Cell-type-specific synaptic scaling mechanisms differentially contribute to associative learning

Fabio Veneto[1,*], Ayça Kepçe[1,2,*], Yue Kris Wu[1,3,#], Julijana Gjorgjieva[1,#]

[1]School of Life Sciences, Technical University of Munich, Freising, Germany;
[2]Elite Master Program in Neuroengineering, Technical University of Munich, Munich, Germany
[3]Zuckerman Mind Brain Behavior Institute, Columbia University, New York, USA
$*$ These authors contributed equally to this work.
[#] Co-senior and co-corresponding authors
Email: yw4297@columbia.edu; gjorgjieva@tum.de.

Abstract

Excitatory synaptic scaling regulates network dynamics by proportionally adjusting excitatory synaptic strengths after sensory perturbations. During associative learning, blocking excitatory scaling in conditioned taste aversion paradigms prolongs generalized aversive responses and delays memory specificity. Recent evidence also implicates inhibitory synaptic scaling in the regulation of network dynamics. Specifically, parvalbumin (PV)-expressing inhibitory neurons, targeting perisomatic regions of excitatory (E) pyramidal neurons, and somatostatin (SST)-expressing neurons, targeting distal dendrites, exhibit distinct scaling responses. This leaves open the question of how complex plasticity mechanisms regulate recurrent excitatory-inhibitory circuit dynamics in associative learning. Using computational approaches, we demonstrate that Hebbian plasticity drives memory generalization to novel stimuli not presented during conditioning. Following conditioning, diverse synaptic scaling mechanisms progressively induce memory specificity, which can be regulated by top-down inputs. Our results reveal that, in the absence of excitatory scaling, PV-to-E scaling can effectively compensate and rescue memory specificity, highlighting the presence of degenerate mechanisms in the brain. Notably, in the process of establishing memory specificity, excitatory scaling and PV-to-E scaling function synergistically, while concurrently opposing SST-to-E scaling. The synergistic and antagonistic plasticity mechanisms are orchestrated to shape the temporal evolution of memory representations, from generalized to precise.

Introduction

Synaptic scaling is considered a crucial homeostatic synaptic plasticity mechanism that adjusts the strength of all incoming synapses to a neuron to stabilize network dynamics in response to sensory perturbations (Turrigiano et al., 1998; Turrigiano, 2008; Pozo and Goda, 2010; Tetzlaff

et al., 2011; Wu et al., 2020; Wen and Turrigiano, 2024). Typically studied for synapses between excitatory neurons, this process either downscales or upscales the synapses to compensate for hyperactive or hypoactive activity, respectively (Turrigiano et al., 1998; Kim et al., 2012; Keck et al., 2013; Torrado Pacheco et al., 2021). Unlike Hebbian plasticity, which can occur from seconds to minutes, synaptic scaling operates on a much slower timescale, unfolding over hours to days (Turrigiano et al., 1998; Ibata et al., 2008; Watt, 2010; Keck et al., 2017). Beyond the homeostatic role of excitatory synaptic scaling in compensating for sensory perturbations, recent studies have revealed the importance of excitatory synaptic scaling in associative learning (Wu et al., 2021). In a classical conditioned taste aversion (CTA) paradigm (Figure 1A), by pairing an aversive unconditioned stimulus (US) with a conditioned stimulus (CS), mice learned the association between the CS and aversion immediately after the conditioning. Four hours after conditioning, mice displayed generalized aversion to a novel test stimulus (TS), a tastant absent during the conditioning. When tested after 24h and 48h with the TS, the generalized aversive behavior diminished, and mice exhibited aversion exclusively to the CS, reflecting the formation of memory specificity. Blocking excitatory synaptic scaling significantly prolonged the generalized aversive behavioral response, with mice continuing to exhibit aversion to TS even 24 hours post-conditioning (Figure 1A). Nonetheless, after blocking excitatory synaptic scaling, mice developed memory specificity after 48h post-conditioning, suggesting that other mechanisms may complementarily achieve the specific refinement of memory representations.

In addition to excitatory synaptic scaling, inhibitory synaptic scaling has also been found to regulate network dynamics in response to sensory perturbations (Kilman et al., 2002; Swanwick et al., 2006; Prestigio et al., 2021). Notably, the scaling of inhibitory synapses is target-dependent (Prestigio et al., 2021), whereby hyperactivity in excitatory pyramidal neurons (E) induces upscaling of inhibitory synapses at perisomatic regions of excitatory neurons while unexpectedly downscaling of inhibitory synapses at dendritic regions. These connectivity preferences have been related to molecularly distinct interneuron subtypes (Tremblay et al., 2016) (Figure 1B). For instance, parvalbumin (PV)-expressing inhibitory neurons primarily innervate perisomatic regions, whereas somatostatin (SST)-expressing inhibitory neurons predominantly target distal dendritic regions (Lazarus and Huang, 2011; Hioki et al., 2013; Dorsett et al., 2021). Yet, how these inhibitory synaptic scaling mechanisms interact with the well-established excitatory synaptic scaling during associative learning remains elusive.

Here, combining analytical calculations and numerical simulations, we demonstrate that rapid Hebbian plasticity drives memory generalization to novel test stimuli that are absent during conditioning. Following conditioning, we find that different forms of synaptic scaling regulate neural dynamics, progressively inducing memory specificity over time. Our findings highlight the critical role of different forms of synaptic scaling in achieving precise memory representations and propose a role for top-down inputs in modulating associate learning. When excitatory scaling is absent, memory

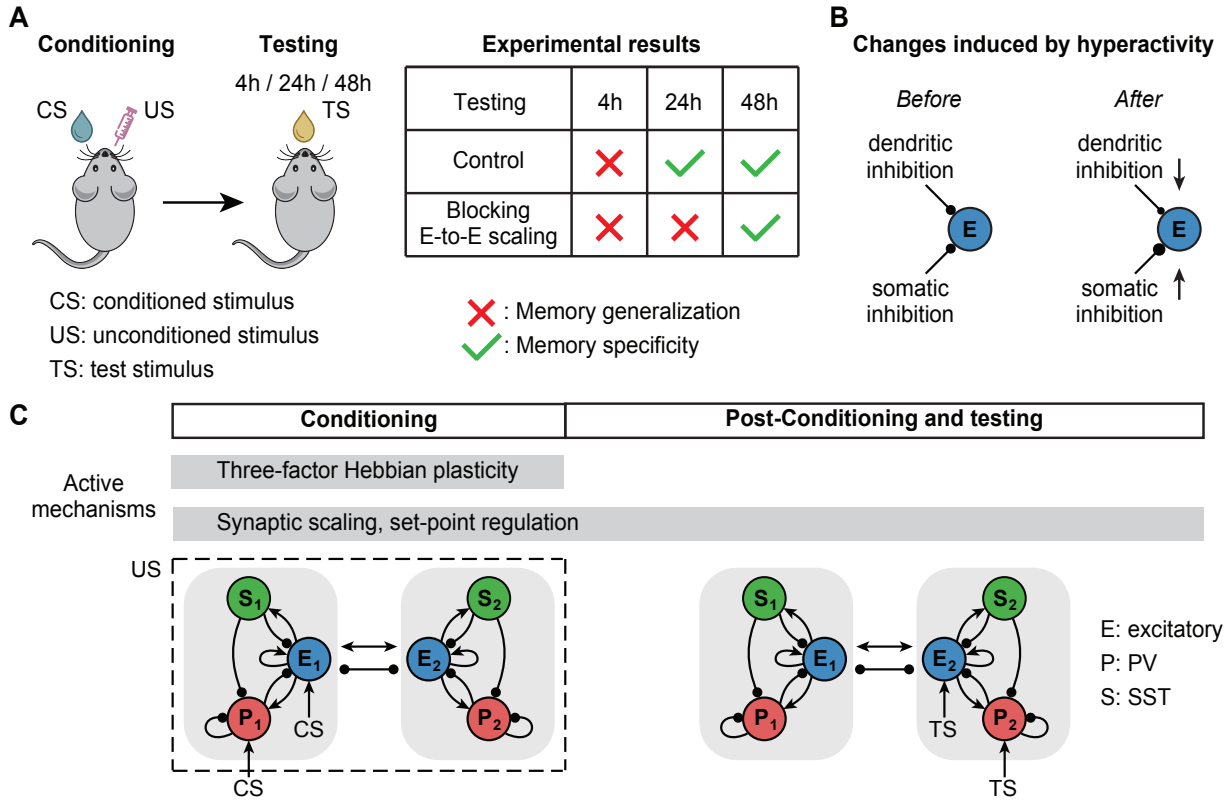

Fig. 1. Experimental paradigm and computational framework for associative learning. **A.** Conditioned taste aversion (CTA) paradigm applied in Wu et al. (2021). Conditioning is induced by pairing an aversive unconditioned stimulus (US) with a conditioned stimulus (CS) (left). Memory generalization and specificity are evaluated by measuring the mouse's aversive behavioral response to a novel test stimulus (TS) at either 4h, 24h, or 48h (left). Mice exhibit an aversive behavioral response to TS at 4h but not at 24h and 48h, indicating a switching from memory generalization to memory specificity (right). When blocking excitatory scaling, memory generalization persists at 24h but diminishes by 48h (right). **B.** Target-specific inhibitory synaptic scaling reported in Prestigio et al. (2021). Hyperactivity in postsynaptic excitatory neurons induces a downscaling of dendritic inhibition while upscaling somatic inhibition. **C.** Schematic of network model with two subnetworks. Each subnetwork consists of one excitatory, one PV and one SST population. Different subnetworks are tuned to different stimuli corresponding to different tastants in the conditioned taste aversion experiments. During conditioning, excitatory (E1) and PV (P1) populations in subnetwork 1 receive additional inputs corresponding to a CS, while the US is present. During the test period, excitatory (E2) and PV (P2) populations in subnetwork 2 receive additional inputs corresponding to a TS. Three-factor Hebbian plasticity operates during conditioning, whereas synaptic scaling and set point regulation mechanisms are active during both conditioning and post-conditioning phases.

specificity can be rescued by PV-to-E scaling, indicating the existence of degenerate mechanisms in the brain. We find that excitatory scaling and PV-to-E scaling work synergistically while counteracting the effects of SST-to-E scaling. This intricate interplay between synergistic and antagonistic plasticity mechanisms drives the temporal evolution of memory specificity, facilitating a smooth transition from generalized to specific memory representations.

Results

To investigate how different plasticity mechanisms – rapid Hebbian and slower forms of synaptic scaling – interact with each other and give rise to memory specificity in associative learning, we developed a rate-based recurrent network model consisting of two interconnected subnetworks. Each subnetwork includes one excitatory (E) population and two distinct inhibitory populations: PV and SST (see Methods, Figure 1C). We assume that different subnetworks are tuned to different stimuli corresponding to different tastants in the conditioned taste aversion experiments. Inspired by experimental studies indicating that PV inhibitory neurons primarily innervate perisomatic regions, while SST inhibitory neurons predominantly target distal dendritic regions, we modeled somatic inhibition to the excitatory population as coming from the PV population and dendritic inhibition as coming from the SST population (Figure 1C) (Lazarus and Huang, 2011; Pfeffer et al., 2013; Dorsett et al., 2021). The network connectivity was designed to incorporate previously reported experimental features, including the absence of inhibitory connections from PV and SST interneurons to SST interneurons (Pfeffer et al., 2013).

Three-factor Hebbian plasticity strengthens excitatory-to-excitatory connections during conditioning

To model the conditioning procedure in the conditioned taste aversion paradigm, the excitatory (E_1) and PV (PV_1) populations in subnetwork 1 receive additional inputs to represent the conditioned stimulus (CS) (Figure 2A) (Ji et al., 2015). Inspired by experimental studies demonstrating that reward or punishment plays a crucial role in learning (Pawlak, 2010; Gerstner et al., 2018), we applied a three-factor Hebbian learning rule to update the E-to-E connection strength during conditioning:

$$\tau_{hebb}\frac{dw_{E_iE_j}}{dt} = \eta r_{E_j}(r_{E_i} - r_{E_i}^{bs}) \qquad i,j \in \{1,2\} \tag{1}$$

$$\eta = \begin{cases} 1 & \text{in the presence of unconditioned stimulus} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where τ_{hebb} is the time constant of Hebbian plasticity, r_{E_i} denotes the activity of the excitatory population in subnetwork i with the superscript 'bs' representing the baseline activity before conditioning, i, j representing the indices of subnetworks. The presence of the aversive unconditioned stimulus (US) determines the third factor η and serves as a gate for Hebbian plasticity, enabling plasticity during the conditioning phase while disabling it elsewhere.

