

Neuron, Volume 80

Supplemental Information

A Theory of the Transition to Critical Period Plasticity: Inhibition Selectively Suppresses Spontaneous Activity

Taro Toyozumi, Hiroyuki Miyamoto, Yoko Yazaki-Sugiyama, Nafiseh Atapour, Takao K. Hensch,
and Kenneth D. Miller

Contents

S1. Supplemental Analysis

S1a. A Simplified, Analytically Tractable Model

S1b. Estimation of the spontaneous-to-visual activity ratio

S2. Supplemental Computational and Experimental Procedures

S2a. Circuit model

S2b. Input statistics

S2c. A plasticity rule and a neuron model

S2d. Receptive field width, Response strength, Contra-bias index, and Gain

S2e. Extracellular recording from freely behaving mice

S2f. LTD Experiments

Supplemental Figures

Figure S1: Results of modeling a linear postsynaptic neuron with a simple activity-dependent plasticity rule

Figure S2: Baseline firing rates and LED-evoked firing rates

Figure S3: The cumulative distributions of the bias-corrected spontaneous-to-visual ratio of firing rates

Figure S4: LTD is intact in *GAD65*-KO mice

Supplemental References

S1. Supplemental Analysis

S1a. A Simplified, Analytically Tractable Model

To try to isolate the essential features that are sufficient to explain pre-CP and CP plasticity, we simplify the model presented in the text to a few essential ingredients. The main ideas we want to capture in this simple model are (1) plasticity occurs through a combination of Hebbian and homeostatic plasticity; (2) before the maturation of inhibition, there is sufficient spontaneous input such that the influence of rare visual input is limited; (3) after the maturation of cortical inhibition at the onset of the critical period, the more frequent but weaker spontaneous activity is largely suppressed by the combination of inhibition and the neuron's nonlinear input-output relation (c.f. Fig. 1B), so that the visual input plays a much stronger role in determining weight development. We show that a simple, minimal implementation of these three ideas suffices to explain the main features of pre-CP and CP plasticity, which demonstrates that these results do not depend on other, more complex aspects of the model presented in the text. We consider a linear neuron model and a simple plasticity rule, which together allow us to analytically describe the determinants of model behavior. This enables us to point to the precise relationships that are needed to explain pre-CP and CP behavior.

For a linear postsynaptic neuron, the output firing rate, y , is given by a weighted sum of input firing rates, i.e., $y = \vec{w} \cdot \vec{x}$ where \vec{x} is the vector of input firing rates and \vec{w} is the vector of synaptic weights. We consider a simple learning rule that combines Hebbian and homeostatic plasticity. The Hebbian term is driven by the covariance between pre- and post-synaptic activities, $\langle \vec{x}y \rangle - \langle \vec{x} \rangle \langle y \rangle$, where the brackets indicate an average over input activity patterns (Sejnowski and Tesauro, 1989). A homeostatic component, proportional to $\vec{w}(y_0 - \langle y \rangle)$, scales synaptic strengths depending on the difference between the average output firing rate $\langle y \rangle$ and a desired or set-point firing rate y_0 . The proportionality to \vec{w} causes the homeostatic term to multiplicatively scale synaptic strengths. Such homeostatic scaling has been experimentally observed (Turrigiano et al., 1998; Stellwagen and Malenka, 2006), including in the visual cortex of developing animals in vivo (Desai et al., 2002; Mrcsic-Flogel et al., 2007; Kaneko et al., 2008). The change in synaptic weights per unit time can then be written as

$$\Delta \vec{w} = \eta \left[Q \vec{w} + \gamma \vec{w} (y_0 - \langle y \rangle) \right],$$

where $Q = \langle \vec{x} \vec{x}^T \rangle - \langle \vec{x} \rangle \langle \vec{x}^T \rangle$ is the input covariance matrix (here, \vec{x} is a column vector and its transpose \vec{x}^T is a row vector), $\gamma = 1$ Hz is a constant that describes the strength of homeostasis relative to Hebbian plasticity, $\eta = 1$ s² is a learning rate, and $y_0 = 1$ Hz.

As in the main text, we assume that the postsynaptic neuron receives only spontaneous input 90% of the time and receives both spontaneous and visual input 10% of the time. In this simple model, we do not explicitly model how inhibition suppresses the contribution of spontaneous activity to plasticity relative to that of visual activity, but simply assume that the combination of cortical inhibition and a nonlinear plasticity rule suppresses synaptic plasticity a fraction f ($0 \leq f \leq 1$) of the time in the absence of visual input. Note that f replaces the variable m from the main text as a quantifier of effective inhibitory strength. The spontaneous and visual inputs are Gaussian random variables that have means $\vec{\mu}^{(S)}$ and $\vec{\mu}^{(V)}$ and covariances $Q^{(S)}$ and $Q^{(V)}$ respectively. Hence, the effective input mean and covariance are given by

$$\begin{aligned} \vec{\mu} &= 0.1(\vec{\mu}^{(V)} + \vec{\mu}^{(S)}) + 0.9(1-f)\vec{\mu}^{(S)}, \\ Q &= 0.1(Q^{(V)} + Q^{(S)}) + 0.9(1-f)Q^{(S)}. \end{aligned}$$

The factor $0.9(1-f)$ qualitatively describes the percentage of time Hebbian plasticity is induced in the absence of visual stimulus, which we refer to in the main text as the gain of spontaneous activity for Hebbian plasticity.

We use the same parameters to model spontaneous and visually-driven input activity as in the main text. In particular, we model monocular deprivation (MD) by assuming that the visual component of deprived-eye activity is decreased in strength by $\frac{1}{2}$, while the distance over which its correlations decay is multiplied by 5 (representing the blurring effect of the closed eyelid, as has been observed in effects of eyelid closure on cortical retinotopic selectivity in vivo (Faguet et al., 2009)). We first consider two extreme values of inhibitory strength to model the pre-CP ($f=0$) and CP ($f=1$), which yield the final receptive fields of the two eyes shown in Fig. S1A for

normal rearing (NR) and for MD of the contralateral eye. In Fig. S1B, we more systematically examine the influence of inhibition on the MD effect. During the pre-CP (weak inhibition; small f), MD suppresses both eyes' responses while maintaining the dominance of the closed eye, with only small changes in CBI. During the CP (strong inhibition; large f), MD enhances the open-eye and suppresses the closed-eye response so that the open eye becomes strongly dominant over the closed eye, with correspondingly large changes in CBI. MD widens the receptive fields relative to NR both during the pre-CP and the CP. This result is consistent with experiments (Smith and Trachtenberg, 2007).

