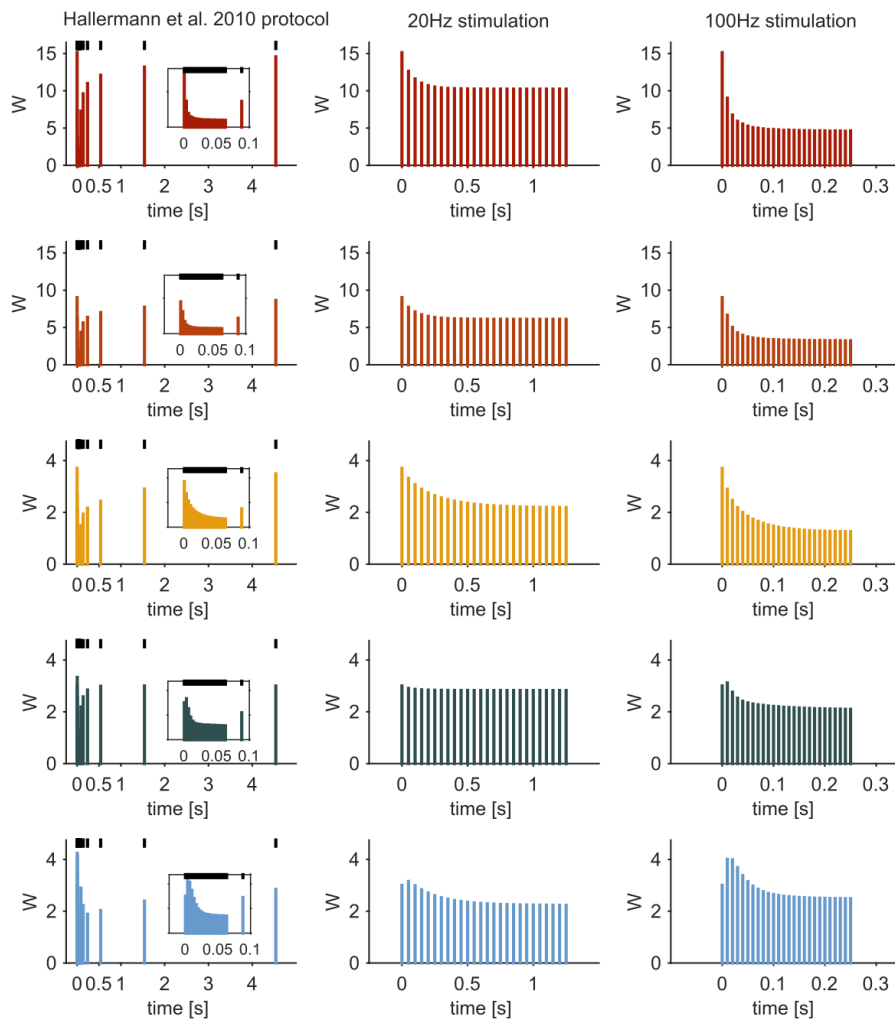
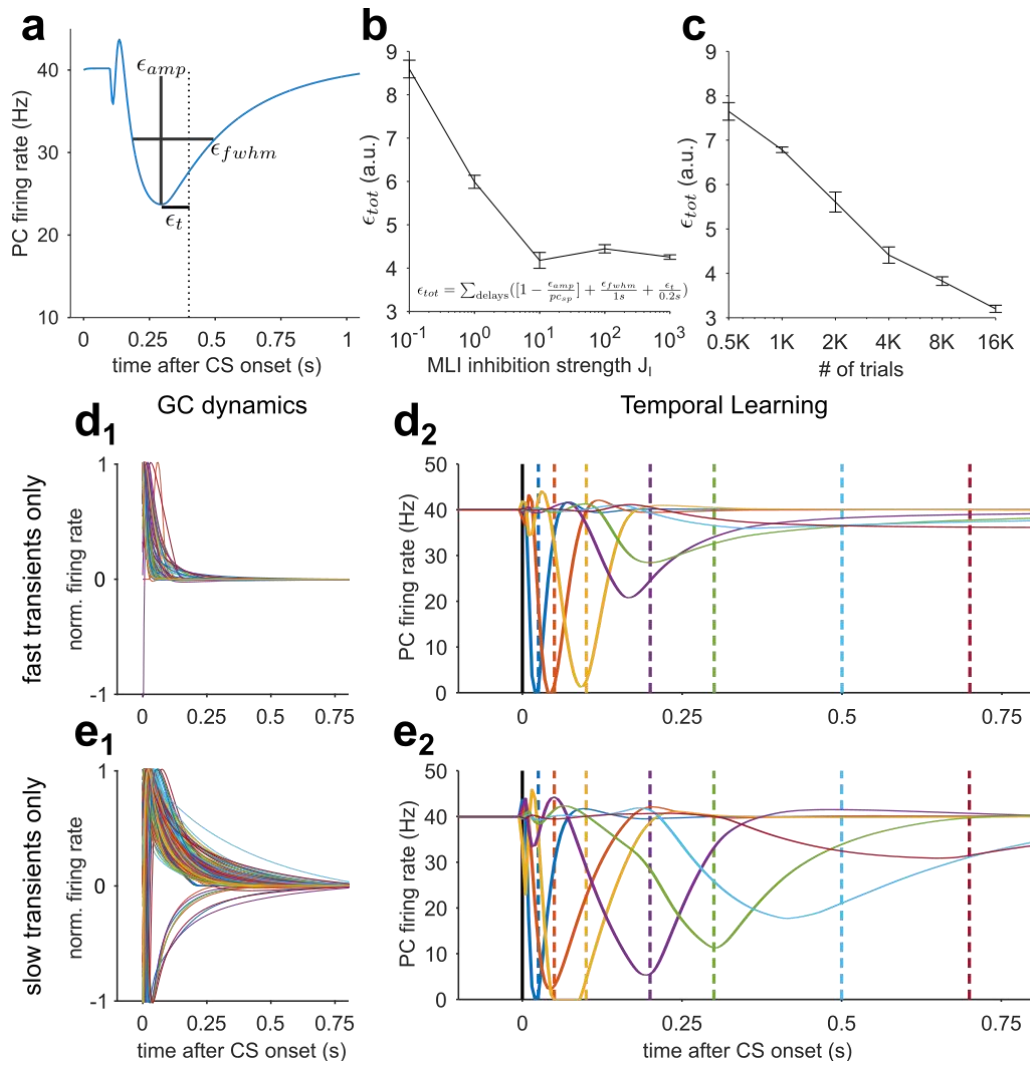


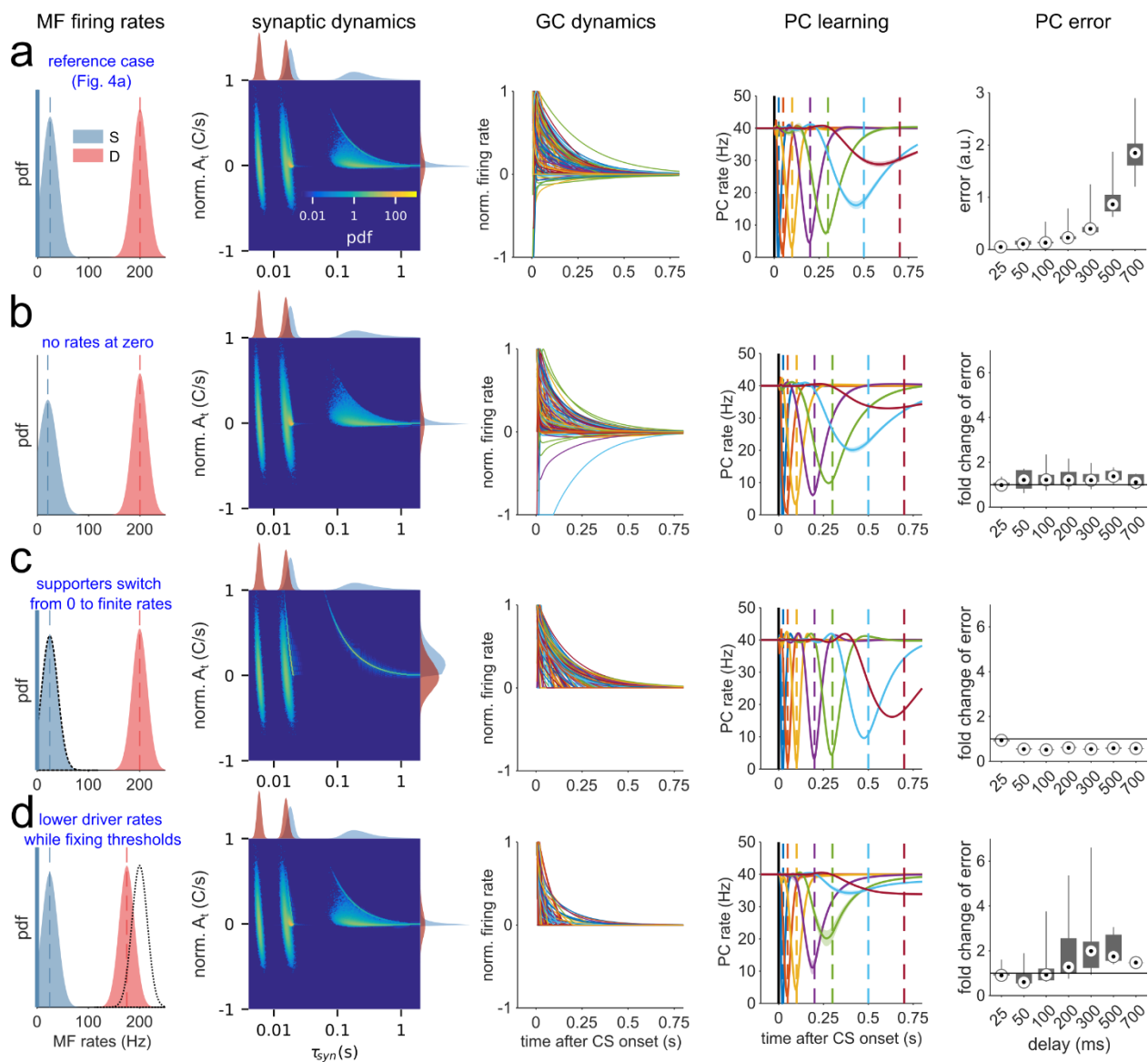
## Supplementary Figures



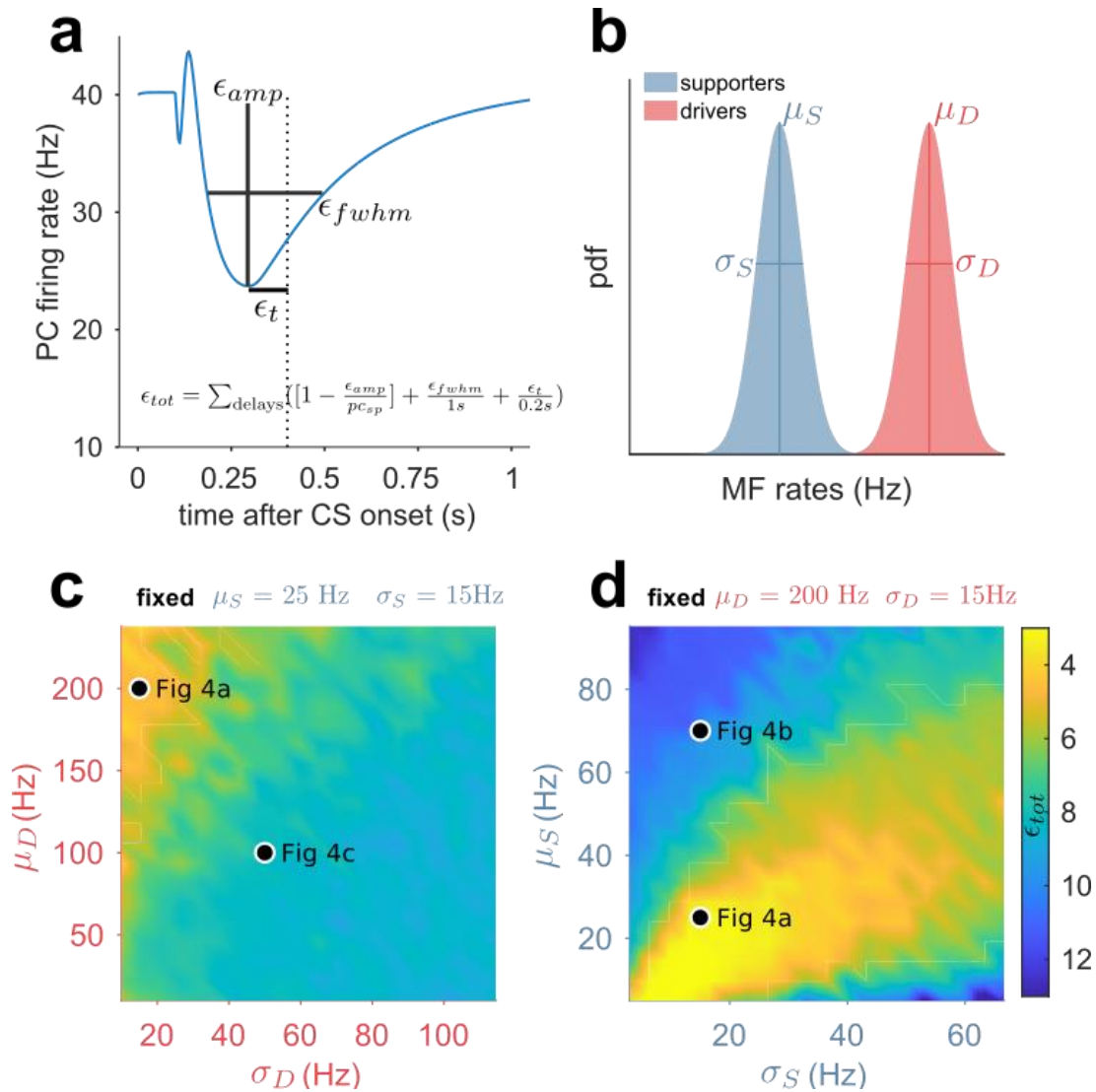
**Supplementary Figure 1** Behavior of model synapses subjected to different stimulation protocols. Each row represents one of the five synapse types from **Fig. 1**. First column: average synaptic weight in response to 300 Hz train followed by increasing intervals ranging from 25 ms to 5 s as in ref.