Challenging Backpropagation: Evidence for Target Learning in the Neocortex

Pau Vilimelis Aceituno[*,1,3], Sander de Haan[*,1,2,3],
Reinhard Loidl[1], Lhea Beumer[1], and Benjamin F. Grewe[1,3,†]

[1]Institute of Neuroinformatics, University of Zurich and ETH Zurich
[2]Institute of Social Ethics, University of Lucerne
[3]ETH AI Center, ETH Zurich
[†]Correspondence: bgrewe@ethz.ch
[*]Equal contribution

September 1, 2025

1 Abstract

Computational neuroscience currently debates two competing hypotheses to explain hierarchical learning in the neocortex: deep learning inspired approximations of the backpropagation algorithm, where neurons adjust synapses to minimize an error, and target learning algorithms, where neurons learn by reducing the feedback needed to achieve a desired target activity. While both hypotheses have been supported by theoretical studies, there is currently no empirical test that compares them directly. Here we provide such test in the mouse neocortex by analyzing in vivo data from pyramidal neurons, finding that the target learning hypothesis more accurately predicts the neural activity during learning. To consolidate our in vivo findings, we conduct additional in vitro experiments that clarify the relationship between algorithmic learning signals and synaptic plasticity within individual pyramidal neurons. By combining in vivo and in vitro data we reveal a critical discrepancy between neocortical hierarchical learning and canonical machine learning.

2 Main

Learning in the cortex occurs through synaptic plasticity, where the strength of connections between neurons change based on experience. Whether learning refers to new behaviors, associations between stimuli, or simply a more efficient encoding of frequent stimuli, learning always requires solving the credit assignment problem: to make changes in neural activity or behavior, the brain must identify the relevant synapses and adjust their strengths accordingly. Since cortical neuronal networks contain multiple layers and complex pathways between sensory input and motor output, tracing back through the network to identify which specific synapses ought to change is difficult (**Fig. 1a**). In canonical artificial neural networks, this is addressed through backpropagation (BP), an algorithm that propagates error gradients backward through the network to adjust each weight in proportion to its contribution to the output error (Rumelhart et al., 1986). While BP has proven highly effective, its biological plausibility remains contentious (Crick, 1989; Lillicrap et al., 2020; Song et al., 2024). BP relies on several assumptions that are difficult to reconcile with the biological properties of the neocortex, including precise, layer-specific error signals, symmetric feedforward and feedback weights (Liao et al., 2016; Lillicrap et al., 2016; Whittington and Bogacz, 2019), and continuous, differentiable activity, all of which contrast with the spiking, asymmetric, and locally governed nature of cortical circuits. Modern BP-inspired frameworks have relaxed these biologically implausible assumptions of classical BP and thereby put BP forth as a hypothesis to explain learning in cortex, through mechanisms such as dendritic error propagation (Sacramento et al., 2018), equilibrium propagation (Scellier