We can analytically describe the model outcome as follows. The only stable fixed point of this learning rule is proportional to the principle eigenvector (the eigenvector with largest eigenvalue), \vec{v} , of the input covariance matrix Q (Oja, 1989; Miller and MacKay, 1994), which we take to be normalized to unit vector length. Assuming the initial set of synaptic weights is not perpendicular to \vec{v} , the synaptic weights converge to a steady-state (fixed point) value \vec{w}_{ss} given by

$$\vec{w}_{ss} = \frac{y_0 + \lambda / \gamma}{\vec{v} \cdot \vec{\mu}} \vec{v},$$

where $\vec{\mu} \equiv \langle \vec{x} \rangle$ and $\lambda > 0$ is the largest eigenvalue of Q (the eigenvalue corresponding to \vec{v}). We further assume that there are no negative or zero correlations between presynaptic neurons. Then the principle eigenvector, \vec{v} , has no negative or zero elements (this is known as the Perron-Frobenius theorem), which guarantees that steady-state synaptic strengths are positive.

The interaction of Hebbian and homeostatic plasticity is the key to the results. As we can see from the fixed-point solution for the synaptic weights, the effects of MD arise through changes in the principal eigenvalue, λ , and the normalized eigenvector, \vec{v} , which are determined by the Hebbian component; and through the change in the mean presynaptic activity $\vec{\mu}$ and thus in $\vec{v} \cdot \vec{\mu}$, which contributes due to the homeostatic component. Of these, only changes in \vec{v} differentially affect the two eyes; the other changes only change the common scale factor applied to both eyes' weights.

The principle eigenvector of the input covariance matrix, \vec{v} , which is also known as the principle component of the input activity, represents the pattern of activity – the set of relative activities across input neurons – that has the most variance in the input activity (the absolute activity levels are then determined by setting the vector length of \vec{v} to 1). Because of positive between-eye correlations, this pattern involves both eyes' inputs being active. If the two eyes have equivalent activities, the two eyes are equally strong in \vec{v} , but to the extent to which the deprived eye has weaker activity than the non-deprived, the deprived eye will be less active in the highest-variance activity pattern and so will be weaker than the non-deprived eye in \vec{v} . The highest-variance activity pattern is centered on the receptive field center, because the inputs at the edges of the receptive field have fewer correlated inputs with which to be co-active. If spatial correlations are narrow, this pattern is more tightly peaked around the receptive field center, while broader correlations lead to a spatially broader and flatter activity pattern. Figure S1A, shows the final receptive fields under various rearing conditions, which in each case are proportional to \vec{v} for that rearing condition. (The number of inputs is larger from the contralateral eye than from the ipsilateral eye as assumed in the main text.) Thus, from Fig. S1A, we can see that, upon MD to the contralateral-eye, the eigenvector, \vec{v} , changes its profile by suppressing the deprived eye components and facilitating the non-deprived eye components. At the same time, the visual input through the closed eyelid induces more spatially uniform input correlation and flattens and spreads the eigenvector.

The eigenvalue λ is the amount of variance that \vec{v} contributed to the input activity. λ decreases under MD because of the loss of input correlation. This decrease of λ reduces \vec{w}_{ss} and thus reduces responses to both eyes. On the other hand, due to the loss of the deprived-eye input activity through $\vec{\mu}$, homeostatic plasticity causes non-selective growth of both eyes' responses. Therefore, the deprived eye will always show less response than the non-deprived eye, but, depending on three factors -- the relative strength of homeostasis, the degree of reduction of the deprived-eye input, and the width of the input correlations in the deprived eye -- both eyes' responses could be suppressed, both could be enhanced, or the deprived-eye could be suppressed while the other eye is enhanced. To cause both eyes' response peaks to be reduced by MD during the pre-CP, we used relatively small γ and y_0 , i.e. relatively weak homeostatic plasticity and a relatively low activity-set point, which

enhances the effect of the decrease in the eigenvalue relative to the effect of homeostasis; and we used relatively wide input correlations from the deprived eye, which flattens the eigenvector. Maturation of inhibition suppressed the spontaneous component, which is balanced between the two eyes, and thus enhanced the difference between the two eyes under MD via the change in the eigenvector, \vec{v} .

This simple model is useful to understand how the balance of spontaneous and visually evoked activity contributes to the activity-dependent plasticity in the course of development. In particular, it already shows the major effects we are trying to understand: before the maturation of inhibition, MD slows retinotopic refinement and reduces both eyes' responses without causing substantial OD shifts; after the maturation of inhibition, MD causes strong OD shifts as well as a slowing of refinement. This simple model does not achieve a confined receptive field (synaptic weights never become zero) and does not address the mechanism by which inhibition modulates the balance of spontaneous and visually evoked activity. In the main text, we use a learning rule that yields confined receptive fields and explicitly model how maturation of inhibition suppresses the relative contribution of spontaneous component over the visually-evoked component, which was described above by the factor $(1-f)$.

S1b. Estimation of the spontaneous-to-visual activity ratio

In the main-text, we estimated the spontaneous-to-visual ratio of firing rates based on extracellular recording from behaving mice. We also explained in the main-text that this ratio is likely overestimated. In this section, we show that, under simple assumptions, the estimated spontaneous-to-visual ratio has a monotonic relation to the true underlying spontaneous-to-visual ratio. Assuming that the visually evoked component linearly adds to spontaneous firing rate, the true spontaneous-to-visual ratio (the spontaneous firing rate divided by the visually evoked firing rate driven by natural scenes) can be written $r_s/(r_s+r_v)$, where r_s is the spontaneous firing rate and r_v is the increment of the firing rate due to an optimal visual stimulus. On the other hand the estimated spontaneous-to-visual ratio (the baseline firing rate divided by the LED-evoked firing rate) is described by $(r_s+q*r_v)/(r_s+q*r_v+b*r_v)$, where q ($0<q<1$) is the frequency of visual input under the free viewing condition and the factor $b*r_v$ with a constant, b ($0<b<1$), describes the increment of firing rate by the LED stimulus. Letting $k=r_s/r_v$, the true ratio is $k/(k+1)$, while the estimated ratio is $(k+q)/(k+q+b)=k'/(k'+1)$. with $k'=(k+q)/b$. It is now easy to see, in this case, that both the true and the estimated spontaneous-to-visual ratios are monotonically increasing functions of k . Thus, one ratio is a monotonically increasing function of the other, and vice versa. Note that the estimated ratio is always greater than the true ratio, because $k'>k$.