<sup>38</sup>. Inset: zoom on 300 Hz train. Second and third column: average synaptic weight in response to trains of 26 stimuli at 20 Hz and 100 Hz, respectively, similar to ref.<sup>39</sup>



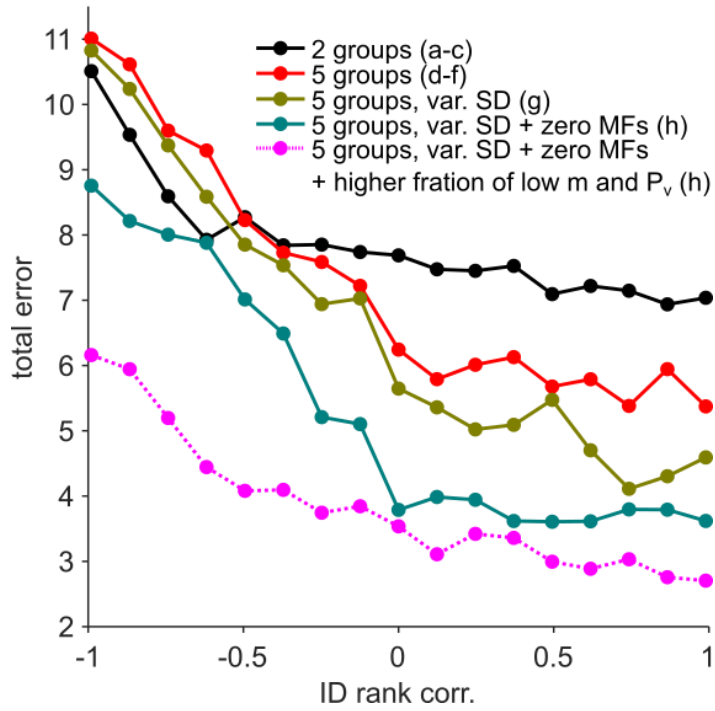
**Supplementary Figure 2 (a)** Definition of error components used to assess learning performance. **(b)** Learning error as a function of MLI-PC inhibition strength. The total error is computed from the individual components according to the formula shown. Black line represents the average over 20 realizations of the network in **Fig.2**. Error bars are SEM. **(c)** Same as **b** for learning error as a function of number of learning iterations. **(d<sub>1</sub>)** As in **Fig.2e**, but without GCs whose transients decayed to 10% of their peak values in more than 150 ms. **(d<sub>2</sub>)** Learning performance when using the temporal basis from panel **d<sub>1</sub>**. **(e<sub>1</sub>)** As in **Fig.2e**, but without GCs whose transients decayed to 10% of their peak values in less than 150 ms. **(e<sub>2</sub>)** Learning performance when using the temporal basis from panel **e<sub>1</sub>**. Note that learning of short delays was not completely abolished because our manipulations did not affect the rising phases of the GC transients.