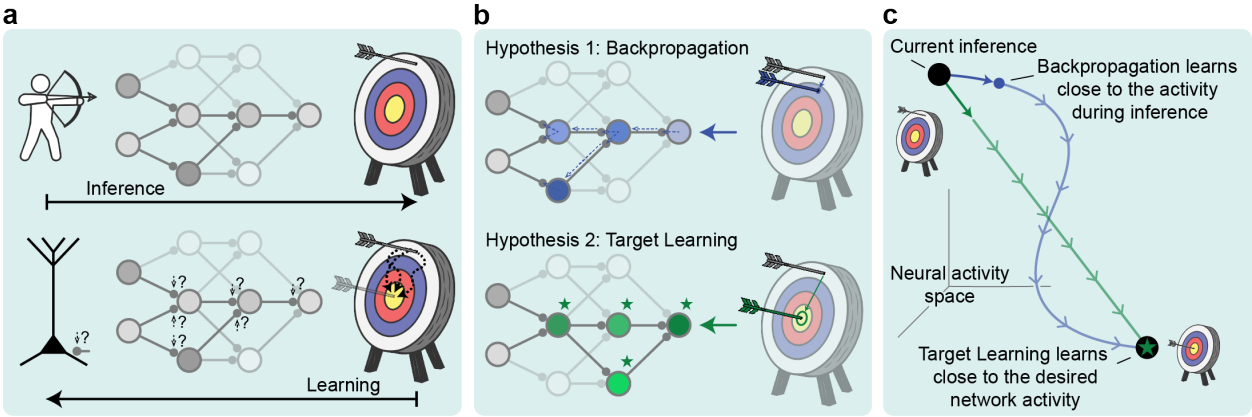

Figure 1: Illustration of BP and TL. (**a**) Illustration of the credit assignment problem in the cortex. (**b**) Illustration of the BP and TL hypotheses on the neural activities during learning in a network. BP relays the gradient information of the error to the individual synapses and minimally deviates from inference activity, while TL first enforces a target activity to every neuron whereafter synapses adjust to facilitate this activity. (**c**) Illustration of the BP and TL hypotheses in neural activity space. BP predicts activity during learning to be close to activity during inference, while TL predicts that activity during learning is close to the desired network activity.

and Bengio, 2017), feedback alignment (Lillicrap et al., 2016; Nøkland, 2016), and others (Whittington and Bogacz, 2017; Akrout et al., 2019; Amit, 2019).

In parallel, a recently growing body of work proposes an alternative cortical learning hypothesis referred to here as target learning (TL). Rather than transmitting explicit error signals, TL frameworks posit that learning occurs by imposing target activity patterns on neurons via feedback or control signals. Synaptic plasticity then adjusts the network such that these target patterns can eventually be produced without the need of an additional learning signal (Rao and Ballard, 1999; Capone et al., 2023; Song et al., 2024). TL thus reframes learning not as the minimization of a global error, but as the local alignment of neuronal activity with desired internal states (**Fig. 1b**) (Meulemans et al., 2021; Song et al., 2024).

Both families of algorithms have advanced our theoretical understanding of learning in neural systems (Körding and König, 2001; Richards et al., 2019; Song et al., 2024). However, different kinds of learning frameworks or circuits can approximate both BP and TL – such as predictive coding (Whittington and Bogacz, 2017; Song et al., 2024), energy models (Scellier and Bengio, 2017), or dendritic error circuits (Sacramento et al., 2018; Meulemans et al., 2021) – emphasizing that theory alone cannot resolve this distinction. Thus, the algorithmic principles employed by the neocortex to solve the credit assignment problem remains an open question, even after decades of debate (Crick, 1989; Lillicrap et al., 2020; Saxe et al., 2021).

In this work, we investigate the BP and TL hypotheses for cortical learning by examining pyramidal neuron (PN) plasticity. To test the two hypotheses, we leverage a crucial difference in their neuronal activity during learning (**Fig. 1c**), and test the predictions of this difference in a recent in vivo dataset (Nguyen et al., 2024). To complement our in vivo findings, we present a detailed mechanistic explanation of learning within individual PNs, which we validate through in vitro experiments.

3 Results

To understand how cortical neuronal networks learn, we start by denoting a mathematical formalism that captures learning rules used by single neurons. With this in mind, we next formulate a test that is capable of differentiating between the BP and TL hypotheses at the neuronal population level. We then justify using an in vivo dataset (Nguyen et al., 2024) upon which we perform the test, and present our findings. To consolidate our findings, we detail the intracellular processes in PNs, theoretically relating teaching signals to plasticity, whereafter we conduct an in vitro experiment to directly test whether our in vivo results are

consistent with in vitro results from individual PNs.

3.1 Theoretical basis

To make a valid comparison between BP and TL, we first take a look at their respective weight update rules and specifically contrast the terms they consist of. Weight update rules formalize how the synaptic strength between two given neurons must change to enforce neuronal activity that produces the correct output given a specific input. In both BP and TL, the learning rule consists of two terms: the presynaptic activity and the necessary change in postsynaptic activity to bring the network closer to the desired output.