S2. Supplemental Computational and Experimental Procedures

S2a. Circuit model

We model a pyramidal cell in V1 that receives input from $N=1184$ LGN neurons. The presynaptic neuron i represent input from either contra- or ipsilateral eye, $e_i \in \{C, I\}$, and from a particular retinotopic position, \vec{z}_i , uniformly spaced on a square grid across a two-dimensional surface of $\{\vec{z} = (z_1, z_2) \mid -1 < z_1 \leq 1, -1 < z_2 \leq 1\}$.

The first $N_C=28*28$ presynaptic neurons represent input from the contralateral eye ($e_i=C$) and the other $N_I=20*20$ neurons represent input from the ipsilateral eye ($e_i=I$). In order to model the anatomical bias toward the contra-lateral eye observed in mice, the number of neurons representing the contralateral eye input is about twice larger than that from the ipsi-lateral eye.

S2b. Input statistics

We assume that the input has both a spontaneous component, $x_i^{(S)}$, and a visual component, $x_i^{(V)}$. We also assume that $x_i^{(S)}$ and $x_i^{(V)}$ are independent Gaussian random variables with mean $\mu_i^{(S)}$ and $\mu_i^{(V)}$ and covariance $Q_{ij}^{(S)}$ and $Q_{ij}^{(V)}$, respectively. The neuron receives the spontaneous and the visual input, $x_i=x_i^{(S)}+x_i^{(V)}$, 10% of the time, and the neuron receives only the spontaneous input, $x_i=x_i^{(S)}$, 90% of the time. This 10% figure indicates that the visual input is sparse (Vinje, 2000) and guarantees a significant contribution of spontaneous input on synaptic plasticity during the pre-CP. Many of the cortical neurons in mice are orientation selective (Sato and Stryker, 2008), so this may imply that the optimal visual stimulus with a preferred orientation at a preferred retinotopic position is rare. Under the normal condition, the input means are uniform across the input neurons and are given by $\mu_i^{(S)}=1$ Hz and $\mu_i^{(V)}=2$ Hz, assuming twice as strong visual input as spontaneous input. The covariance of input depends on the corresponding eyes and also on the retinotopic distance, $|z_i-z_j|$, of the two presynaptic neurons,

i.e., $Q_{ij}^{(S)} = \text{Cov}[x_i^{(S)}, x_j^{(S)}] = c_{e_i, e_j} \mu_i^{(S)} \mu_j^{(S)} G(\bar{z}_i - \bar{z}_j, L)$ and

$$Q_{ij}^{(V)} = \text{Cov}[x_i^{(V)}, x_j^{(V)}] = c_{e_i, e_j} \mu_i^{(V)} \mu_j^{(V)} G(\bar{z}_i - \bar{z}_j, L), \text{ where } G(\bar{z}, L) = \frac{1}{2\pi L^2} \exp\left(-\frac{|\bar{z}|^2}{2L^2}\right) \text{ is a Gaussian}$$

distribution with isotropic standard deviation $L=0.06$; the within-eye coefficients are set to $c_{e_i=e_j} = 2\pi L^2$ so that the diagonal components, $Q_{ii}^{(S)} = \mu_i^{(S)} \mu_i^{(S)}$, are invariant to the width, L , of the correlation; the between-eye coefficients are set to $c_{e_i \neq e_j} = 0.5 c_{e_i=e_j}$ indicating half as strong between eye-correlation as the within-eye correlation. Those correlation coefficients, c_{e_i, e_j} , are common to both spontaneous and visual components. The spatial correlation profile is modeled as a Gaussian function of the retinotopic distance so that inputs from nearby retinotopic positions are more strongly correlated than distant ones. This spatial correlation structure produces localized (rather than full field) RFs under the competitive Hebbian learning rule, which will be introduced in the next section. The interocular correlations are based on the observation of strong inter-ocular correlations in LGN spontaneous activity, which are induced by cortical feedback: (Ohshiro and Weliky, 2006; Weliky and Katz, 1999).

If the between-eye coefficient for the visual input were higher than that of spontaneous input, binocular visual deprivation would reduce between-eye correlations, which would tend to shift the OD toward the anatomically stronger contralateral eye (for example, BD during the pre-CP would shift OD towards the contralateral eye, contrary to the observations of Smith and Trachtenberg, 2007). The assumption that the spatial structure of covariance is the same in spontaneous and visually evoked signals was introduced for simplicity. The model produces similar results as long as the two covariance functions have similar widths. If the spontaneous covariance is much wider than the visually evoked covariance, the model predicts a similar level of retinotopic refinement under MD and under monocular inactivation (MI) during the pre-CP, because the visually evoked covariance driven by the residual input through the closed eyelid under MD may be as broad as the spontaneous covariance in this case. This runs counter to the findings of (Smith and Trachtenberg, 2007). If the spontaneous covariance is much narrower than the visually evoked covariance, the model predicts that refinement of retinotopy should slow or reverse after maturation of inhibition, which again runs counter to the experimental findings. The parameterization of the input covariance proportional to the input mean square is consistent with the multiplication of input by a constant factor.

Only the visual component is modulated by visual deprivation. Under MD, where one eyelid is closed by suturing, we assume that the visual input from the closed-eye is described by $\tilde{x}_i^{(V)} = A_{el} G(\bar{z}_i, L_{el}) * x_i^{(V)}$, where the original visual input, $x_i^{(V)}$, is convolved by an eyelid filter $A_{el} G(\bar{z}_i, L_{el})$. The closed-eye input strength is reduced by a factor of $A_{el}=0.5$ and is also blurred by a spatial Gaussian filter with $L_{el}=0.3$. To the extent to which the closed eyelid passes light, it would be expected to filter out all higher spatial frequencies leaving only very blurry, large-scale changes in luminance, which would yield greatly broadened correlations, as we model. In our model this residual visual stimulus is important to explain the restricted refinement of RF under MD but nearly normal refinement under MI as observed during the pre-CP (Smith and Trachtenberg, 2007). Thus, the mean input strength from the deprived eye is reduced to $\tilde{\mu}_i^{(V)} = A_{el} \mu_i^{(V)}$; the within-deprived-eye covariance becomes $\tilde{Q}_{ij}^{(V)} = c_{e_i=e_j} \tilde{\mu}_i^{(V)} \tilde{\mu}_j^{(V)} G(\bar{z}_i - \bar{z}_j, \sqrt{L^2 + 2L_{el}^2})$, where e_i and e_j are both the deprived eye; and the between-eye covariance becomes $Q_{ij}^{(V)} = c_{e_i \neq e_j} \tilde{\mu}_i^{(V)} \mu_j^{(V)} G(\bar{z}_i - \bar{z}_j, \sqrt{L^2 + L_{el}^2})$, with e_i the deprived eye. The between-eye covariance and within-deprived-eye covariance have standard deviation of $\sqrt{L^2 + L_{el}^2}$ and $\sqrt{L^2 + 2L_{el}^2}$ because the eyelid filter is once and twice convolved with the normal covariance, respectively. Under MI by pharmacologically inactivating retinal activity or by enucleating one eye, the visual component from the inactivated eye becomes zero, which corresponds to setting $A_{el}=0$ in the MD scenario. The spontaneous component is assumed to be unchanged under MI for simplicity. Note that we are modeling spontaneous activity in LGN (Ohshiro and Weliky, 2006; Weliky and Katz, 1999) rather than in retina as the source of spontaneous input.