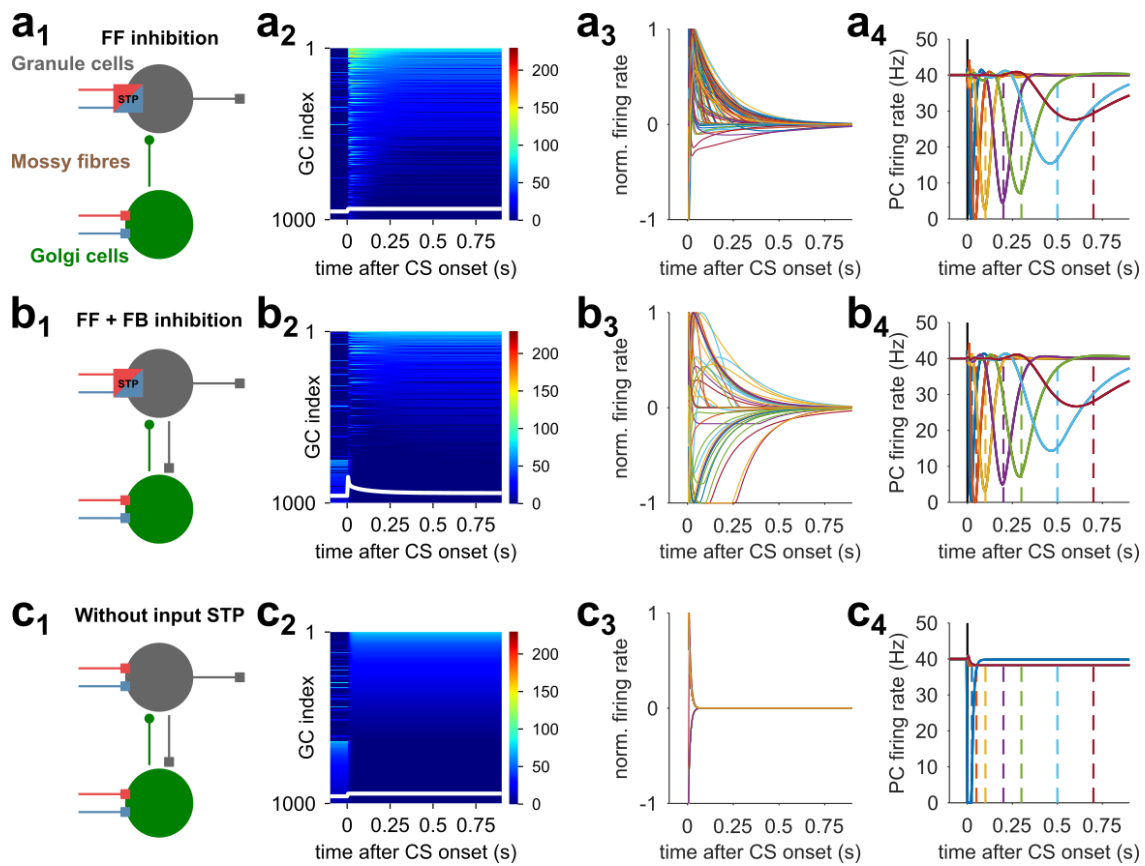
**Supplementary Figure 3** Additional examples of MF rate distributions and their impact on learning. **(a)** Same as **Fig. 4a** (i.e. reference case, see legend of **Fig. 4** for details). **(b)** Without zero firing rates. **(c)** Simulations in which all supporter inputs are set to zero before the CS and switch to finite firing rates during the CS. **(d)** Simulations in which GC thresholds are set as in **a** (blue and dashed firing rate distributions), but eyelid conditioning is carried out with lower driver inputs (blue and red firing rate distributions).