However, the terms do not capture the same activations in BP and TL. While the postsynaptic term reflects a change from current to desired activity that is similar in both algorithms, the presynaptic term differs in a more consequential way. In BP, the presynaptic activation represents the activity during previous stimulus presentation. In contrast, TL imposes a target activity onto the presynaptic neuron and uses this instead of the activity that occurs during stimulus presentation. This target activity equals the activity that the network will converge to after training (**Fig. 1c**). We formalize the weight update rules as:

$$\Delta w = \begin{cases} a_{\text{pre}}\delta^{BP}_{\text{post}}, & \text{for BP} \quad (1a) \\ a^{\star}_{\text{pre}}\delta^{TL}_{\text{post}}, & \text{for TL} \quad (1b) \end{cases}$$

where a_{pre} reflects the presynaptic neuronal activity when the stimulus is presented, $a^{\star}_{\text{pre}}$ represents the target presynaptic activity, and $\delta^{BP}_{\text{post}}$ together with $\delta^{TL}_{\text{post}}$ indicate how the postsynaptic neuron should change its activity for BP and TL respectively. A mathematical derivation for both rules can be found in the **Supp. Text 1**.

When comparing **Eq. 1a** with **Eq. 1b**, their similarity is apparent. In fact, the computational machinery of BP can be used to implement TL (Ernoult et al., 2022; LeCun, 1986) and vice versa (Whittington and Bogacz, 2017; Scellier and Bengio, 2017). However, doing so requires adding mechanisms and parameters to the basic algorithms, making the modified algorithms less parsimonious (**Supp. Text 2**). Therefore, any empirical evidence that favors the basic version of either BP or TL could only be supportive of the other algorithm if the extra mechanisms are also accounted for in the data. Thus, despite the similarities, the basic principle guiding how the two algorithms steer neuronal activity toward desired outcomes predicts fundamentally different observations of neuronal activity during training phases. These differences allow for the formulation of two incompatible and testable hypotheses:

- The neuronal activity during learning **resembles the activity during inference** – the last activity in response to a stimulus – if the cortex learns through BP.
- The neuronal activity during learning **resembles the activity after learning** – the target state – if the cortex learns through TL.

To evaluate these hypotheses, we next introduce a dataset capturing neuronal activity during both inference and learning phases, enabling a direct test of these predictions.

3.2 Cortical reactivations isolate activity during inference and learning

To distinguish whether cortical learning aligns better with TL or BP predictions, we need experimental data that captures neuronal activity during learning. We thus have two main requirements: (1) the data must reflect long-term neuronal changes driven by learning, and (2) the neuronal activity that drives this learning needs to be identifiable.

We adopt data from a recent experimental study which recorded somatic L2/3 PN activity from the mouse lateral visual cortex with two-photon calcium imaging (Nguyen et al., 2024). Here, mice were presented two stimuli in random order, where each stimulus lasts for two seconds followed by a gray screen for fifty-eight seconds before the next stimulus onset (**Fig. 2a**). During one session, the mouse was exposed to a total of 120 stimulus presentations, sixty for every stimulus. Importantly, the inter-stimuli intervals contain cortical reactivations: spontaneous reoccurrences of neuronal activity patterns resembling those seen during stimulus presentations (Rothschild et al., 2017). These reactivations were classified according to which of the two

stimuli they resembled most, and reactivations are considered local in time. We now analyze whether this data fulfills our requirements.

For requirement (1), we assess whether the recordings capture long-term neuronal changes driven by learning. Because the in vivo dataset uses a passive exposure paradigm, one can argue that no behavioral learning is present, and thereby no learning is required. However, the study shows that the representations of the two stimuli in neuronal activity space orthogonalize during the sessions, where the representations remain stable over days, indicating an underlying optimization goal that is achieved. This orthogonalization is also visible in the overlap of highly active neurons decreasing during training (**Fig. 2b**). This orthogonalization could arise from simple learning paradigms such as Hebbian-like learning rules (Dayan and Abbott, 2005), which would require not require sophisticated algorithms such as TL or BP. To assess whether this is the case, we note that Hebbian-like learning rules would maintain the relative ranking of neuronal activations during learning, because highly active postsynaptic neurons potentiate more (Frémaux et al., 2010). The data shows significant changes in the identity of the most active neurons, suggesting that changes in activity are driven by learning algorithms more complex than classical Hebbian learning (**Fig. 2c**).