S2c. A plasticity rule and a neuron model

The firing rate of the postsynaptic neuron is described by y and the firing rate of the i th presynaptic neuron is described by x_i . The output firing rate is given by $y = \left[\sum_{i=1}^N w_i x_i - m \bar{y}_{\text{nb}} \right]_+$, where $[\]_+$ is a threshold nonlinearity that rectifies negative arguments to guarantee positive firing rates, w_i is synaptic strength from the presynaptic neuron i , m is the inhibitory synaptic strength from the neighboring neurons, and \bar{y}_{nb} is a long-term average firing rate of the neighboring neurons. Although the average firing rate of the neighboring neurons is generally different from the average firing rate of the output neuron \bar{y} , we assume a homogeneous population of neurons and approximate \bar{y}_{nb} by \bar{y} (See the following for the definition). Thus $m \bar{y}_{\text{nb}} \approx m \bar{y}$ describes the strength of inhibitory input to the output neuron; we model the maturation of inhibition by increasing m at the onset of the CP from $m=0$ to $m=5$ if not stated otherwise.

Learning rule and temporal averages

We introduce two kinds of averages: First $\langle F \rangle$ describes the instantaneous average with respect to the input. Next, \bar{F} describes a long-term running average with a time constant $\tau=1$ (this corresponds to 200 time steps in simulations: see, following), i.e., $\bar{F}(t) = \int_{-\infty}^t e^{-(t-t')/\tau} F(t') dt'$. We consider in this paper an activity dependent learning rule for excitatory synapses within the constraint that synaptic strengths are positive, i.e., $w_i \geq 0$. Synaptic strengths are modified by the following plasticity rule:

$\langle \Delta w_i \rangle = \eta \left[\langle x_i f(y) \rangle - a \langle x_i \rangle \langle f(y) \rangle + \gamma w_i (y_0 - \bar{y}) \right]$, where $f(y) = [y - \theta]_+$ is a threshold-linear function with threshold $\theta = 2$ Hz, $\eta = 1.0 \cdot 10^{-3} \text{ Hz}^{-2}$ is a learning rate, $a = 1.001$ is a constant slightly larger than unity to cause synaptic competition, $\gamma = 2$ Hz is a constant that sets the strength of homeostatic plasticity, and $y_0 = 1.2$ Hz is the set point of the output firing rate for the homeostatic plasticity. If some synaptic strengths are updated to negative values, they are set to 0 instead to guarantee positive strengths. The visual-plus-spontaneous input is presented with probability $q = 0.1$ (Pattern 1) and the spontaneous-only input with probability $1 - q$ (Pattern 2). We update synaptic strength by $q \langle \Delta w_i \rangle_1 + (1 - q) \langle \Delta w_i \rangle_2$ at each time step of duration $\Delta t = 5.0 \cdot 10^{-3}$, where $\langle \rangle_1$ and $\langle \rangle_2$ describes the instantaneous averages under the two patterns. This way of updating synapses using the average change is justified when the learning rate is sufficiently slow compared to the typical duration of each pattern so that synaptic strengths are relatively constant during the repeated sampling of the patterns.

This learning rule is a combination of Hebbian and homeostatic plasticity. The first term $\langle x_i f(y) \rangle - a \langle x_i \rangle \langle f(y) \rangle$ is a Hebbian term that modifies synaptic strengths according to the correlation of pre- and postsynaptic activity. The parameter a controls the size of the RF (the area with nonzero synaptic strengths) and also determines the sensitivity to weak input correlation. If $a = 1$, any nonzero correlation between x_i (presynaptic factor) and $f(y)$ (postsynaptic factor) can potentiate synapses and the size of the RF tends to become large. On the other hand, for $a > 1$, the synaptic strength of an ineffective input is steadily decreased by the Hebbian term, so we need to introduce a lower limit of zero for synaptic strength, $w_i \geq 0$, for all i , so that such synaptic strengths do not take negative values. If a is too large, the learning depends only on the strongest peak of the correlation and becomes insensitive to the correlation width.

The nonlinearity, $f(y) = [y - \theta]_+$, suppresses Hebbian plasticity when the output firing rate is smaller than the Hebbian threshold, θ . To see how inhibition suppresses the contribution of the spontaneous component along with the Hebbian threshold, it is useful to consider a special case of $a = 1$. Because x_i is jointly Gaussian, the Hebbian term in this case is also described (the Bussgang theorem) as

$$\text{Cov} \left[x_i, h \left(\sum_i w_i x_i \right) \right] = g \text{Cov} \left[x_i, \sum_i w_i x_i \right], \text{ where Cov denotes covariance, } h(u) = [u - m \bar{y} - \theta]_+ \text{ is a}$$

threshold-linear function, and the gain factor, $g = \left\langle h' \left(\sum_i w_i x_i \right) \right\rangle$, modulates the magnitude of Hebbian

plasticity. Hence, apart from the non-trivial multiplication factor g that depends on the inhibitory strength, this learning rule is equivalent to the standard Hebbian plasticity rule for a linear neuron model (Oja, 1989; Sejnowski and Tesauro, 1989; Miller and MacKay, 1994). The gain factor g describes, with the current set of

parameters, the percentage of time the output firing rate exceeds the Hebbian threshold, which decreases rapidly when the net input falls below the Hebbian threshold. Hence, weighted by the probability of receiving visual input q , the net Hebbian change of synaptic strengths is described by

$$qg_1\bar{c}_1 + (1-q)g_2\bar{c}_2 \propto \frac{g_2}{g_1}(q\bar{c}_1 + (1-q)\bar{c}_2) + \left(1 - \frac{g_2}{g_1}\right)q\bar{c}_1,$$