During the simulation of conditioning, the CS leads to an increase in the excitatory activity in subnetwork 1 (r_{E_1}) (Figure 2B). Despite not being directly stimulated by the CS, the excitatory activity in subnetwork 2 (r_{E_2}) also increases, albeit to a lesser extent, due to recurrent excitatory connections between E_1 and E_2 (Figure 2B). During conditioning, in the presence of the US, E-to-E connection

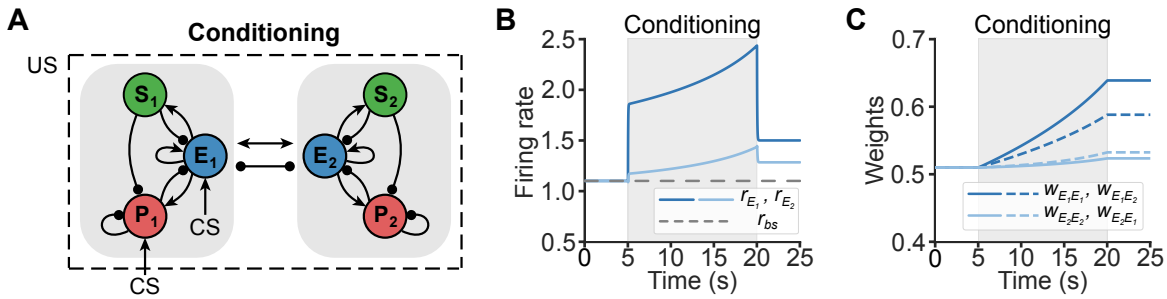

Fig. 2. Hebbian plasticity enhances excitatory activity and strengthens E-to-E connections during conditioning. **A.** Network schematic of the conditioning phase. During conditioning, E and PV populations of subnetwork 1 receive additional inputs that correspond to the conditioned stimulus. During this phase, the unconditioned stimulus is present, modulating Hebbian plasticity. **B.** Activity of excitatory population in subnetwork 1 (r_{E_1}) and subnetwork 2 (r_{E_2}). Conditioning is applied during the interval from 5 to 20s by increasing the inputs to E and PV populations in subnetwork 1. The dashed line represents the baseline activity level measured before conditioning. **C.** Excitatory to excitatory connection strength during conditioning. Different connections are indicated by the differently colored lines.

strengths increase through Hebbian plasticity, with the strongest enhancement observed in the connection strength within the excitatory population of subnetwork 1 ($w_{E_1E_1}$) (Figure 2C).

Memory undergoes transient generalization caused by Hebbian plasticity before gradually achieving specificity

Together with Hebbian plasticity acting on excitatory-to-excitatory synapses during conditioning, we incorporated synaptic scaling at the connections from E to E synapses, i.e., excitatory scaling. This is consistent with experimental findings that hyperactivity (hypoactivity) of excitatory neurons leads to downscaling (upscaling) of E-to-E synapses. Synaptic scaling adjusts synaptic weights to maintain stable activity levels, preventing activity from becoming excessively low or high (Turrigiano et al., 1998; Kim et al., 2012; Keck et al., 2013; Torrado Pacheco et al., 2021). This process is generally considered to be multiplicative and independent of presynaptic activity (Turrigiano, 2008). Following previous computational studies (Van Rossum et al., 2000), we modeled the change of connection strength from the excitatory population in subnetwork j to the excitatory population in subnetwork i via synaptic scaling as follows:

$$\tau_{ss}^{EE}\frac{dw_{E_iE_j}}{dt} = (1 - \frac{r_{E_i}}{\theta_{E_i}})w_{E_iE_j}. \tag{3}$$

Here, τ_{ss} represents the time constant of synaptic scaling for individual type of connections, θ_{E_i} denotes the target firing rate or the set point of the excitatory population in the subnetwork i.

To investigate how different inhibitory synaptic scaling, discovered experimentally to excitatory dendrites and somas (Prestigio et al., 2021), collectively affect associative learning, we implemented synaptic scaling from PV-to-E and SST-to-E synapses. In line with the observed decrease in somatic inhibition induced by the hyperactivity of postsynaptic excitatory neurons (Prestigio

et al., 2021), PV-to-E synapses are scaled by:

$$\tau_{ss}^{EP}\frac{dw_{E_iP_j}}{dt} = -(1 - \frac{r_{E_i}}{\theta_{E_i}})w_{E_iP_j}. \tag{4}$$

Similarly, consistent with the observed increase in dendritic inhibition resulting from the hyperactivity of postsynaptic excitatory neurons (Prestigio et al., 2021), SST-to-E synapses are scaled by:

$$\tau_{ss}^{ES}\frac{dw_{E_iS_j}}{dt} = (1 - \frac{r_{E_i}}{\theta_{E_i}})w_{E_iS_j}. \tag{5}$$

The set points of the two excitatory populations were allowed to change (Leman et al., 2025), to reflect distinct dynamics and activity levels that emerge due to direct stimulation of one population during the conditioning phase. In particular, the set points were jointly determined by the corresponding activity and the set point regulator β according to the following dynamics:

$$\tau_\theta\frac{d\theta_{E_i}}{dt} = (-\theta_{E_i} + r_{E_i}) + (-\theta_{E_i} + \beta_{E_i}) \tag{6}$$

where τ_θ is the time constant governing the plasticity of the set points. The set point regulator β can be considered a form of a global homeostatic mechanism that uniformly regulates the activity of the entire network and is dynamically updated according to:

$$\tau_\beta\frac{d\beta_{E_i}}{dt} = -\beta_{E_i} + r_{E_i} \tag{7}$$

$$\beta_{E_i} \rightarrow \beta_{E_i} - k \;\text{ at }\; t^{cond_onset} \tag{8}$$

where τ_β denotes the time constant governing the plasticity of the set point regulator and k is a free parameter that determines the magnitude of the abrupt decrease in the set point regulator β of excitatory populations in both subnetworks at the onset of conditioning t^{cond_onset}. Conditioning raises the excitatory activity r_E, thereby increasing both the set point θ and the set point regulator β. In contrast, this sudden reduction in β counteracts the increases induced by conditioning and functions as a homeostatic mechanism to globally regulate the overall activity level.

To evaluate memory specificity after conditioning, we presented a test stimulus (TS) to the network by providing excitatory inputs to the E and PV populations in subnetwork 2 at three distinct test time points (4h, 24h, and 48h), and measured the activity of the excitatory population in subnetwork 1, denoted by $r_{E_1}^{test}$ (Figure 3A). This activity is compared to a reference activity, $r_{E_1}^{ref}$, obtained by simulating the network under identical initial conditions but without applying conditioning (Figure S1A). We observed that $r_{E_1}^{test}$ exceeds $r_{E_1}^{ref}$ during TS presentation at 4h (Figure 3B), whereas, at 24h and 48h, $r_{E_1}^{test}$ is smaller than $r_{E_1}^{ref}$ (Figure 3B). These results suggest that, following conditioning, the memory initially generalizes to test stimuli but eventually becomes specific over time.

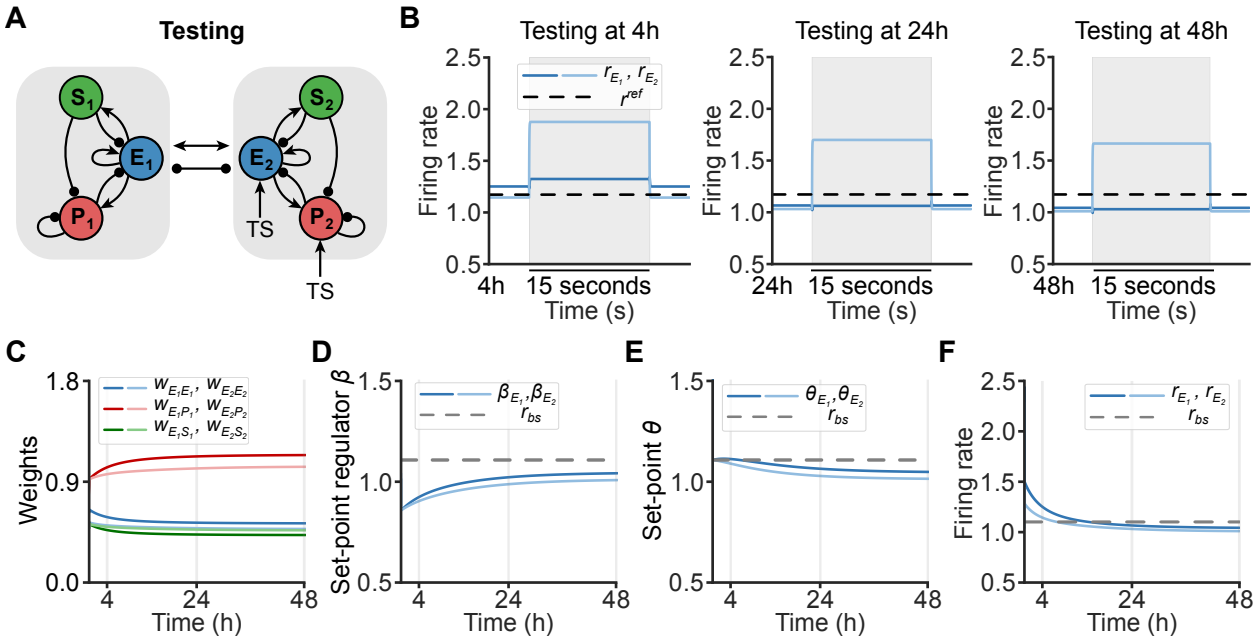

Fig. 3. Memory gradually transitions from generalization to specificity. **A.** Network schematic of the testing phase. After conditioning, E (E2) and PV (P2) populations of subnetwork 2 receive additional inputs that correspond to the test stimulus. The unconditioned stimulus is not presented during this phase. **B.** Responses of excitatory populations in subnetwork 1 and subnetwork 2 when presenting a test stimulus for 15s (gray) at 4h (left), 24h (middle) and 48h (right). The black horizontal lines indicate the reference activity (r^{ref}), measured by the excitatory population in subnetwork 1 in response to a test stimulus under identical initial weights conditions but without plasticity. Here, a moderate value of k = 0.25 is applied. **C.** Different connection strengths (E-to-E, PV-to-E and SST-to-E) after conditioning. **D.** Evolution of set points regulators of excitatory population in subnetwork 1 (β_{E_1}) and subnetwork 2 (β_{E_2}) after conditioning up to 48h. The gray horizontal dashed line represents the baseline activity level measured before conditioning. **E.** Same as D but for set point θ. **F.** Activity of excitatory population in subnetwork 1 (r_{E_1}) and subnetwork 2 (r_{E_2}) after conditioning.

Although synaptic scaling is active throughout the entire simulation, including the conditioning phase, its slow time constant renders changes in the E-to-E connection strength during conditioning negligible. Following conditioning, the weights evolve solely due to the different forms of synaptic scaling, undergoing significant changes over time: E-to-E weights decrease, PV-to-E weights increase, and SST-to-E weights decrease (Figure 3C). Following the abrupt decrease at the onset of conditioning, the set point regulators gradually rise over time, driven by increased excitatory activity in the early post-conditioning period (Figure 3D). Due to the set point regulators being lower than the baseline activity, the set points, decreased throughout the post-conditioning period, eventually stabilizing at a new steady state lower than the initial set points (Figure 3E). Consequently, following conditioning, the excitatory activity, along with PV and SST activity, gradually decreases over time and converges towards the new set points (Figure 3F). Our results indicate that synaptic scaling gradually reshapes network connectivity after conditioning, driving a progressive reduction in excitatory, PV, and SST activity (Figure S2) as the system stabilizes to a new equilibrium.

Characterization of the temporal evolution of memory representations

To characterize how memory representations dynamically evolve over time, we aimed to describe the network's response to the test stimulus following conditioning in the model. To that end, we introduced a procedure consisting of two phases for calculating the weights in the network during and post-conditioning, followed by the testing phase where we defined a "Generalization Index" to measure the degree of memory specificity or generalization (Figure 4A).

During conditioning (Phase 1), the excitatory firing rate can be well approximated by an exponential function (see Methods), capturing the simulated activity dynamics (Figure S3). We derived the evolution of E-to-E synaptic weights during this phase from solving the dynamics of the three-factor Hebbian learning rule, producing values that closely match those observed in simulations (Figure S3). Given that synaptic scaling operates on a substantially longer timescale than the duration of conditioning, its effects are negligible in this phase. After conditioning (Phase 2), excitatory firing rates can also be well described by an exponential function allowing us to compute the set point regulator β. Subsequently, we derived the set point θ from the obtained β. This allows us to accurately determine the synaptic weight evolution during post-conditioning (Figure S4).

To quantify the degree of memory specificity or generalization, we defined a new measure, called the Generalization Index (GI):

$$GI = \frac{r_{E_1}^{test} - r_{E_1}^{ref}}{r_{E_1}^{ref}}. \tag{9}$$

The GI quantifies the relative change in the excitatory population activity of subnetwork 1 between the test and the previously defined reference conditions. A positive GI (e.g. $r_{E_1}^{test} > r_{E_1}^{ref}$) suggests that a memory has been generalized, a negative GI (e.g. $r_{E_1}^{test} \leq r_{E_1}^{ref}$) indicates that a memory is specific. The magnitude of GI reflects the strength of memory specificity or generalization. Applying the above procedure, we found that the GI gradually transitions from positive to negative (Figure 4B). This transition suggests that, following conditioning, the memory initially generalizes to test stimuli but gradually becomes specific over time. Taken together, our procedure provides a quantitative characterization of the temporal dynamics underlying evolving memory representations, revealing how these representations are progressively reshaped over time.