For requirement (2), the original study suggests that the reactivations are predictive of the observed long-term changes in activity, which would suggest that reactivations are the drivers of learning, in line with previous studies showing a correlation between learning and reactivations (Chen and Wilson, 2023), as well as with direct causal tests where blocking reactivations prevented animals from learning altogether (Jadhav et al., 2012), while prolonging them enhanced learning (Fernández-Ruiz et al., 2019). To verify this at the single neuron level, we tested possible predictors of activity changes in subsequent stimulus-evoked responses. Specifically, we assessed how well previous stimulus-evoked responses, reactivations, or their differences predicted the sign change of the current stimulus-evoked event, indicating the direction of learning. We found that the difference between the previous stimulus response and its associated reactivation best predicts these changes (**Fig. 2d**), suggesting that the neuronal activity during reactivations drives learning at the single neuron level.

We conclude that the dataset from Nguyen et al. (2024) provides neuronal recordings capturing long-term plasticity, allowing us to isolate activity patterns during inference and learning, which are critical for testing our algorithmic hypotheses.

3.3 Target learning explains reactivations in cortical networks

Drawing on our previous theoretical distinctions between BP and TL and associated hypotheses (**Eq. 1** and **Section 3.1**), we can now make specific predictions about the trajectory of neural activity in the in vivo dataset from (Nguyen et al., 2024). Since the dataset includes recordings of both stimulus-evoked neuronal responses (during inference) and cortical reactivations (during learning), we specify our predictions as:

- If learning is based on BP, neuronal population activity during reactivations will resemble the most recent stimulus-evoked response for a given stimulus.
- If learning is based on TL, neuronal population activity during reactivations will resemble the stimulus-evoked activity at the late-learning stage, when neuronal representations have stabilized.

Note that these predictions are not distinguishable during the late stages of learning, as the activity during inference is by definition very close to the activity after learning, resulting in neuronal activities during reactivations that are very similar to both previous and final stimulus evoked responses. Therefore, we focus our analysis solely on early reactivations. Every step of our data analysis is detailed in **Methods 1**.

To test our hypotheses, we first project the stimulus-evoked responses into a low-dimensional space, reducing the amount of variability that is unrelated to our question. We define the first dimension as the vector from early to late stimulus-evoked responses, reflecting the learning progression. This learning-progression dimension reflects how neuronal responses evolve from the early initial activity states to stabilized and orthogonalized late training states, and therefore is the dimension of principle interest to study activity changes. We set subsequent dimensions as principal components orthogonal to this vector. We present the stimulus-evoked responses (for one example mouse stimulus) in **Fig. 2e**, starting at early points and converging toward late responses. We illustrate how BP aligns its hypothesized reactivations with the most