where \bar{c}_1 and \bar{c}_2 are the standard linear Hebbian factors ($\text{Cov}[\bar{x}, \bar{w} \cdot \bar{x}]$) and g_1 and g_2 are the gain factors under Pattern 1 (with visual input) and Pattern 2 (without visual input), respectively. We can see that the spontaneous-to-visual ratio of gains, g_1/g_2 , linearly interpolates two extreme conditions: the pre-CP-like modification, $q\bar{c}_1 + (1-q)\bar{c}_2$, where both patterns with and without visual input contribute to Hebbian plasticity, and the CP-like modification, $q\bar{c}_1$, where only the visual pattern contributes to Hebbian plasticity. The switching from pre-CP like modification to CP-like modification robustly arises upon maturation of inhibition if the following conditions are satisfied. During the pre-CP, the spontaneous input should have a larger contribution to plasticity than the sparse visually driven input. Because the input covariance is proportional to the square of the mean input firing rate in the model, this gives the condition that, before inhibitory maturation, $q[(\mu^{(S)})^2 + (\mu^{(V)})^2] < (g_2/g_1)(1-q)(\mu^{(S)})^2$. At the onset of the CP, maturation of inhibition should significantly reduce the contribution of the spontaneous input to plasticity by suppressing the spontaneous-to-visual ratio of gains, g_1/g_2 . This gives the condition that, after inhibitory maturation, $q[(\mu^{(S)})^2 + (\mu^{(V)})^2] \gg (g_2/g_1)(1-q)(\mu^{(S)})^2$. This analysis also well describes the outcome when the constant a is greater than but close to 1 (c.f. Fig. 3F).

This analysis makes clear that the key requirement for the transition from pre-CP to CP plasticity is that the maturation of inhibition should strongly reduce the ratio of gains, $\frac{g_2}{g_1}$. This reduction in ratio will be achieved if inhibition is strong enough to largely eliminate V1 spontaneous activity. It will also be achieved if inhibition is weaker so that weak spontaneous activity remains, but the contribution of that spontaneous activity to Hebbian plasticity is suppressed relative to the contribution of visually-induced activity (or equivalently, the contribution of visually-induced activity is enhanced relative to the contribution of spontaneous activity). The function $f(y)$ we use, which is zero (no contribution to Hebbian plasticity) for y less than the threshold θ , yields relative suppression of plasticity for low firing rates and thus for spontaneous activity relative to visually-induced activity. This could arise if low postsynaptic firing rates yield calcium entry too weak to evoke LTD or LTP as in (Artola et al., 1990) (however, the physiological conditions that induce LTD are controversial (Huang et al., 2005; Hager and Dringenberg, 2010; Froc et al., 2000; Jiang et al., 2003)). Alternatively, relative enhancement of the contribution of visually-induced activity could be achieved if the function $f(y)$ nonlinearly enhances Hebbian synaptic changes for larger postsynaptic firing rates. This has been observed: there is an increase in LTP per postsynaptic spike for higher postsynaptic firing rates (Sjöström et al., 2001).

The second term, $\gamma w_i(y_0 - \bar{y})$, is a homeostatic term that scales synaptic strengths and shift the output firing rate into an appropriate range around y_0 . When the average output rate, \bar{y} , is greater than the set point, y_0 , this term scales down the synaptic strengths; otherwise scales up the strengths to compensate for perturbations to the average activity level. This kind of homeostatic synaptic scaling has widely been observed (Sejnowski and Tesauro, 1989). Although, the Hebbian term alone is typically unstable (strong synapses tend to drive the postsynaptic neuron harder and grow even stronger) and can saturate all nonzero synaptic strengths to their maximum strength, the homeostatic term controls the overall scaling of synapses and sets the strengths proportional to the Hebbian term, yielding a bell shaped tuning curve in the current setting. Thus, the parameter values of the coefficient γ and the set point y_0 only contribute to the scaling of strengths without influencing their relative strength.

The induction of homeostatic plasticity is known to be relatively slow compared to the loss of responsiveness from the deprived eye (Turrigiano et al., 1998; Maffei et al., 2004; Mrsic-Flogel et al., 2007; Kaneko et al., 2008; Ibata et al., 2008). According to these experiments, we introduced a long-term running average, \bar{y} , of the output firing rate. Because the homeostatic term compares \bar{y} to the set point, it takes about the averaging time constant τ to adapt its strength after a sudden change in the input statistics caused by visual deprivations.

However, if this adaptation rate $1/\tau$ is too slow compared to the learning speed of the synaptic update rule, this generally causes the oscillation of synaptic strengths – the homeostatic term drive synapses beyond the set point until \bar{y} has been updated - and learning does not converge. Note that replacing the long-term averaged firing rate, \bar{y} , by an instantaneous average, $q\langle y \rangle_1 + (1-q)\langle y \rangle_2$, does not alter the fixed point of the learning rule, and, hence, does not alter main results presented in this paper.

S2d. Receptive field width, Response strength, Contra-bias index, and Gain

In Fig. S3, we plotted several quantities, defined as follows. The width of the RF is the standard deviation of the spatial distribution of synaptic strength from each eye, i.e., $\text{Width}_e = \sqrt{\frac{\sum_{i \in \{i|e_i=e\}} R_i^2 w_i}{\sum_{i \in \{i|e_i=e\}} w_i}}$, where i

indexes the inputs and R_i is the retinotopic distance from the RF center and $e_i \in \{C, I\}$ the eye (either contralateral or ipsilateral) of the i^{th} input. The response strength from each eye was measured by the maximum synaptic strength times the density of synapses (N_e synapses in the retinotopic area of 4), i.e.,

$\text{Response}_e = \max_{i \in \{i|e_i=e\}} w_i * (N_e / 4)$. We take response strength proportional to the peak of the RF, rather than

the sum over the RF, because we are comparing principally to measurements (Smith and Trachtenberg, 2007) from optical imaging, in which each pixel measures response over a local collection of cells, using narrow bars (relative to RF sizes) for stimuli, which will drive only a small portion of an RF at any given time. Given nonlinear input/output functions in which stronger inputs yield supralinearly stronger responses, due to threshold and/or noise, the strongest responses, driven by neurons stimulated on their RF peaks, will tend to dominate the optical response. The contra-bias index was $\text{CBI} = \text{Response}_c / (\text{Response}_c + \text{Response}_i)$. The gain was the percentage of time the output activity exceeded the Hebbian threshold, as described above in the section describing the learning rule.