The global homeostatic mechanism adjusts the set points and the strength of memory specificity

The global homeostatic mechanism in our model (Eq. 8), governed by the parameter k, influences the set point regulators and thus the new set points. For a small k, in the absence of the global homeostatic mechanism, the set points θ slightly increase throughout the post-conditioning period and stabilize at a new steady state moderately above the baseline activity (i.e., the initial set points) (Figure 5A). Following conditioning, excitatory activity decreases and approaches the new

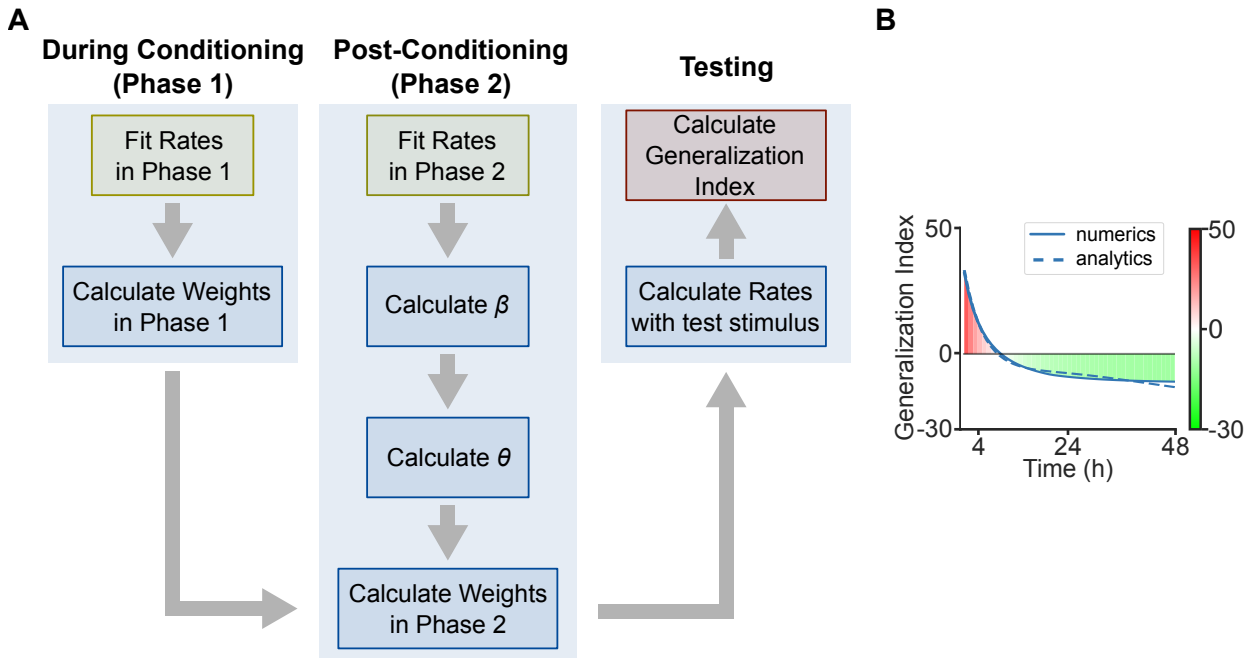

Fig. 4. A procedure to assess the evolution of memory representations. **A.** Workflow chart for calculating the Generalization Index (GI) (Eq. 9), see main text. **B.** Evolution of the GI after conditioning. Numerical results (solid line) represent GI measurements taken hourly post-conditioning, while analytical results are derived from continuous GI calculations using the procedure described in Figure 4A. The GI shifts from positive to negative, indicating the transition from memory generalization to memory specificity.

set points (Figure 5B). The GI remains positive throughout the post-conditioning period consistent with memory generalization (Figure 5C). In contrast, increasing the influence of the global homeostatic mechanism (large k), suppresses the set points θ and excitatory activity (Figure 5D, E). In this case, the GI shifts from positive to negative throughout the post-conditioning period (Figure 5F) and reaches a lower value compared to the immediate k condition (Figure 3C), indicating enhanced memory specificity. Taken together, these results suggest that the global homeostatic mechanism significantly influences the set points and regulates the degree of memory specificity.

Top-down inputs regulate memory specificity

In addition to bottom-up inputs driven by sensory stimuli, primary sensory cortical areas also receive abundant top-down inputs from higher-order regions which influence neuronal processing in local recurrent circuits (Johnson and Burkhalter, 1997; Garrett et al., 2014). To investigate how top-down inputs influence associative learning, we applied an additional input to SST populations (a common target of top-down inputs) during conditioning (Batista-Brito et al., 2018; Shen et al., 2022). When introducing an inhibitory top-down input to both SST populations (Figure 6A), we found that inhibition of SST interneurons disinhibits excitatory neurons, leading to a drastic increase in the firing rates of both subnetworks during conditioning (Figure 6B). However, after

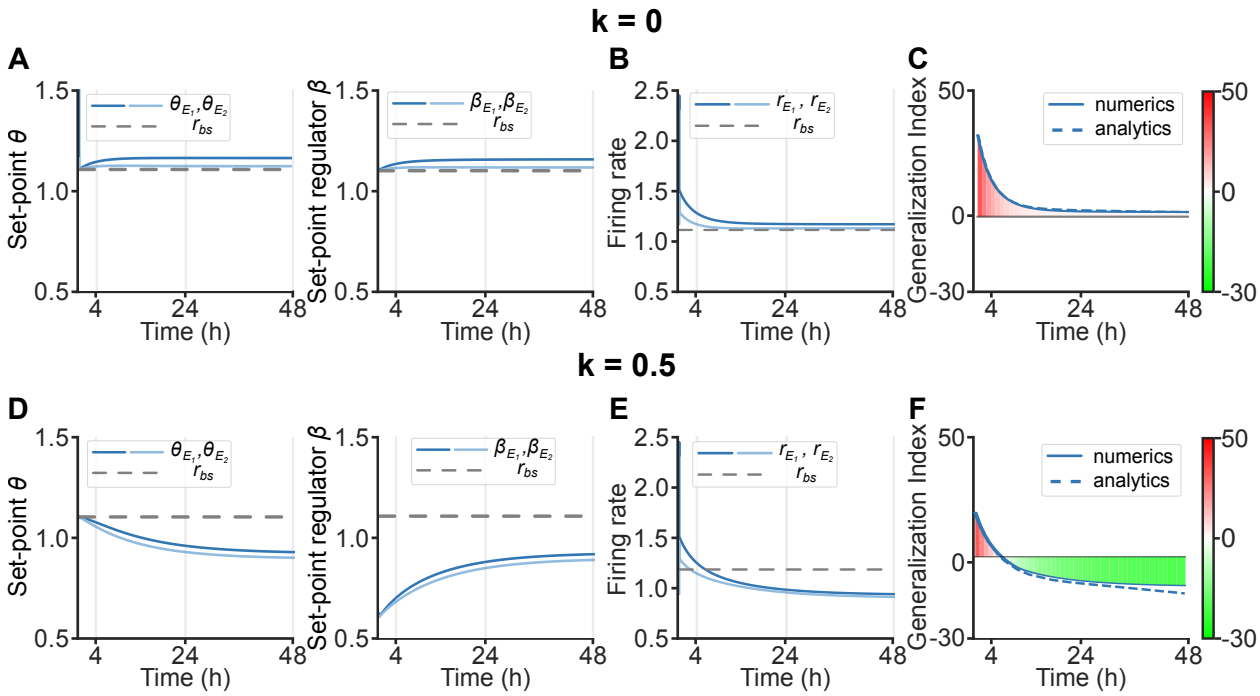

Fig. 5. The global homeostatic mechanism regulates the emergence of memory specificity. **A.** Evolution of set point θ (left) and set point regulator β (right) in subnetwork 1 and subnetwork 2 after conditioning up to 48h for $k = 0$, corresponding to the absence of the global homeostatic mechanism. The gray horizontal dashed line represents the baseline activity level measured before conditioning. **B.** Evolution of excitatory firing rate in subnetwork 1 and subnetwork 2 after conditioning up to 48h for $k = 0$. **C.** Evolution of the GI after conditioning. For $k = 0$, GI remains positive after conditioning. A positive GI indicates memory generalization, whereas a negative GI represents memory specificity. **D - F.** Same as (A - C) but for $k = 0.5$. The GI transitions from positive to negative after 4h after conditioning.

conditioning, excitatory activity rapidly declines (Figure 6C). Furthermore, due to the large initial change in firing rates, the synaptic weights undergo substantial modification (Figure 6D). The GI transitions from positive to negative earlier in the post-conditioning period than in the absence of inhibitory top-down input, indicating a faster emergence of memory specificity (Figure 6E). In contrast, when applying excitatory top-down input to both SST populations in the same amount during conditioning (Figure 6F), the excitatory firing rate of subnetwork 1 slightly increases, while subnetwork 2 decreases (Figure 6G). After conditioning, both subnetworks' excitatory activity gradually decline below baseline levels (Figure 6H). Although the firing rates at 48h show negligible differences compared to the case with inhibitory top-down inputs, the changes in synaptic weights are significantly smaller (Figure 6I), leading to a marked difference in the GI (Figure 6J).

These findings suggest that top-down inhibition of SST would transiently enhance excitatory activity but paradoxically accelerate memory specificity, whereas top-down excitation of SST confines the degree of increase in excitatory activity and delays the refinement of memory representations.

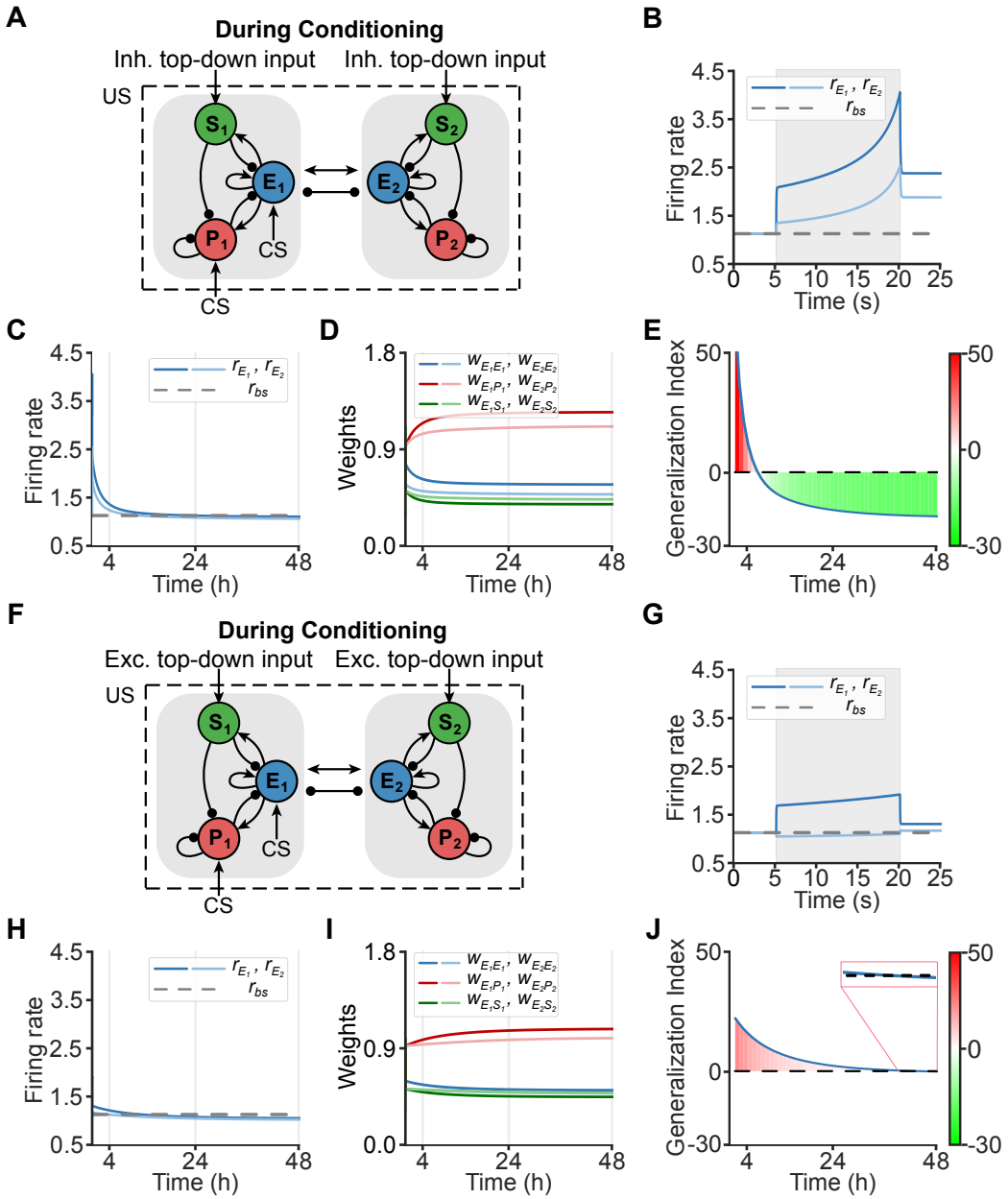

Fig. 6. Top-down inputs influence memory specificity. **A.** Network schematic of the conditioning phase in the presence of inhibitory top-down inputs. During conditioning, the E and PV populations of subnetwork 1 receive additional inputs that correspond to the conditioned stimulus, while the SST population of both subnetworks 1 and 2 receives additional inhibitory top-down inputs. **B.** Activity of excitatory population in subnetwork 1 (r_{E_1}) and subnetwork 2 (r_{E_2}) during conditioning in the presence of inhibitory top-down inputs. Conditioning is marked by the gray interval from 5 to 20s. The gray horizontal dashed line represents the baseline activity level measured before conditioning. **C.** Activity of excitatory population in subnetwork 1 (r_{E_1}) and subnetwork 2 (r_{E_2}) after conditioning in the presence of inhibitory top-down inputs. **D.** Different connection strengths (E-to-E, PV-to-E and SST-to-E) after conditioning. **E.** Evolution of the Generalization Index (GI) after conditioning in the presence of inhibitory top-down inputs. A positive GI indicates memory generalization, whereas a negative GI represents memory specificity. **F - J.** Same as (A - E) but in the presence of excitatory top-down inputs.

