a training session. The stimulus is active during the period indicated by the green patch and cyan lines show reward times. The response of the naïve network, indicated by thick black line, decays quickly back to zero following stimulation. As training progresses, the responses to both inputs increase until the appropriate activity level is reached at the time of reward at which time learning stops.

Figure S4 Network trained with overlapping inputs. Naïve network structure is the same as in figure 3, but the input patterns are set so that several neurons respond binocularly (binocular neurons marked with arrows). Neurons in the trained network respond as with monocular training, except that binocular neurons have learned appropriate responses for both reward times. See supplemental figure S5 for plot of a binocular neuron's activity elicited by both inputs.

Figure S5 Response of a binocular neuron to left and right eye stimulation. This plot demonstrates the normalized activity of a single binocularly responsive neuron to both left (blue line) and right (red line) eye stimulation. The stimulus is active during the period indicated by the green patch. The reward times used during training for both stimuli are shown by ticks at 1000 and 1600 ms. The binocular neuron participates in computation of both temporal representations and responds with different time constants depending on which input is presented to the network. The neuron shown is one of the binocularly responsive neurons from Fig. S4.

Figure S6 Training with local inhibition. This plot shows the results of training the rate-based model with local reward inhibition. The results are virtually identical to training using inhibition based on average activity in the entire recurrent layer (see Fig. 3).

Figure S7 Noise Correlations. Spontaneously active neurons in a network trained with a reward time of 1400 ms have higher correlation coefficients ($r = 0.309$) than in a naïve network ($r = -0.001$). Plots of the cross-correlograms show a flat temporal profile in the naïve case (black) and a region of increased correlation between neurons in the trained network (gray). An exponential fit (dashed line) shows the region of increased correlation. Note correlation values are calculated across grouped neural populations (see Supplemental Methods).

Figure S8 Inhibition can decreases evoked firing rates. This plot demonstrates one way homeostatic mechanisms could be used by the brain to limit maximum firing rates of neurons in trained networks. Neurons in the recurrent layer (labelled E in inset) were trained to respond to left and right eye stimulation with two different reward times (shown by dashed lines). In the left column, neurons in E receive only excitatory feed-forward and lateral inputs. In the right column, these same neurons excite a smaller population of inhibitory neurons (labelled I)

that project back into the recurrent layer. The inhibition decreases the magnitude of evoked response with little impact on temporal representations. Connections between layers E and I were established randomly and weights were set heuristically to demonstrate this result. Stimulation is active during the periods indicated by green bars.

Figure S9 Sustained decrease. This plot demonstrates how a population of neurons trained to show sustained increase in firing rates can engender a sustained decrease when embedded in a network including inhibitory neurons. As shown in the inset, recurrent layer neurons (E) drive inter-neurons (I) that inhibit the activity of a population of spontaneously spiking neurons (S). The rasters and histograms in the left column show spiking activity of the neurons in S without inhibition; this activity is not affected by the feed-forward stimulation of E (indicated by green bars). Including inhibitory connection between I and S (right column) results in a decreased firing rate of S that persists until the time of reward (dashed line) for both left and right-eye stimulation (top and bottom rows). Here, reward is still encoded by recurrent excitatory weights and the sustained decrease is a derivative form of representation.