S2e. Extracellular recording from freely behaving mice

To acquire extracellular spikes, tetrodes of four nichrome wires (13 μm) were stereotaxically implanted into the binocular zone of V1 (at lambda, 3.0 mm lateral to the midline). Signals from each electrode were band-pass filtered (0.15 – 9 KHz) and digitized at 40 KHz (MAP system, Plexon, TX). The electrode tip positions were verified histologically using Nissl staining. Isolation of single units was performed manually using a multi-dimensional cluster cutting software (Off-line sorter, Plexon). Clusters of spikes with a clear refractory period (> 2 ms) were included in the analysis. Putative pyramidal neurons were discriminated by waveform features (spike width, shape, and autocorrelation (Barthó et al., 2004)). Periodic full-field visual stimuli were provided by a light emitting diode (LED, peak 510nm, 5000 mcd) on the ceiling of the recording chamber. Peri-event time histograms were made using NeuroExplorer (Nex Technology, MA). Diazepam (in 50% propylene glycol/saline; 20 mg/kg i.p.) was administered acutely under halothane anesthesia. Chronic-treated *GAD65-KO* (KOc) mice received repeated, daily intracerebroventricular injections over six consecutive days around P28 and then later were implanted with electrodes for recording as adults ($> P60$).

S2f. LTD Experiments

Electrophysiological recordings *in vitro* were obtained at P24–35, using our standard methods (Renger et al, 2002). In brief, coronal slices (350 μm) of binocular mouse visual cortex were cut and incubated (>1 h, 33°C) in equilibrated (95% $\text{O}_2/5\%$ CO_2) artificial cerebrospinal fluid (ACSF), containing (in mM): 119 NaCl, 2.5 KCl, 1.3 MgSO_4 , 10 NaH_2PO_4 , 26.2 NaHCO_3 , 2.5 CaCl_2 , and 11 glucose. Slices were transferred to a recording chamber superfused with the same ACSF (33°C). A two-pathway experiment was designed in which standard coronal slices were split from white matter up to layer 4 (Renger et al, 2002). Half-maximal field potentials evoked from layer 4 by a bipolar glass stimulating electrode filled with ACSF were recorded through a glass electrode filled with 1 M NaCl (1–3 M Ω) in layer 2/3. The independence of inputs stimulated from either side of the cut was confirmed by additive responses in layer 2/3 showing no interaction when closely paired stimuli (< 50 msec) were presented across the two sites. Stable, baseline responses were recorded (>25 min, 0.1 Hz) before applying low-frequency stimulation (LFS; 900 pulses, 1 Hz) to induce LTD on the test pathway, while the other, control pathway was switched off during conditioning. To determine saturated LTD levels, repeated LFS (3x) was applied to one pathway in a separate set of un-split WT and KO slices.

Supplemental Figures

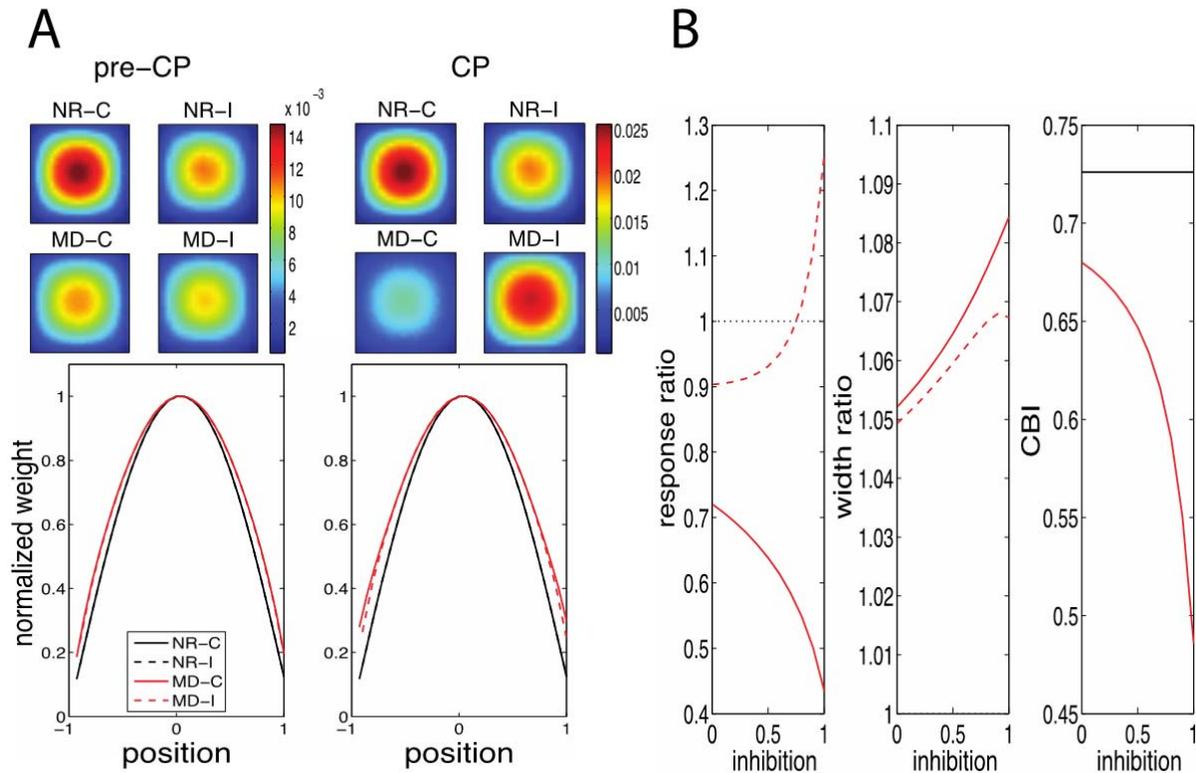


Fig. S1: Results of modeling a linear postsynaptic neuron with a simple plasticity rule, Related to Fig. 2-3. Results of modeling a linear postsynaptic neuron with a simple activity-dependent plasticity rule, for which the final receptive field is proportional to the principle eigenvector of the input covariance matrix. (A) The final receptive field from either eye (C: contralateral; I: ipsilateral) is shown under normal rearing (NR) and under monocular deprivation (MD) of the contralateral eye, during the pre-CP ($f=0$) and the CP ($f=1$). The bottom panels show peak-normalized one-dimensional horizontal sections of the two-dimensional receptive fields shown in the upper panels (NR: black; MD: red; contra eye: solid; ipsi eye: dashed). (B) Quantifications of the MD effect on the receptive fields, as a function of the strength of inhibition f . The left two panels plot the ratio of the value under MD to the value under NR of the peak response (left) and the receptive field width (middle). The right panel shows the contralateral-bias index (CBI) under NR (black) and under MD (red). See Supplemental Computational and Experimental Procedures, S2d for definitions of response, width, and CBI.