Synaptic scaling is essential for associative learning, with distinct contributions from specific cell types

Next, we investigated the role of synaptic scaling in associative learning by blocking all synaptic scaling mechanisms. In the absence of synaptic scaling mechanisms, both firing rates and weights stay unchanged during the post-conditioning period (Figure 7A), $r_{E_1}^{test}$ exceeded r^{ref} (Figure 7B), and the GI remains positive (Figure 7C), suggesting memory generalization. Together, these results indicate that synaptic scaling is crucial for achieving memory specificity.

But to which extent do the different types of synaptic scaling affect associative learning? When selectively blocking E-to-E scaling (Figure 7D), we found that excitatory firing rates of both subnetworks gradually converge to levels close to their baseline (Figure 7E). The GI shifts from positive to negative, albeit at a later time point compared to when all scaling mechanisms are present. This indicates that the memory eventually becomes specific (Figure 7F), as shown experimentally (Wu et al., 2021). Blocking PV-to-E scaling (Figure 7G) elevates excitatory firing rates of both subnetworks constantly beyond their baseline levels (Figure 7H), resulting in a positive GI and hence generalized memories (Figure 7I). In contrast, when SST-to-E scaling is blocked (Figure 7J), the excitatory firing rates of both subnetworks promptly decrease to levels below their baseline (Figure 7K). This results in the GI transitioning from positive to negative earlier in the post-conditioning period than with intact scaling mechanisms (Figure 7L), indicating a more rapid emergence of memory specificity.

To assess the robustness of the observed results to parameter selection, we conducted numerous simulations using different initial weight conditions (see Methods). We found that, in 93% of initial weight conditions, PV-to-E scaling is essential to achieve memory specificity, whereas blocking SST-to-E scaling always accelerates the transition to memory specificity (see Methods, Figure S5, S6). Together, these results suggest that E-to-E and PV-to-E scaling operate synergistically, while SST-to-E scaling acts antagonistically, to collectively regulate the timing of memory specificity.

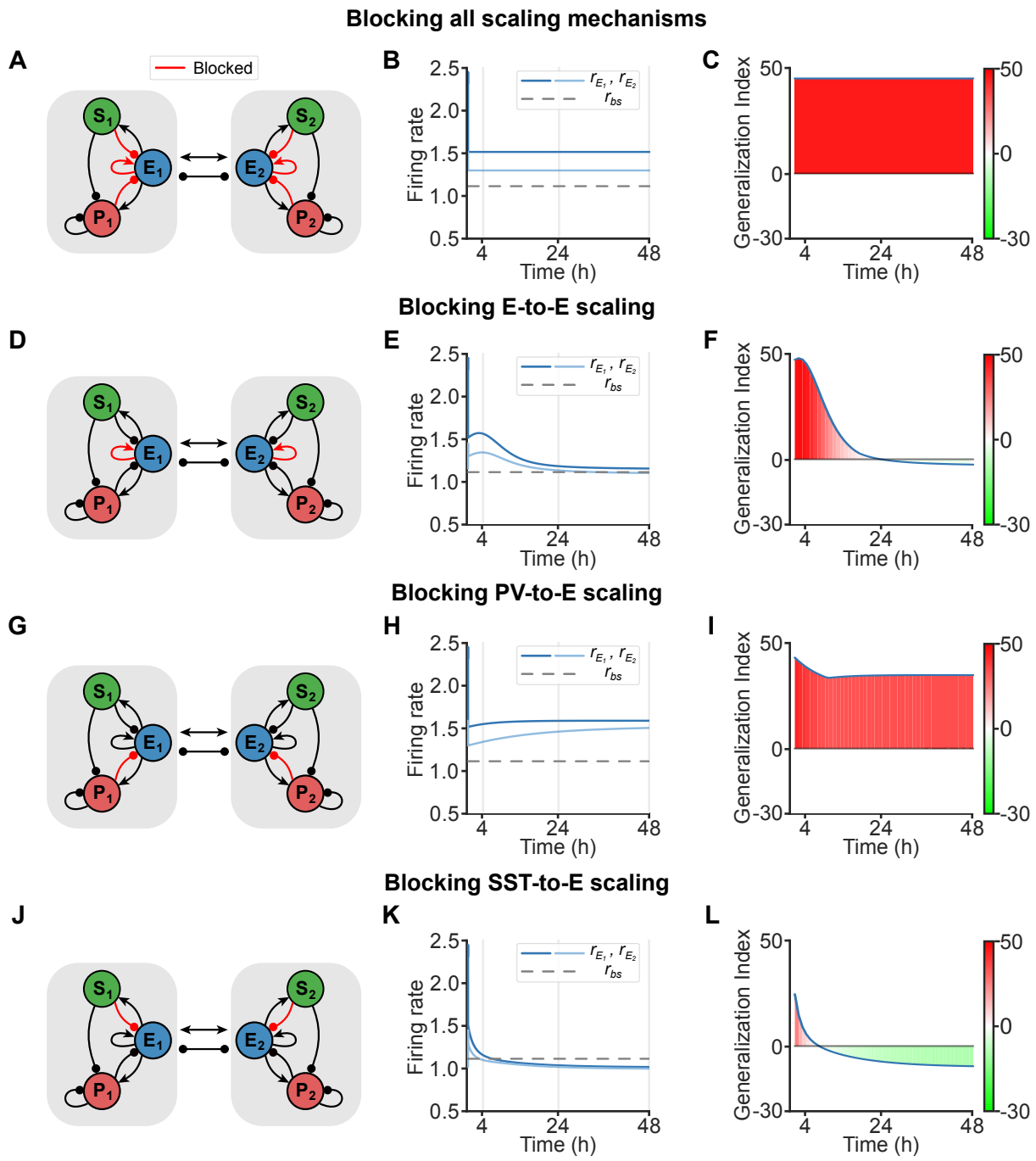

Fig. 7. Cell-type-specific synaptic scaling contributions to memory refinement in associative learning. **A.** Network schematic when blocking all scaling mechanisms (red connections). Cross-connections are also blocked accordingly. **B.** Activity of excitatory population in subnetwork 1 (r_{E_1}) and subnetwork 2 (r_{E_2}) after conditioning when blocking all scaling mechanisms. The gray horizontal dashed line represents the baseline activity level measured before conditioning. **C.** Evolution of the Generalization Index (GI) after conditioning when blocking all scaling mechanisms. A positive GI indicates memory generalization, whereas a negative GI represents memory specificity. **D - F.** Same as (A - C) but for blocking E-to-E scaling. **G - I.** Same as (A - C) but for blocking PV-to-E scaling. **J - L.** Same as (A - C) but for blocking SST-to-E scaling.

Discussion

Here, we investigated how different plasticity mechanisms shape associative learning in recurrent circuits comprising multiple interneuron types. Using analytical and computational approaches, we demonstrated that brief conditioning induces memory generalization through Hebbian plasticity. Following conditioning, different forms of synaptic scaling progressively establish memory specificity over time. Specifically, E-to-E and PV-to-E scaling function synergistically, but counteract SST-to-E scaling, to collectively govern the timing of memory refinement. Our findings reveal the cell-type-specific contributions of synaptic scaling and propose a role for top-down modulation in regulating associative learning.

Our study revealed several key insights into the mechanisms and consequences of associative learning. We demonstrated that different forms of synaptic scaling – a relatively slow process – are crucial for establishing memory specificity. This finding aligns with experimental observations showing that memory specificity emerges only several hours after conditioning (Wu et al., 2021). In the context of conditioned taste aversion, over time, the gradual fading of memory generalization may reduce food avoidance along with increasing hunger. Faster mechanisms, such as Hebbian inhibitory plasticity, could accelerate the elimination of food avoidance, but at the cost of a higher risk of encountering aversive food. In contrast, slower mechanisms, like synaptic scaling, may be more beneficial for animals to minimize risk while avoiding starvation. In addition to the gustatory cortex, similar associative learning paradigms have been applied in other sensory cortical regions. For instance, by pairing specific sounds with a foot shock in the auditory cortex (Letzkus et al., 2011), by associating specific visual stimuli with rewards in the visual cortex (Pakan et al., 2018), by pairing specific odors with rewards in the olfactory cortex (Ottenheimer et al., 2023). Given the ubiquity of the cortical circuit motifs we modeled (Tremblay et al., 2016), our findings have the potential to provide broad insights into associative learning across sensory cortical regions.

Going beyond capturing existing experimental data, our model proposes a critical role of top-down influences in associative learning. Our findings demonstrate that a global, unspecific top-down signal, mediated by the unconditioned stimulus (e.g., punishment or reward), acts as a gate for the Hebbian learning process. In addition to these global signals, more specific top-down inputs, such as those related to attention and that target particular cell types (Park et al., 2025), can profoundly influence neural activity. These specific inputs can flexibly shift different subnetworks either into a long-term potentiation (LTP)-dominated or long-term depression (LTD)-dominated regime, thereby shaping associative learning and the timing of memory specificity emergence. Notably, while both types of inputs affect learning, global, unspecific top-down signals exert minimal influence on activity levels, whereas more specific, cell-type-targeted inputs affect learning by strongly modulating activity. Our model thus highlights that distinct sources of top-down inputs can act in parallel, each contributing to learning in mechanistically different ways.

Furthermore, our computational model allowed us to test and identify the cell-type-specific contributions to associative learning. Specifically, E-to-E scaling and PV-to-E scaling operate synergistically while opposing SST-to-E scaling. Our findings indicate that disabling excitatory synaptic scaling while preserving all forms of inhibitory synaptic scaling achieves memory specificity within 48 hours after conditioning. In contrast, disabling PV-to-E synaptic scaling while maintaining other scaling mechanisms prevents the establishment of memory specificity within the same period. These results highlight a powerful role of different forms of inhibitory synaptic scaling in facilitating precise associative learning.

To preserve computational and analytical tractability, we made several simplifications. Biological neurons possess complex morphologies and exhibit a non-uniform distribution of ion channels and synaptic inputs across their dendritic trees (Jiang et al., 2015; Peng et al., 2021). Parvalbumin (PV) and somatostatin (SST) interneurons exhibit distinct targeting patterns on pyramidal neurons, influencing their computational properties. PV interneurons primarily innervate the perisomatic region, including the soma and proximal dendrites, while, in contrast, SST interneurons target distal dendrites (Dorsett et al., 2021; Schneider-Mizell et al., 2025). These dendritic nonlinearities and localized synaptic interactions play a crucial role in integrating synaptic inputs and shaping neuronal output (Poirazi et al., 2003; London and Häusser, 2005; Larkum et al., 2009), influencing network dynamics and learning processes. Nevertheless, by using point neuron models combined with the incorporation of known connectivity properties, our work provides valuable insight into how different synaptic scaling mechanisms influence associative learning.