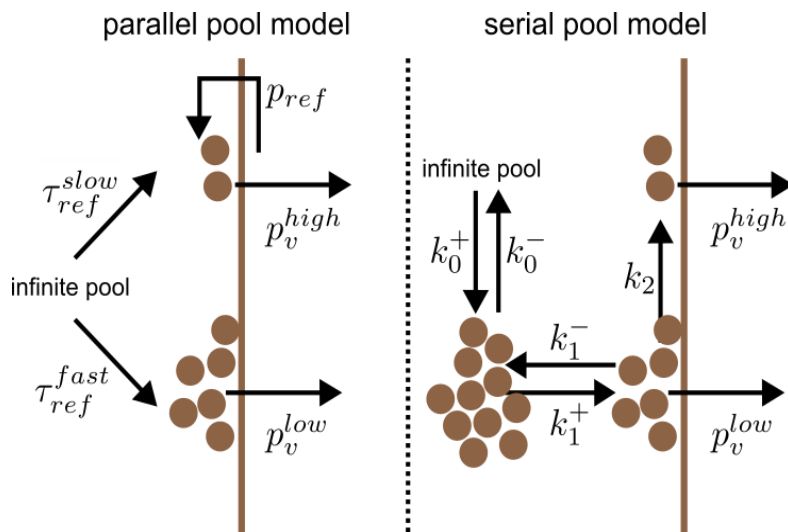
**Supplementary Figure 4** Scan of MF firing rate parameters show distinct roles for driver and supporter inputs. **(a)** Definition of error components used to assess learning performance. The total error is computed from the individual components according to the formula shown. **(b)** Definition of the MF firing rate parameters.  $\mu_D$ ,  $\mu_S$  and  $\sigma_D$ ,  $\sigma_S$  denote means and standard deviations of the MF firing rate distributions of drivers (red) and supporters (blue) respectively. **(c)** Scan over driver parameters while keeping supporter parameters fixed. **(d)** Scan over supporter parameters while keeping driver parameters fixed. In **c** and **d** the total learning error is color coded and parameter configurations corresponding to rows in **Fig. 4** are indicated by black dots.



**Supplementary Figure 5** Average total error over seven delay intervals for varying rank-correlation between the  $m$  category and the  $p_v$  category for all scenarios shown in **Fig. 5**.



**Supplementary Figure 6** Eyelid learning is not significantly affected by simple GoC feedback. **(a<sub>1</sub>)** Scheme of CC input layer with GoC inhibition that acts in a purely feed-forward manner. This configuration is functionally identical to the reduced model used in the main text. **(a<sub>2</sub>, a<sub>3</sub>)** GC temporal basis (as in Fig. 2c). White line in **a<sub>2</sub>** indicates GoC activity. **(a<sub>4</sub>)** PC eyelid response learning. **(b<sub>1-4</sub>)** Same as row **a**, but with GC-GoC feedback connections. **(c<sub>1-4</sub>)** Same as row **b**, but without MF-GC STP.



**Supplementary Figure 7** Scheme comparing the parallel pool model used here with the serial pool model from ref.<sup>38</sup>. The symbols  $k_0^+$ ,  $k_0^-$ ,  $k_1^+$ ,  $k_1^-$  and  $k_2$  are rate constants described in ref.<sup>38</sup>.