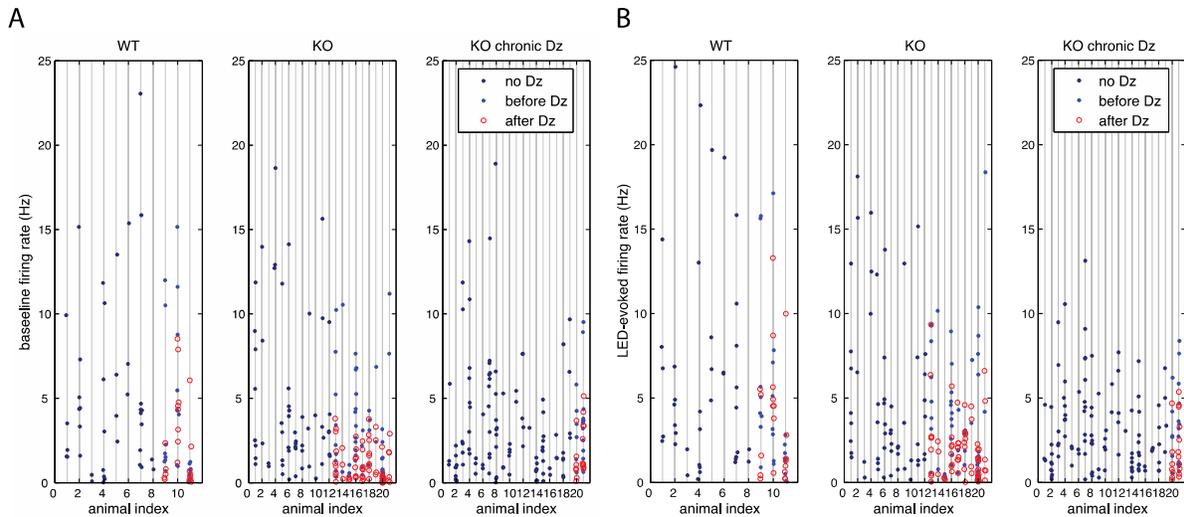


Fig. S2: Baseline and LED-evoked firing rates, Related to Fig. 4. Baseline firing rates (A) and LED-evoked firing rates (B) in Wild-type mice (WT), *GAD65*-KO mice (KO), and *GAD65*-KO mice that were chronically treated with diazepam (KO chronic Dz). Many animals are recorded without any acute diazepam treatment (dark-blue dots). In order to study the effect of enhanced inhibition on visual response ratio, a small subset of animals was recorded before (blue dots) and after (red circles) acute administration of diazepam. Each mark shows firing rate from one recording unit. Note that the sets of recording units before and after the diazepam treatment were not always identical because some recording units were lost/gained during the acute administration of diazepam.

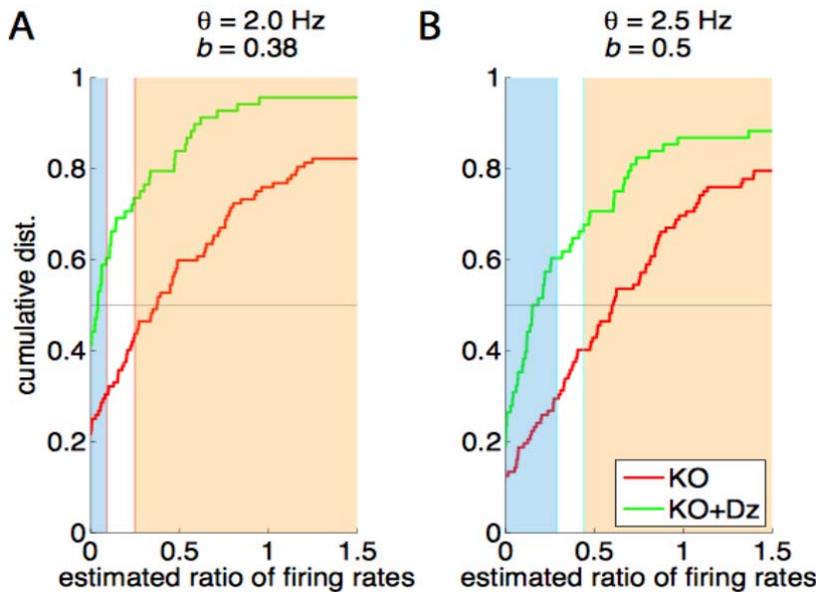


Fig. S3: The cumulative distributions of the bias-corrected spontaneous-to-visual ratio of firing rates, Related to Fig. 4.

The cumulative distributions of the bias-corrected spontaneous-to-visual ratio of firing rates in *GAD65*-KO mice (KO) and *GAD65*-KO mice acutely treated with diazepam (KO+Dz), for two different choices of underlying parameters. The underlying true spontaneous-to-visual ratio of firing rates were estimated, as described in Sec. S1b, from the experimental data presented in Fig. 4D. The estimation was based on the assumed frequency, $q=0.1$, of visual input under behaving conditions, the unknown effectiveness, b ($0 < b < 1$), of the LED stimulation in

comparison to the preferred visual stimulus for individual cells, and the unknown Hebbian threshold, θ . The estimation of the true spontaneous-to-visual ratio of firing rates for majority of cells fell in a regime that reproduced pre-CP and CP plasticity for KO and KO+Dz animals, respectively, if the parameters b and θ were chosen appropriately. There was a range of b that achieves this for a given choice of Hebbian threshold values, of which one in the middle of the range was chosen for each choice of θ illustrated here (c.f. Fig. 3E), $\theta=2.0$ Hz (A) and $\theta=2.5$ Hz (B). Cells in the light-blue and yellow areas showed large enough and almost no ocular dominance shift, respectively, in the simulation of Fig. 3E.