In our work, we primarily investigated various synaptic scaling mechanisms during associative learning while excluding long-term inhibitory Hebbian plasticity. Although inhibitory synapses are known to undergo modifications driven by Hebbian plasticity (Froemke et al., 2007; D'amour and Froemke, 2015; Hennequin et al., 2017; Lagzi et al., 2021; Schulz et al., 2021; Wu et al., 2022; Miehl and Gjorgjieva, 2022; Festa et al., 2024), experiments suggest that the timescale of long-term inhibitory plasticity might be too rapid to explain the prolonged duration of memory generalization and the gradual emergence of memory specificity at 48 hours after blocking E-to-E scaling (Wu et al., 2021). Therefore, we postulate that this delayed emergence of memory specificity is likely driven by slower processes, such as inhibitory synaptic scaling, as proposed in our study.

In addition, beyond the three cell types (E, PV, and SST) included in our model, several other inhibitory interneuron subtypes have been identified (Wilmes and Clopath, 2019; Hertäg and Sprekeler, 2020; Pardi et al., 2020; Canto-Bustos et al., 2022; Veit et al., 2023; Palmigiano et al., 2023; Hartung et al., 2024; Naumann et al., 2025). Among these, vasoactive intestinal peptide (VIP)-expressing interneurons are a prominent class often incorporated into canonical microcircuit motifs (Pfeffer et al., 2013; Waitzmann et al., 2024). VIP interneurons primarily inhibit SST cells and are known to receive top-down inputs, which can significantly impact recurrent network dynamics (Fu et al., 2014; Zhang et al., 2014; Dipoppa et al., 2018; Garrett et al., 2020; Bastos

et al., 2023; Furutachi et al., 2024). Although VIP interneurons were not explicitly modeled in our study, by providing dedicated inputs to SST to emulate top-down modulation, our results suggest a pivotal role of top-down modulation in shaping associative learning.

Together, our work offers new insights into how distinct plasticity mechanisms interact to shape associative learning, highlights the significant impact of top-down influences and synaptic scaling, and reveals the cell-type-specific contributions to the establishment of precise memory representations.

Methods

Rate-based population model

To investigate the role of cell-type-specific synaptic scaling in associative learning, we constructed a rate-based population model comprising two subnetworks. Each subnetwork includes one excitatory, one PV, and one SST population. Different subnetworks are tuned to different stimuli corresponding to different tastants in the conditioned taste aversion experiments. The dynamics of the network can be described as follows (Richter and Gjorgjieva, 2022):

$$\tau \frac{d\boldsymbol{r}}{dt} = -\boldsymbol{r} + [\boldsymbol{W}\boldsymbol{r} + \boldsymbol{g} - \rho]_+ , \tag{10}$$

where τ is a diagonal matrix containing the time constants of firing rate dynamics for different populations, $\boldsymbol{r}$ is a vector containing the firing rates of different populations, $\boldsymbol{g}$ is a vector containing the inputs to different populations, and ρ is a vector containing the rheobases of different populations, $[]_+$ is a rectified function.

$$\boldsymbol{r} = \begin{bmatrix} r_{E_1} \\ r_{P_1} \\ r_{S_1} \\ r_{E_2} \\ r_{P_2} \\ r_{S_2} \end{bmatrix}, \boldsymbol{\tau} = \begin{bmatrix} \tau_{E_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \tau_{P_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & \tau_{S_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \tau_{E_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \tau_{P_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \tau_{S_2} \end{bmatrix}, \boldsymbol{g} = \begin{bmatrix} g_{E_1} \\ g_{P_1} \\ g_{S_1} \\ g_{E_2} \\ g_{P_2} \\ g_{S_2} \end{bmatrix}, \quad \boldsymbol{\rho} = \begin{bmatrix} \rho_{E_1} \\ \rho_{P_1} \\ \rho_{S_1} \\ \rho_{E_2} \\ \rho_{P_2} \\ \rho_{S_2} \end{bmatrix}.$$

$\boldsymbol{W}$ is the connectivity matrix defined as follows:

$$\boldsymbol{W} = \begin{bmatrix} w_{E_1E_1} & -w_{E_1P_1} & -w_{E_1S_1} & w_{E_1E_2} & -w_{E_1P_2} & -w_{E_1S_2} \\ w_{P_1E_1} & -w_{P_1P_1} & -w_{P_1S_1} & w_{P_1E_2} & -w_{P_1P_2} & -w_{P_1S_2} \\ w_{S_1E_1} & 0 & 0 & w_{S_1E_2} & 0 & 0 \\ w_{E_2E_1} & -w_{E_2P_1} & -w_{E_2S_1} & w_{E_2E_2} & -w_{E_2P_2} & -w_{E_2S_2} \\ w_{P_2E_1} & -w_{P_2P_1} & -w_{P_2S_1} & w_{P_2E_2} & -w_{P_2P_2} & -w_{P_2S_2} \\ w_{S_2E_1} & 0 & 0 & w_{S_2E_2} & 0 & 0 \end{bmatrix}.$$

To model the conditioned taste aversion experimental paradigm, specifically, to simulate the conditioned stimulus, additional inputs are provided to the excitatory (E_1) and PV (P_1) populations in the subnetwork 1 via increasing g_{E_1} and g_{P_1} by $g_{\Delta E}$ and $g_{\Delta P}$, respectively. Similarly, to simulate the test stimulus, additional inputs are applied to the excitatory (E_2) and PV (P_2) populations in the subnetwork 2 via increasing g_{E_2} and g_{P_2} by $g_{\Delta E}$ and $g_{\Delta P}$, respectively. Parameter values for two subnetworks are the same unless mentioned otherwise.

Three-factor Hebbian plasticity

Motivated by experimental studies showing that reward or punishment plays a decisive role in learning (Pawlak, 2010; Yagishita et al., 2014; He et al., 2015; Gerstner et al., 2018), we modeled Hebbian plasticity using a three-factor learning rule as follows:

$$\tau_{hebb}\frac{dw_{E_iE_j}}{dt} = \eta r_{E_j}(r_{E_i} - r_{E_i}^{bs}),$$

where τ_{hebb} is the time constant of Hebbian plasticity, $r_{E_i}^{bs}$ represents the baseline activity of the excitatory population in subnetwork i before conditioning, and $i, j \in \{1, 2\}$, representing the indices of subnetworks.

The third factor η is determined by the presence of the unconditioned aversive stimulus. More specifically,

$$\eta = \begin{cases} 1, & \text{in the presence of unconditioned stimulus,} \\ 0, & \text{otherwise.} \end{cases}$$

Thus, the third factor serves as a gate for Hebbian plasticity, enabling it during the conditioning phase while disabling it elsewhere.

Synaptic scaling

The dynamics of the connection strength governed by synaptic scaling from the excitatory population in subnetwork j to the excitatory population in subnetwork i is given by (Van Rossum et al., 2000):

$$\tau_{ss}^{EE}\frac{dw_{E_iE_j}}{dt} = (1 - \frac{r_{E_i}}{\theta_{E_i}})w_{E_iE_j}.$$

Similarly, for PV-to-E synaptic scaling, we have:

$$\tau_{ss}^{EP}\frac{dw_{E_iP_j}}{dt} = -(1 - \frac{r_{E_i}}{\theta_{E_i}})w_{E_iP_j}.$$

And for SST-to-E synaptic scaling, we have:

$$\tau_{ss}^{ES}\frac{dw_{E_iS_j}}{dt} = (1 - \frac{r_{E_i}}{\theta_{E_i}})w_{E_iS_j}.$$

Here, τ_{ss} represents the time constant of synaptic scaling for individual type of connections, θ_{E_i} denotes the target firing rate or the set point of the excitatory population in the subnetwork i.

Plasticity of set points

Set points of excitatory populations are subject to plastic changes, governed by the following dynamics:

$$\tau_\theta \frac{d\theta_{E_i}}{dt} = (-\theta_{E_i} + r_{E_i}) + (-\theta_{E_i} + \beta_{E_i}),$$

where τ_θ is the time constant governing the plasticity of the set points. The set point θ_{E_i} evolves based on the current activity r_{E_i} and the set point regulator β_{E_i}. The set point regulator β is dynamically updated according to:

$$\tau_\beta \frac{d\beta_{E_i}}{dt} = -\beta_{E_i} + r_{E_i},$$

$$\beta_{E_i} \rightarrow \beta_{E_i} - k \text{ at } t^{cond_onset},$$

where τ_β denotes the time constant governing the plasticity of the set point regulator and k is a free parameter that determines the magnitude of the abrupt decrease in the set point regulator β of excitatory populations in both subnetworks at the onset of conditioning. Conditioning raises the activity r_E, thereby increasing both the set point θ and the set point regulator β, in contrast, this sudden reduction in β counteracts the increases induced by conditioning and functions as a homeostatic mechanism to globally regulate the overall activity level (Kaleb et al., 2021).

Analytical procedure

To thoroughly characterize the temporal evolution of memory specificity and generalization – specifically, how the network responds to the test stimulus following conditioning – we introduced a procedure to determine how set point regulators, set points, weights, and rates during conditioning and after conditioning evolve dynamically. In this procedure (Figure 4A), we defined two phases, and assumed that the firing rates of excitatory populations during conditioning (Phase 1) and after conditioning (Phase 2) in the absence of the test stimulus, which are experimentally measurable, are known. First, we formulate the time-variant firing rate of the excitatory population in the subnetwork i during conditioning (Phase 1) as an exponential function as follows:

$$\hat{r}_{E_i}^{(1)} = a_i^{(1)} e^{(-b_i^{(1)} t)} + c_i^{(1)}, \tag{11}$$

where a_i^1, b_i^1, and c_i^1 are coefficients obtained by fitting the parameterized functions to the excitatory firing rates of the subnetwork i during conditioning in the simulation. Superscripts '(1)' indicate Phase 1 corresponding to the conditioning period.

By solving the three-factor Hebbian learning equation (Eq. 1), we can obtain $\hat{w}^{(1)}_{E_iE_j}$ during conditioning as:

$$\hat{w}^{(1)}_{E_iE_j} = f^{(1)}_{w_{E_iE_j}}(t) + C^{(1)}, \tag{12}$$

where

$$f^{(1)}_{w_{E_iE_j}}(t) = \frac{\eta}{\tau_{Hebb}}\left[\frac{a^{(1)}_i a^{(1)}_j}{b^{(1)}_i + b^{(1)}_j} e^{(b^{(1)}_i + b^{(1)}_j)t} + \frac{a^{(1)}_j c^{(1)}_i - a^{(1)}_j r_{bs}}{b^{(1)}_j} e^{b^{(1)}_j t} + \frac{a^{(1)}_i c^{(1)}_j}{b^{(1)}_i} e^{b^{(1)}_i t} + (c^{(1)}_i c^{(1)}_j - c^{(1)}_j r_{bs})t\right], \tag{13}$$

$$C^{(1)} = w_{E_iE_j}(0) - \frac{\eta}{\tau_{Hebb}}\left[\frac{a^{(1)}_i a^{(1)}_j}{b^{(1)}_i + b^{(1)}_j} + \frac{a^{(1)}_j c^{(1)}_i - a^{(1)}_j r_{bs}}{b^{(1)}_j} + \frac{a^{(1)}_i c^{(1)}_j}{b^{(1)}_i}\right], \tag{14}$$

where $w_{E_iE_j}(0)$ is the initial value before conditioning. Note that here we disregarded the effects of synaptic scaling on the weights during conditioning due to its relatively slow timescale compared to the duration of the conditioning.

We then approximate the excitatory firing rates after conditioning (Phase 2) in a similar way:

$$\hat{r}^{(2)}_{E_i} = a^{(2)}_i e^{(-b^{(2)}_i t)} + c^{(2)}_i. \tag{15}$$

Given k, which represents the magnitude of the abrupt decrease in the set point regulator β, the set point regulator $\hat{\beta}$ after conditioning can be determined analytically as follows:

$$\hat{\beta}^{(2)}_{E_i} = (r_{bs} - k - \frac{a^{(2)}_i}{1 - b^{(2)}_i \tau_\beta} - c^{(2)}_i)e^{-\frac{t}{\tau_\beta}} + \frac{a^{(2)}_i}{1 - b^{(2)}_i} e^{-b^{(2)}_i t} + c^{(2)}_i. \tag{16}$$

Using the $\hat{\beta}$ from the above equation, the set points of excitatory populations $\hat{\theta}$ after conditioning can be calculated as follows:

$$\hat{\theta}^{(2)}_{E_i} = L^{(2)}_{E_i} e^{-\frac{2}{\tau_\beta}t} + M^{(2)}_{E_i} e^{-b^{(2)}_i t} + N^{(2)}_{E_i} e^{-\frac{t}{\tau_\beta}} + c^{(2)}_i, \tag{17}$$

with

$$L^{(2)}_{E_i} = r_{bs} - \frac{a^{(2)}_i + \frac{a^{(2)}_i}{1-b^{(2)}_i \tau_\beta}}{2 - \tau_\theta b^{(2)}_i} - \frac{r_{bs} - k - \frac{a^{(2)}_i}{1-b^{(2)}_i \tau_\beta} - c^{(2)}_i}{2 - \frac{\tau_\theta}{\tau_\beta}} - c^{(2)}_i, \tag{18}$$

$$M^{(2)}_{E_i} = \frac{a^{(2)}_i + \frac{a^{(2)}_i}{1-b^{(2)}_i \tau_\beta}}{2 - \tau_\theta b^{(2)}_i}, \tag{19}$$

$$N^{(2)}_{E_i} = \frac{r_{bs} - k - \frac{a^{(2)}_i}{1-b^{(2)}_i \tau_\beta} - c^{(2)}_i}{2 - \frac{\tau_\theta}{\tau_\beta}}. \tag{20}$$

Using the above obtained $\hat{\theta}$, the weights after conditioning can be computed as follows:

$$\hat{w}^{(2)}_{E_iE_j} = w^{init}_{E_iE_j} e^{\int_0^t f_i(x)dx}, \tag{21}$$

$$\hat{w}^{(2)}_{E_iP_j} = -w^{init}_{E_iP_j} e^{\int_0^t f_i(x)dx}, \tag{22}$$

$$\hat{w}^{(2)}_{E_iS_j} = w^{init}_{E_iS_j} e^{\int_0^t f_i(x)dx}, \tag{23}$$

with

$$f_i(x) = \frac{1}{\tau_{ss}}(1 - \frac{a^{(2)}_i e^{-b^{(2)}_i x} + c^{(2)}_i}{L^{(2)}_{E_i} e^{-\frac{2}{\tau_\beta}t} + M^{(2)}_{E_i} e^{-b^{(2)}_i t} + N^{(2)}_{E_i} e^{-\frac{t}{\tau_\beta}} + c^{(2)}_i}), \tag{24}$$

where $w^{init}_{E_iE_j}$, $w^{init}_{E_iP_j}$, and $w^{init}_{E_iS_j}$ denote the weights right after conditioning from the excitatory, PV and SST population in the subnetwork j to the excitatory population in the subnetwork i. We obtain $w^{init}_{E_iE_j}$ from Eq. 12. In contrast, $w^{init}_{E_iP_j}$ and $w^{init}_{E_iS_j}$ are identical to their initial values $w_{E_iP_j}(0)$ and $w_{E_iP_j}(0)$, respectively, as they remain unchanged during conditioning.

From the above-obtained weights, assuming all populations exhibit non-zero firing rates, the steady-state firing rate for any given input at any given time after conditioning can be calculated as follows:

$$\boldsymbol{r} = (\boldsymbol{I} - \boldsymbol{W})^{-1}(\boldsymbol{g} - \boldsymbol{\rho}). \tag{25}$$

By substituting the inputs from the test stimulus condition into the above equation, we can determine the responses to the test stimulus over time, providing a comprehensive description of the evolution of memory specificity and generalization.

Sensitivity analysis

To assess the robustness of our results to parameter selection, we conducted grid simulations across the model's parameter space by changing each parameter from 0.01 to 1.01 with increments of 0.1. We further constrained the parameters for the cross-connections to be smaller than those for within connections, resulting in a total of 72728 simulations. To ensure computational feasibility, we performed these simulations and evaluated memory specificity in models with varying initial conditions of either plastic or static weights. We only focused on models that meet two experimental conditions: (1) a transition from memory generalization to memory specificity occurs between 4h and 24h, and (2) when E-to-E scaling is blocked, this transition occurs between 24h and 48h. Models that meet these criteria are labeled as respecting 'experimentally-matched conditions' (in total 828 out of 58087 in Figure S5, 258 out of 14641 in Figure S6). Using these selected models, we evaluated whether the role of cell-specific synaptic scaling mechanisms aligns with the instance presented in the Results Section (Figure 7) by blocking either PV-to-E or SST-to-E scaling. If blocking PV-to-E scaling diminishes memory specificity and blocking SST-to-E

scaling accelerates the transition from memory generalization to memory specificity in these selected models, we labeled them as 'aligned conditions' (Figure S5, S6). We found that 96% of the selected models (1041 out of 1086) exhibit the same cell-type-specific contributions as in the presented example (Figure 7), suggesting that our results are robust to parameter selection.

Numerical simulation

Simulations were performed in Python with Numba (Lam et al., 2015). Differential equations were implemented by Euler method with a timestep of 0.1 milliseconds. All simulation parameters are listed in Table S1.

Data Availability

The simulation code is publicly available at https://github.com/comp-neural-circuits/cell-type-specific-synaptic-scaling.

Contributions

Y.K.W. and J.G. designed research; F.V. and A.K. performed the numerical simulations; F.V. and Y.K.W. performed the analytic calculations; F.V., A.K., Y.K.W., and J.G. wrote the paper.

Acknowledgments

We thank Gina Turrigiano and Chi-Hong Wu as well as members of the Computation in Neural Circuits Group for helpful discussions. This work was funded by the European Research Council (Grant Agreement No. 804824 to J.G.) and by the DFG in the Collaborative Research Centre 1080 (to J.G.). A.K. was also supported by TUM and the Elite Network of Bavaria. Y.K.W. was also supported by the Add-on Fellowship of the Joachim Herz Foundation.

References

Bastos G, Holmes JT, Ross JM, Rader AM, Gallimore CG, Wargo JA, et al. Top-down input modulates visual context processing through an interneuron-specific circuit. Cell Reports 2023 Sep;42(9):113133. doi: https://doi.org/10.1016/j.celrep.2023.113133.

Batista-Brito R, Zagha E, Ratliff JM, Vinck M. Modulation of cortical circuits by top-down processing and arousal state in health and disease. Current Opinion in Neurobiology 2018 Oct;52:172–181. doi: https://doi.org/10.1016/j.conb.2018.06.008.

Canto-Bustos M, Friason FK, Bassi C, Oswald AMM. Disinhibitory Circuitry Gates Associative Synaptic Plasticity in Olfactory Cortex. Journal of Neuroscience 2022;42(14):2942–2950. doi: https://doi.org/10.1523/JNEUROSCI.1369-21.2021.

Dipoppa M, Ranson A, Krumin M, Pachitariu M, Carandini M, Harris KD. Vision and Locomotion Shape the Interactions between Neuron Types in Mouse Visual Cortex. Neuron 2018 May;98(3):602–615.e8. doi: https://doi.org/10.1016/j.neuron.2018.03.037.

Dorsett C, Philpot BD, Smith SL, Smith IT. The Impact of SST and PV Interneurons on Nonlinear Synaptic Integration in the Neocortex. eneuro 2021 Sep;8(5):ENEURO.0235–21.2021. doi: https://doi.org/10.1523/ENEURO.0235-21.2021.

D'amour J, Froemke R. Inhibitory and Excitatory Spike-Timing-Dependent Plasticity in the Auditory Cortex. Neuron 2015 Apr;86(2):514–528. doi: https://doi.org/10.1016/j.neuron.2015.03.014.

Festa D, Cusseddu C, Gjorgjieva J. Structured stabilization in recurrent neural circuits through inhibitory synaptic plasticity. bioRxiv 2024; doi: https://doi.org/10.1101/2024.10.12.618014.

Froemke RC, Merzenich MM, Schreiner CE. A synaptic memory trace for cortical receptive field plasticity. Nature 2007 Nov;450(7168):425–429. doi: https://doi.org/10.1038/nature06289.

Fu Y, Tucciarone J, Espinosa J Sheng N, Darcy D, Nicoll R, et al. A Cortical Circuit for Gain Control by Behavioral State. Cell 2014 Mar;156(6):1139–1152. doi: https://doi.org/10.1016/j.cell.2014.01.050.

Furutachi S, Franklin AD, Aldea AM, Mrsic-Flogel TD, Hofer SB. Cooperative thalamocortical circuit mechanism for sensory prediction errors. Nature 2024 Sep;633(8029):398–406. doi: https://doi.org/10.1038/s41586-024-07851-w.

Garrett M, Manavi S, Roll K, Ollerenshaw DR, Groblewski PA, Ponvert ND, et al. Experience shapes activity dynamics and stimulus coding of VIP inhibitory cells. eLife 2020 feb;9:e50340. doi: https://doi.org/10.7554/eLife.50340.

Garrett ME, Nauhaus I, Marshel JH, Callaway EM. Topography and Areal Organization of Mouse Visual Cortex. Journal of Neuroscience 2014;34(37):12587–12600. doi: https://doi.org/10.1523/JNEUROSCI.1124-14.2014.

Gerstner W, Lehmann M, Liakoni V, Corneil D, Brea J. Eligibility traces and plasticity on behavioral time scales: experimental support of neohebbian three-factor learning rules. Frontiers in neural circuits 2018;12:53. doi: https://doi.org/10.3389/fncir.2018.00053.

Hartung J, Schroeder A, Péréz Vázquez RA, Poorthuis RB, Letzkus JJ. Layer 1 NDNF interneurons are specialized top-down master regulators of cortical circuits. Cell Reports 2024 May;43(5):114212. doi: https://doi.org/10.1016/j.celrep.2024.114212.

He K, Huertas M, Hong S, Tie X, Hell J, Shouval H, et al. Distinct Eligibility Traces for LTP and LTD in Cortical Synapses. Neuron 2015 Nov;88(3):528–538. doi: https://doi.org/10.1016/j.neuron.2015.09.037.

Hennequin G, Agnes EJ, Vogels TP. Inhibitory Plasticity: Balance, Control, and Codependence. Annual Review of Neuroscience 2017;40(Volume 40, 2017):557–579. doi: https://doi.org/10.1146/annurev-neuro-072116-031005.

Hertäg L, Sprekeler H. Learning prediction error neurons in a canonical interneuron circuit. eLife 2020 aug;9:e57541. doi: https://doi.org/10.7554/eLife.57541.

Hioki H, Okamoto S, Konno M, Kameda H, Sohn J, Kuramoto E, et al. Cell Type-Specific Inhibitory Inputs to Dendritic and Somatic Compartments of Parvalbumin-Expressing Neocortical Interneuron. The Journal of Neuroscience 2013 Jan;33(2):544–555. doi: https://doi.org/10.1523/JNEUROSCI.2255-12.2013.

Ibata K, Sun Q, Turrigiano GG. Rapid Synaptic Scaling Induced by Changes in Postsynaptic Firing. Neuron 2008 Mar;57(6):819–826. doi: https://doi.org/10.1016/j.neuron.2008.02.031.

Ji Xy, Zingg B, Mesik L, Xiao Z, Zhang LI, Tao HW. Thalamocortical Innervation Pattern in Mouse Auditory and Visual Cortex: Laminar and Cell-Type Specificity. Cerebral Cortex 2015 05;26(6):2612–2625. doi: https://doi.org/10.1093/cercor/bhv099.

Jiang X, Shen S, Cadwell CR, Berens P, Sinz F, Ecker AS, et al. Principles of connectivity among morphologically defined cell types in adult neocortex. Science 2015;350(6264):aac9462. doi: https://doi.org/10.1126/science.aac9462.

Johnson RR, Burkhalter A. A Polysynaptic Feedback Circuit in Rat Visual Cortex. Journal of Neuroscience 1997;17(18):7129–7140. doi: https://doi.org/10.1523/JNEUROSCI.17-18-07129.1997.

Kaleb K, Pedrosa V, Clopath C. Network-centered homeostasis through inhibition maintains hippocampal spatial map and cortical circuit function. Cell Reports 2021 Aug;36(8):109577. doi: https://doi.org/10.1016/j.celrep.2021.109577.

Keck T, Keller GB, Jacobsen RI, Eysel UT, Bonhoeffer T, Hübener M. Synaptic scaling and homeostatic plasticity in the mouse visual cortex in vivo. Neuron 2013;80(2):327–334. doi: https://doi.org/10.1016/j.neuron.2013.08.018.

Keck T, Toyoizumi T, Chen L, Doiron B, Feldman DE, Fox K, et al. Integrating Hebbian and homeostatic plasticity: the current state of the field and future research directions. Philosophical Transactions of the Royal Society B: Biological Sciences 2017 Mar;372(1715):20160158. doi: https://doi.org/10.1098/rstb.2016.0158.

Kilman V, Van Rossum MC, Turrigiano GG. Activity deprivation reduces miniature IPSC amplitude by decreasing the number of postsynaptic GABAA receptors clustered at neocortical synapses. Journal of Neuroscience 2002;22(4):1328–1337. doi: https://doi.org/10.1523/jneurosci.22-04-01328.2002.

Kim J, Tsien RW, Alger BE. An improved test for detecting multiplicative homeostatic synaptic scaling. PloS one 2012;7(5):e37364. doi: https://doi.org/10.1371/journal.pone.0037364.

Lagzi F, Bustos MC, Oswald AM, Doiron B. Assembly formation is stabilized by Parvalbumin neurons and accelerated by Somatostatin neurons. bioRxiv 2021; doi: https://doi.org/10.1101/2021.09.06.459211.

Lam SK, Pitrou A, Seibert S. Numba: a LLVM-based Python JIT compiler. In: Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC Austin Texas: ACM; 2015. p. 1–6. doi: https://doi.org/10.1145/2833157.2833162.