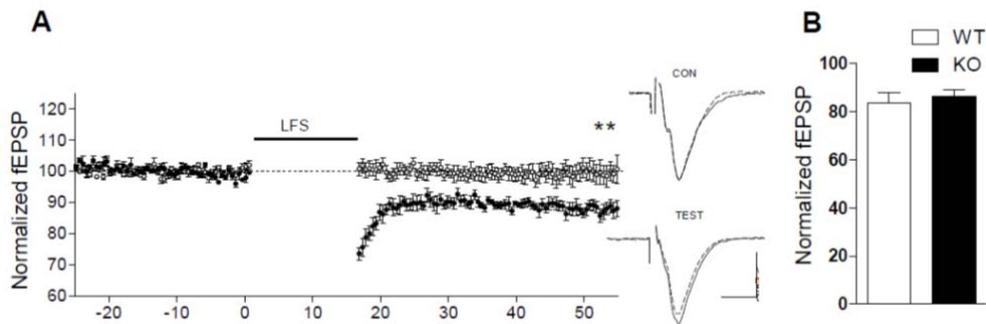


Fig. S4: LTD is intact in *GAD65*-KO mice, Related to Discussion

(A) Only the test pathway was persistently depressed by LFS (black circles), revealing a clear LTD in *GAD65*-KO mouse visual cortex ($n = 6$ slices). Moreover, unaltered control responses (white circles; ** $p < 0.01$, t test 40 min post-LFS) verified independence of the two pathways as well as general health of the KO slices, confirming our original observations (Hensch et al, 1998). Scale bar, 500 μ V, 5 msec. (B) Equal saturation of LTD by repeated LFS (3x) to one pathway in WT and *GAD65*-KO mice ($n = 7$ slices each, $p < 0.6$, t test). Error bars are SEM.

Supplemental References

Artola, A., Bröcher, S., and Singer, W. (1990). Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. *Nature* 347, 69-72.

Barthó, P., Hirase, H., Monconduit, L., Zugaro, M., Harris, K.D., and Buzsáki, G. (2004). Characterization of neocortical principal cells and interneurons by network interactions and extracellular features. *J Neurophysiol* 92, 600-608.

Desai, N.S., Cudmore, R.H., Nelson, S.B., and Turrigiano, G.G. (2002). Critical periods for experience-dependent synaptic scaling in visual cortex. *Nat Neurosci* 5, 783-789.

Faguet, J., Maranhao, B., Smith, S.L., and Trachtenberg, J.T. (2009). Ipsilateral eye cortical maps are uniquely sensitive to binocular plasticity. *J Neurophysiol* 101, 855-861.

Froc, D.J., Chapman, C.A., Trepel, C., and Racine, R.J. (2000). Long-term depression and depotentiation in the sensorimotor cortex of the freely moving rat. *J Neurosci* 20, 438-445.

Hager, A.M., and Dringenberg, H.C. (2010). Assessment of different induction protocols to elicit long-term depression (LTD) in the rat visual cortex in vivo. *Brain Res* 1318, 33-41.

Hensch, T.K., Fagiolini, M., Mataga, N., Stryker, M.P., Baekkeskov, S., and Kash, S.F. (1998). Local GABA circuit control of experience-dependent plasticity in developing visual cortex. *Science* 282, 1504-1508.

- Huang, Y.Z., Edwards, M.J., Rounis, E., Bhatia, K.P., and Rothwell, J.C. (2005). Theta burst stimulation of the human motor cortex. *Neuron* 45, 201-206.
- Ibata, K., Sun, Q., and Turrigiano, G.G. (2008). Rapid synaptic scaling induced by changes in postsynaptic firing. *Neuron* 57, 819-826.
- Jiang, B., Akaneya, Y., Hata, Y., and Tsumoto, T. (2003). Long-term depression is not induced by low-frequency stimulation in rat visual cortex in vivo: a possible preventing role of endogenous brain-derived neurotrophic factor. *J Neurosci* 23, 3761-3770.
- Kaneko, M., Stellwagen, D., Malenka, R.C., and Stryker, M.P. (2008). Tumor necrosis factor-alpha mediates one component of competitive, experience-dependent plasticity in developing visual cortex. *Neuron* 58, 673-680.
- Maffei, A., Nelson, S.B., and Turrigiano, G.G. (2004). Selective reconfiguration of layer 4 visual cortical circuitry by visual deprivation. *Nat Neurosci* 7, 1353-1359.
- Miller, K.D., and MacKay, D.J.C. (1994). The role of constraints in Hebbian learning. *Neural Comput* 6, 100-126.
- Mrsic-Flogel, T.D., Hofer, S.B., Ohki, K., Reid, R.C., Bonhoeffer, T., and Hübener, M. (2007). Homeostatic regulation of eye-specific responses in visual cortex during ocular dominance plasticity. *Neuron* 54, 961-972.
- Oja, E. (1989). Neural networks, principal components, and subspaces. *International journal of neural systems* 1, 61-68.
- Ohshiro, T., and Weliky, M. (2006). Simple fall-off pattern of correlated neural activity in the developing lateral geniculate nucleus. *Nat Neurosci* 9, 1541-1548.
- Renger, J.J., Hartman, K.N., Tsuchimoto, Y., Yokoi, M., Nakanishi, S., and Hensch, T.K. (2002). Experience-dependent plasticity without long-term depression by type 2 metabotropic glutamate receptors in developing visual cortex. *Proc Natl Acad Sci U S A* 99, 1041-1046.
- Sato, M., and Stryker, M.P. (2008). Distinctive features of adult ocular dominance plasticity. *J Neurosci* 28, 10278-10286.
- Sejnowski, T.J., and Tesauro, G. (1989). The Hebb rule for synaptic plasticity: algorithms and implementations. In J. Byrne and W. O. Berry, eds. *Neural Models of Plasticity* (New York: Academic Press), pp. 94-103.
- Sjöström, P.J., Turrigiano, G.G., and Nelson, S.B. (2001). Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron* 32, 1149-1164.
- Smith, S.L., and Trachtenberg, J.T. (2007). Experience-dependent binocular competition in the visual cortex begins at eye opening. *Nat Neurosci* 10, 370-375.
- Stellwagen, D., and Malenka, R.C. (2006). Synaptic scaling mediated by glial TNF-alpha. *Nature* 440, 1054-1059.
- Turrigiano, G.G., Leslie, K.R., Desai, N.S., Rutherford, L.C., and Nelson, S.B. (1998). Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature* 391, 892-896.
- Vinje, W.E. (2000). Sparse Coding and Decorrelation in Primary Visual Cortex During Natural Vision. *Science* 287, 1273-1276.
- Weliky, M., and Katz, L.C. (1999). Correlational Structure of Spontaneous Neuronal Activity in the Developing Lateral Geniculate Nucleus in Vivo. *Science* 285, 599-604.