Larkum ME, Nevian T, Sandler M, Polsky A, Schiller J. Synaptic Integration in Tuft Dendrites of Layer 5 Pyramidal Neurons: A New Unifying Principle. Science 2009;325(5941):756–760. doi: https://doi.org/10.1126/science.1171958.

Lazarus MS, Huang ZJ. Distinct maturation profiles of perisomatic and dendritic targeting GABAergic interneurons in the mouse primary visual cortex during the critical period of ocular dominance plasticity. Journal of Neurophysiology 2011 Aug;106(2):775–787. doi: https://doi.org/10.1152/jn.00729.2010.

Leman DP, Cary BA, Bissen D, Lane BJ, Shanley MR, Wong NF, et al. Rapid prey capture learning drives a slow resetting of network activity in rodent binocular visual cortex. bioRxiv 2025; doi: https://doi.org/10.1101/2025.04.11.648036.

Letzkus JJ, Wolff SBE, Meyer EMM, Tovote P, Courtin J, Herry C, et al. A disinhibitory microcircuit for associative fear learning in the auditory cortex. Nature 2011 Dec;480(7377):331–335. doi: https://10.1038/nature10674.

London M, Häusser M. DENDRITIC COMPUTATION. Annual Review of Neuroscience 2005;28(Volume 28, 2005):503–532. doi: https://doi.org/10.1146/annurev.neuro.28.061604.135703.

Miehl C, Gjorgjieva J. Stability and learning in excitatory synapses by nonlinear inhibitory plasticity. PLOS Computational Biology 2022 Dec;18(12):e1010682. doi: https://doi.org/10.1371/journal.pcbi.1010682.

Naumann LB, Hertäg L, Müller J, Letzkus JJ, Sprekeler H. Layer-specific control of inhibition by NDNF interneurons. Proceedings of the National Academy of Sciences 2025;122(4):e2408966122. doi: https://doi.org/10.1073/pnas.2408966122.

Ottenheimer DJ, Hjort MM, Bowen AJ, Steinmetz NA, Stuber GD. A stable, distributed code for cue value in mouse cortex during reward learning 2023 Jun; doi: https://10.7554/elife.84604.2.

Pakan JM, Francioni V, Rochefort NL. Action and learning shape the activity of neuronal circuits in the visual cortex. Current Opinion in Neurobiology 2018;52:88–97. doi: https://doi.org/10.1016/j.conb.2018.04.020.

Palmigiano A, Fumarola F, Mossing DP, Kraynyukova N, Adesnik H, Miller KD. Common rules underlying optogenetic and behavioral modulation of responses in multi-cell-type V1 circuits. bioRxiv 2023; doi: https://doi.org/10.1101/2020.11.11.378729.

Pardi MB, Vogenstahl J, Dalmay T, Spanò T, Pu DL, Naumann LB, et al. A thalamocortical top-down circuit for associative memory. Science 2020;370(6518):844–848. doi: https://doi.org/10.1126/science.abc2399.

Park E, Mosso MB, Barth AL. Neocortical somatostatin neuron diversity in cognition and learning. Trends in Neurosciences 2025 Jan;p. S0166223624002492. doi: https://doi.org/10.1016/j.tins.2024.12.004.

Pawlak V. Timing is not everything: neuromodulation opens the STDP gate. Frontiers in Synaptic Neuroscience 2010;2. doi: https://doi.org/https://doi.org/10.3389/fnsyn.2010.00146.

Peng H, Xie P, Liu L, Kuang X, Wang Y, Qu L, et al. Morphological diversity of single neurons in molecularly defined cell types. Nature 2021 Oct;598(7879):174–181. doi: https://doi.org/10.1038/s41586-021-03941-1.

Pfeffer CK, Xue M, He M, Huang ZJ, Scanziani M. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. Nature Neuroscience 2013 Aug;16(8):1068–1076. doi: https://doi.org/10.1038/nn.3446.

Poirazi P, Brannon T, Mel BW. Pyramidal Neuron as Two-Layer Neural Network. Neuron 2003 Mar;37(6):989–999. doi: https://doi.org/10.1016/S0896-6273(03)00149-1.

Pozo K, Goda Y. Unraveling Mechanisms of Homeostatic Synaptic Plasticity. Neuron 2010 May;66(3):337–351. doi: https://doi.org/10.1016/j.neuron.2010.04.028.

Prestigio C, Ferrante D, Marte A, Romei A, Lignani G, Onofri F, et al. REST/NRSF drives homeostatic plasticity of inhibitory synapses in a target-dependent fashion. Elife 2021;10:e69058. doi: https://doi.org/10.7554/eLife.69058.

Richter LMA, Gjorgjieva J. A circuit mechanism for independent modulation of excitatory and inhibitory firing rates after sensory deprivation. Proceedings of the National Academy of Sciences 2022;119(32):e2116895119. doi: https://doi.org/10.1073/pnas.2116895119.

Schneider-Mizell CM, Bodor AL, Brittain D, Buchanan J, Bumbarger DJ, Elabbady L, et al. Inhibitory specificity from a connectomic census of mouse visual cortex. Nature 2025 Apr;640(8058):448–458. doi: https://doi.org/10.1038/s41586-024-07780-8.

Schulz A, Miehl C, Berry I Michael J, Gjorgjieva J. The generation of cortical novelty responses through inhibitory plasticity. eLife 2021 oct;10:e65309. doi: https://doi.org/10.7554/eLife.65309.

Shen S, Jiang X, Scala F, Fu J, Fahey P, Kobak D, et al. Distinct organization of two cortico-cortical feedback pathways. Nature Communications 2022 Oct;13(1):6389. doi: https://doi.org/10.1038/s41467-022-33883-9.

Swanwick CC, Murthy NR, Kapur J. Activity-dependent scaling of GABAergic synapse strength is regulated by brain-derived neurotrophic factor. Molecular and Cellular Neuroscience 2006;31(3):481–492. doi: https://doi.org/10.1016/j.mcn.2005.11.002.

Tetzlaff C, Kolodziejski C, Timme M, Wörgötter F. Synaptic Scaling in Combination with Many Generic Plasticity Mechanisms Stabilizes Circuit Connectivity. Frontiers in Computational Neuroscience 2011;volume 5 - 2011. doi: https://doi.org/10.3389/fncom.2011.00047.

Torrado Pacheco A, Bottorff J, Gao Y, Turrigiano GG. Sleep Promotes Downward Firing Rate Homeostasis. Neuron 2021 Feb;109(3):530–544.e6. doi: https://doi.org/10.1016/j.neuron.2020.11.001.

Tremblay R, Lee S, Rudy B. GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. Neuron 2016 Jul;91(2):260–292. doi: https://doi.org/10.1016/j.neuron.2016.06.033.

Turrigiano GG. The Self-Tuning Neuron: Synaptic Scaling of Excitatory Synapses. Cell 2008 Oct;135(3):422–435. doi: https://doi.org/10.1016/j.cell.2008.10.008.

Turrigiano GG, Leslie KR, Desai NS, Rutherford LC, Nelson SB. Activity-dependent scaling of quantal amplitude in neocortical neurons. Nature 1998;391(6670):892–896. doi: https://doi.org/10.1038/36103.

Van Rossum MCW, Bi GQ, Turrigiano GG. Stable Hebbian Learning from Spike Timing-Dependent Plasticity. The Journal of Neuroscience 2000 Dec;20(23):8812–8821. doi: https://doi.org/10.1523/JNEUROSCI.20-23-08812.2000.

Veit J, Handy G, Mossing DP, Doiron B, Adesnik H. Cortical VIP neurons locally control the gain but globally control the coherence of gamma band rhythms. Neuron 2023 Feb;111(3):405–417.e5. doi: https://doi.org/10.1016/j.neuron.2022.10.036.

Waitzmann F, Wu YK, Gjorgjieva J. Top–down modulation in canonical cortical circuits with short-term plasticity. Proceedings of the National Academy of Sciences 2024;121(16):e2311040121. doi: https://doi.org/10.1073/pnas.2311040121.

Watt AJ. Homeostatic plasticity and STDP: keeping a neuron's cool in a fluctuating world. Frontiers in Synaptic Neuroscience 2010;2. doi: https://doi.org/10.3389/fnsyn.2010.00005.

Wen W, Turrigiano GG. Keeping Your Brain in Balance: Homeostatic Regulation of Network Function. Annual Review of Neuroscience 2024 Aug;47(1):41–61. doi: https://doi.org/10.1146/annurev-neuro-092523-110001.

Wilmes KA, Clopath C. Inhibitory microcircuits for top-down plasticity of sensory representations. Nature Communications 2019 Nov;10(1):5055. doi: https://doi.org/10.1038/s41467-019-12972-2.

Wu CH, Ramos R, Katz DB, Turrigiano GG. Homeostatic synaptic scaling establishes the specificity of an associative memory. Current biology 2021;31(11):2274–2285. doi: https://doi.org/10.1016/j.cub.2021.03.024.

Wu YK, Hengen KB, Turrigiano GG, Gjorgjieva J. Homeostatic mechanisms regulate distinct aspects of cortical circuit dynamics. Proceedings of the National Academy of Sciences 2020 Sep;117(39):24514–24525. doi: https://doi.org/10.1073/pnas.1918368117.

Wu YK, Miehl C, Gjorgjieva J. Regulation of circuit organization and function through inhibitory synaptic plasticity. Trends in Neurosciences 2022 Dec;45(12):884–898. doi: https://doi.org/10.1016/j.tins.2022.10.006.

Yagishita S, Hayashi-Takagi A, Ellis-Davies GCR, Urakubo H, Ishii S, Kasai H. A critical time window for dopamine actions on the structural plasticity of dendritic spines. Science 2014;345(6204):1616–1620. doi: https://doi.org/10.1126/science.1255514.

Zhang S, Xu M, Kamigaki T, Do JPH, Chang WC, Jenvay S, et al. Long-range and local circuits for top-down modulation of visual cortex processing. Science 2014;345(6197):660–665. doi: https://doi.org/10.1126/science.1254126.

Supplementary Material

Table S1: Model Parameters

Network dynamics and network connectivity			
Symbol	**Value**	**Unit**	**Description**
τ_E	20	ms	time constant of E rate dynamics
τ_P	5	ms	time constant of PV rate dynamics
τ_S	10	ms	time constant of SST rate dynamics
$W_{EE_{ii}}$	0.51	a.u.	connection strength from E to E within subnetworks
$W_{EE_{ij}}$	0.51	a.u.	connection strength from E to E across subnetworks
$W_{EP_{ii}}$	0.91	a.u.	connection strength from PV to E within subnetworks
$W_{EP_{ij}}$	0.41	a.u.	connection strength from PV to E across subnetworks
$W_{ES_{ii}}$	0.51	a.u.	connection strength from SST to E within subnetworks
$W_{ES_{ij}}$	0.31	a.u.	connection strength from SST to E across subnetworks
$W_{PE_{ii}}$	0.3	a.u.	connection strength from E to PV within subnetworks
$W_{PE_{ij}}$	0.1	a.u.	connection strength from E to PV across subnetworks
$W_{PP_{ii}}$	0.2	a.u.	connection strength from PV to PV within subnetworks
$W_{PP_{ij}}$	0.1	a.u.	connection strength from PV to PV across subnetworks
$W_{PS_{ii}}$	0.3	a.u.	connection strength from SST to PV within subnetworks
$W_{PS_{ij}}$	0.1	a.u.	connection strength from SST to PV across subnetworks
$W_{SE_{ii}}$	0.4	a.u.	connection strength from E to SST within subnetworks
$W_{SE_{ij}}$	0.1	a.u.	connection strength from E to SST across subnetworks
ρ	1.5	a.u.	rheobase current for E, PV, and SST
Plasticity mechanisms			
τ_{hebb}	4	min	time constant of three-factor Hebbian learning
τ_θ	24	h	time constant of target activity dynamics
τ_β	28	h	time constant of target activity regulator dynamics
τ_{ss}	8	h	time constant of synaptic scaling dynamics
k	0.25	a.u.	amplitude of the decrease in target activity regulator
Inputs			
g_E	4.5	a.u.	background input to E
g_P	3.2	a.u.	background input to PV
g_S	3	a.u.	background input to SST
$g_{\Delta E}$	1	a.u.	additional input to E
$g_{\Delta P}$	0.5	a.u.	additional input to P
$g_{\Delta S}$	0.5 (-0.5)	a.u.	additional top-down excitatory (inhibitory) input to SST

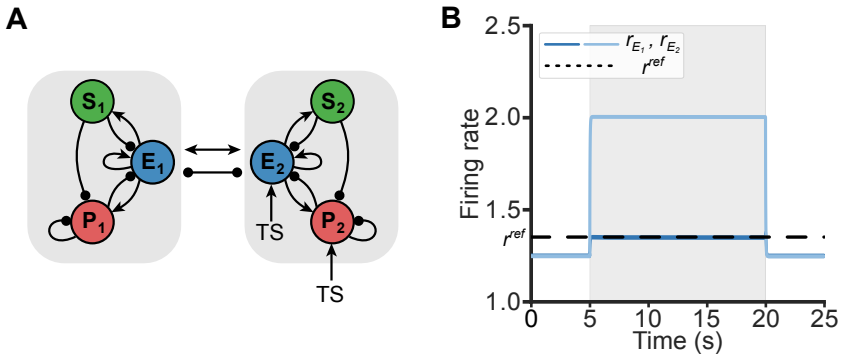

Fig. S1. Definition of reference activity. **A.** Network schematic when applying a conditioned stimulus to subnetwork 2 in the absence of unconditioned stimulus. **B.** The reference activity r^{ref} is determined by measuring the excitatory population activity of subnetwork 1 in response to the test stimulus in the absence of plasticity. The test stimulus is applied during the period marked in gray, and modeled by providing additional inputs to the excitatory and PV population of subnetwork 2.

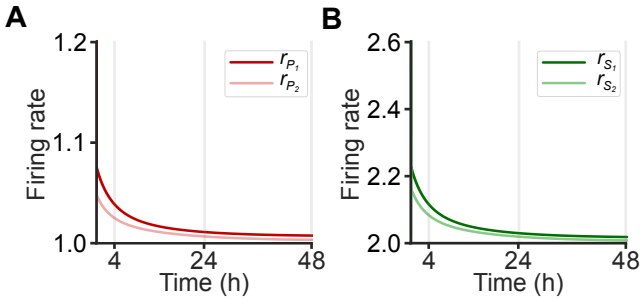

Fig. S2. Inhibitory population activity post-conditioning. **A.** Activity of PV population in subnetwork 1 (r_{P_1}) and subnetwork 2 (r_{P_2}) after conditioning. **B.** Same as A but for SST population.

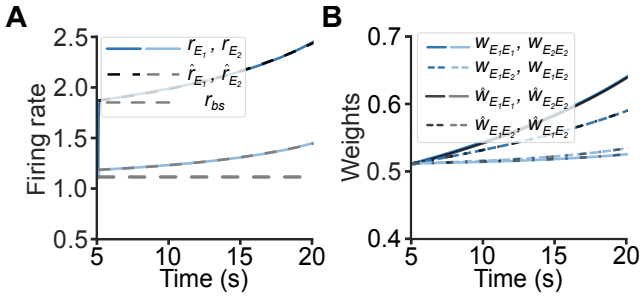

Fig. S3. Comparisons between numerical simulations and analytical calculations during conditioning in Phase 1. **A.** Simulated excitatory activity of subnetwork 1 (r_{E_1}, same as in Figure 2) and subnetwork 2 (r_{E_2}, same as in Figure 2) during conditioning, and fitted excitatory activity of subnetwork 1 ($\hat{r}_{E_1}$) and subnetwork 2 ($\hat{r}_{E_2}$) during conditioning. The horizontal dashed line represents the baseline activity level measured before conditioning. **B.** Numerical simulated excitatory weight evolution (same as in Figure 2) during conditioning, and analytical calculated excitatory weights evolution during conditioning.

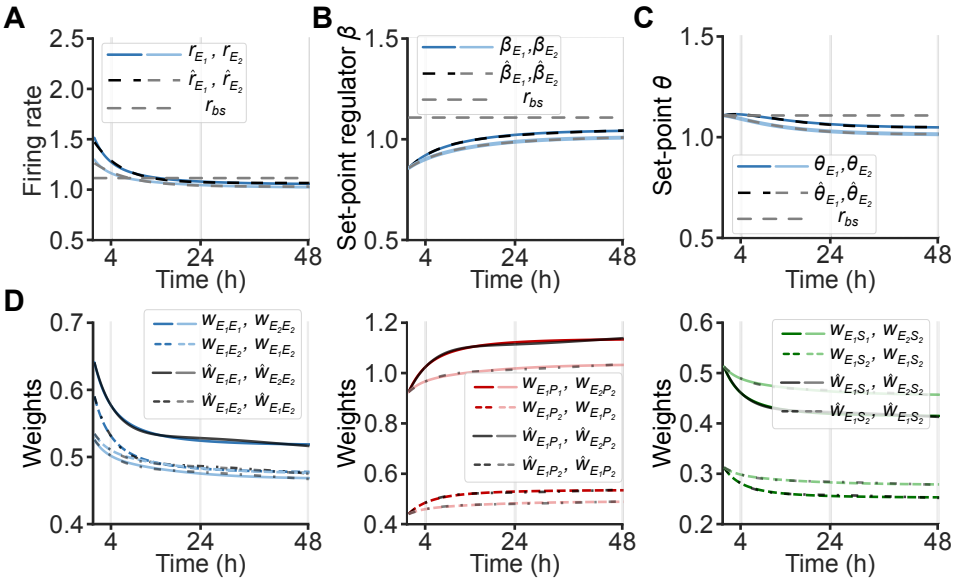

Fig. S4. Comparisons between numerical simulations and analytical calculations post-conditioning in Phase 2. **A.** Simulated excitatory activity of subnetwork 1 (r_{E_1}, same as in Figure 3) and subnetwork 2 (r_{E_2}, same as in Figure 3) after conditioning, and fitted excitatory activity of subnetwork 1 ($\hat{r}_{E_1}$) and subnetwork 2 ($\hat{r}_{E_2}$) after conditioning. The horizontal dashed line represents the baseline activity level measured before conditioning. **B.** Same as A but for set point regulators after conditioning. **C.** Same as B but for set points after conditioning. **D.** (Left) Numerical simulated excitatory weight evolution (same as in Figure 3) after conditioning, and analytical calculated excitatory weights evolution after conditioning. (Middle, Right) Same as Left but for PV-to-E and SST-to-E weights respectively.

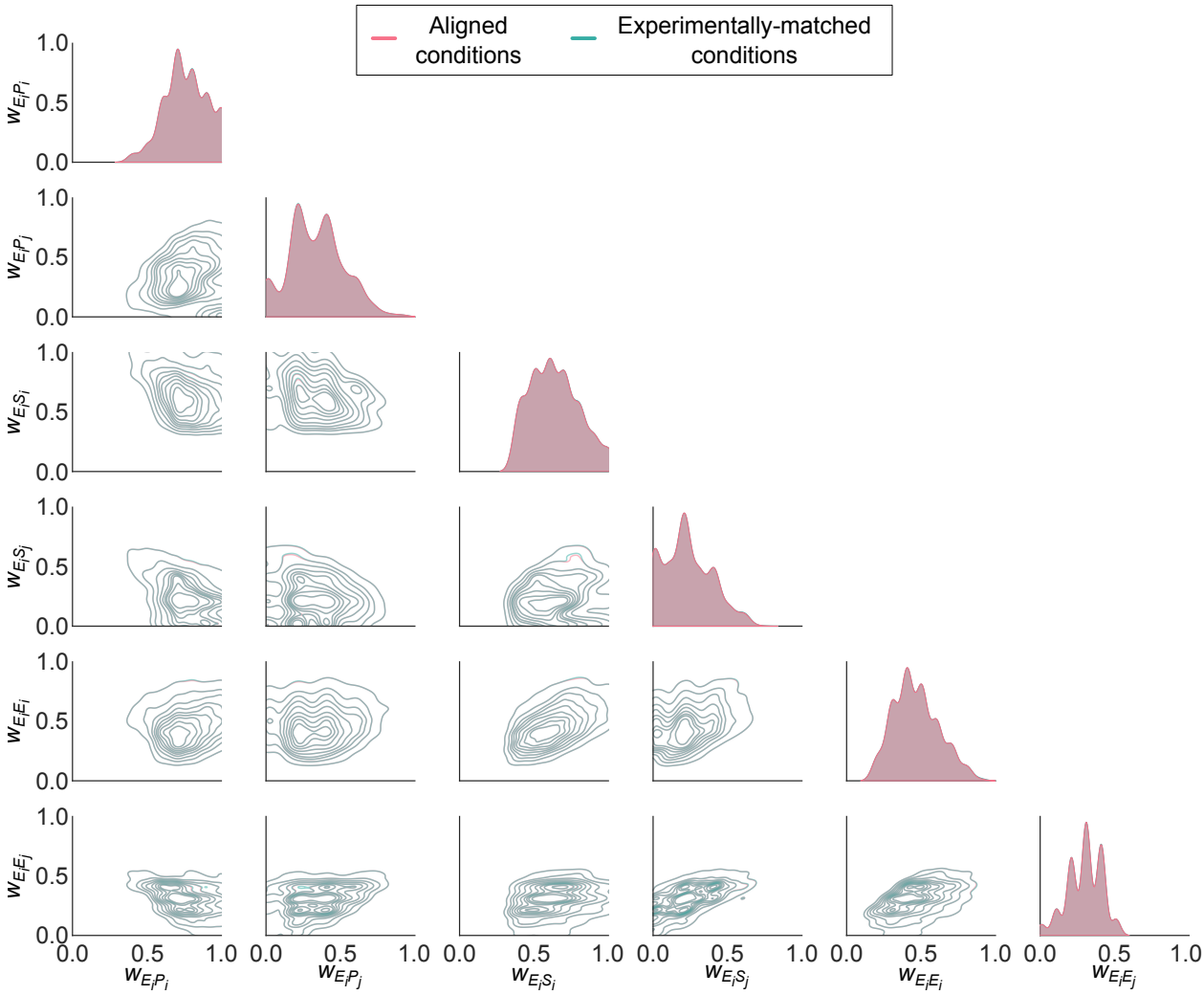

Fig. S5. Pairwise density plots of plastic weight parameter distributions for experimentally matched models and aligned models. Experimentally matched models are defined as those in which the transition from memory generalization to specificity occurs between 4h and 24h when all scaling mechanisms are active and shifts to between 24h and 48h when E-to-E scaling is blocked. Aligned models are a subset of experimentally matched models, characterized by a diminished memory specificity when PV-to-E scaling is blocked and an accelerated transition when SST-to-E scaling is blocked. The x- and y-axes represent specific plastic connection weights. The substantial overlap between the two distributions suggests that cell-type-specific contributions to associative learning demonstrated in Figure 7 are robust to parameter selection. Here, 98% of the experimentally matched models (811 out of 828) are aligned models.

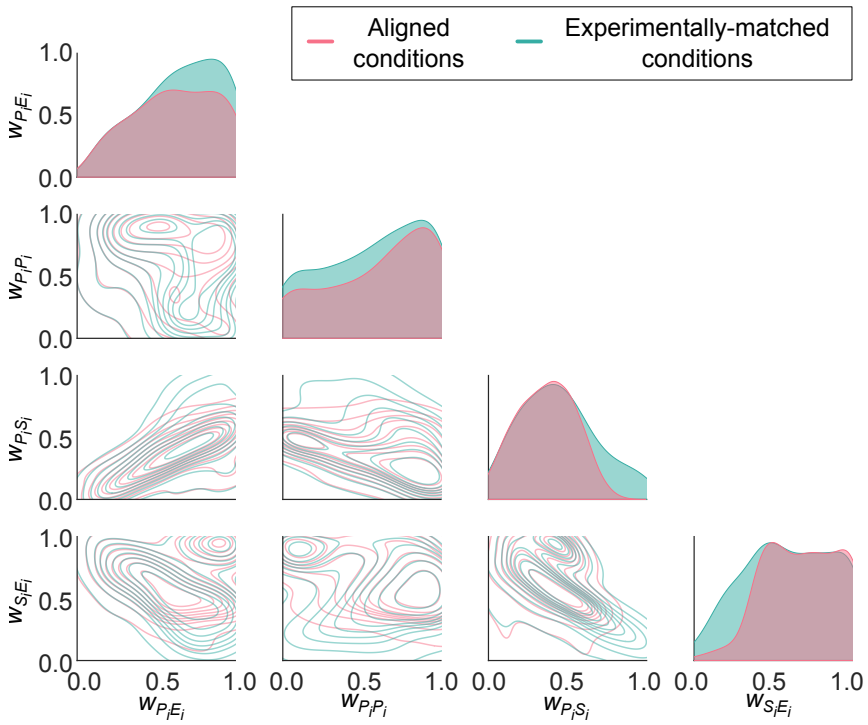

Fig. S6. Pairwise density plots of static weight parameter distributions for experimentally matched models and aligned models. Experimentally matched models are defined as those in which the transition from memory generalization to specificity occurs between 4h and 24h when all scaling mechanisms are active and shifts to between 24h and 48h when E-to-E scaling is blocked. Aligned models are a subset of experimentally matched models, characterized by a diminished memory specificity when PV-to-E scaling is blocked and an accelerated transition when SST-to-E scaling is blocked. The x- and y-axes represent specific static connection weights. The substantial overlap between the two distributions suggests that cell-type-specific contributions to associative learning demonstrated in Figure 7 are robust to parameter selection. Here, 89% of the experimentally matched models (230 out of 258) are aligned models.