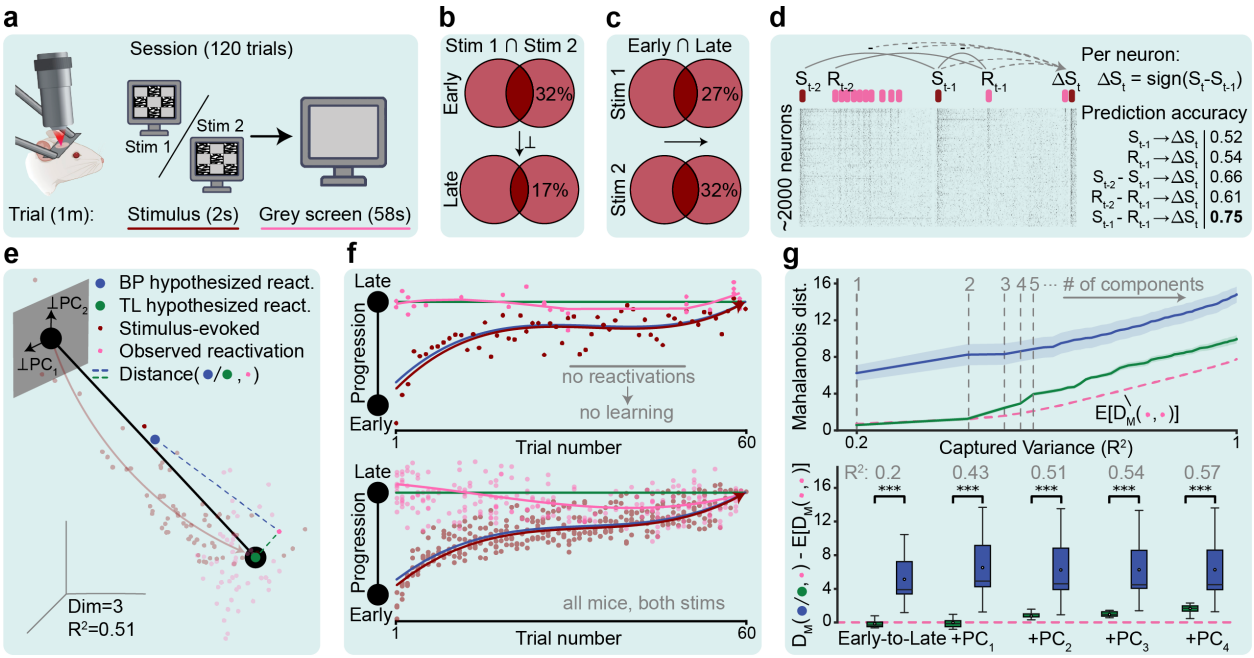

Figure 2: Target learning aligns with in vivo cortical learning dynamics. (**a**) Experimental setup from Nguyen et al. (2024) uses two alternating stimuli with resting periods, recording neuronal activity via two-photon calcium imaging to detect reactivations from stimulus-evoked responses. (**b**) The overlap between the top 5 percent of most active neurons for each stimuli decreases from early to late training (Early: $32.2 \pm 0.7\%$ SEM, Late: $16.9 \pm 0.9\%$ SEM) suggesting orthogonalization. (**c**) The identities of the top 5 percent most active neurons change significantly for each stimuli (Stim 1: $27.0 \pm 1.8\%$ SEM, Stim 2: $31.8 \pm 4.0\%$ SEM). (**d**) Reactivations relative to preceding stimuli best predict activity changes (pred. acc. = 0.75 vs. random = 0.5). (**e**) Stimulus-evoked responses and reactivations trace trajectories in PC space, with the first component reflecting early-to-late changes. Distances are measured between hypothesized and observed reactivations. (**f**) Reactivations align with TL along the early-to-late component for one mouse (top) and all mice (bottom), with no learning observed without reactivations. (**g**) Mahalanobis distance shows TL explains the observed reactivations better than BP after subtracting expected distances as the number of PCs is increased, capturing a larger fraction of variance in the reactivations, (first five shown, Wilcoxon, $p < 0.0001$).

recent stimulus-evoked activity, while in TL the network activity present in late stimulus-evoked responses corresponds to the target used for the reactivations.

We observe that the reactivations align with the late rather than recent stimulus-evoked responses, in line with (Nguyen et al., 2024), thus supporting TL rather than BP. We show the stability of reactivations around the TL hypothesis along the temporal learning dimension from early to late responses in **Fig. 2f**. This additionally justifies our data requirements by demonstrating the absence of learning when no reactivations are present. To quantify the distinction between TL and BP, we next assess the distance between the hypothesized and observed reactivations, and test which hypothesis better fits the observed reactivations. We use the Mahalanobis distance (intuitively, the z-score in high dimensions) which measures the separation between data points while accounting for the inverse covariance in the neuronal activity, enhancing robustness against noisy and uninformative dimensions in the reactivation patterns. We add principal components incrementally to account for more variance of the neuronal data beyond our principal learning-progression dimension, and observe that TL consistently has a lower distance than BP (**Fig. 2g**), indicating that the TL-predicted reactivations are closer to the observed reactivations. To account for the intrinsic increase in variability when considering more and more dimensions, we subtract the minimum expected distance due to the variance in the reactivations (pink dotted line) from the distances measured by the two hypotheses. Within this corrected space, we observe that the TL hypothesis results in a Mahalanobis distance that

is within the range of the expected distance in the space of the first two dimensions (covering a fraction of variance in the data of $R^2 = 0.2$, and $R^2 = 0.43$ respectively), and remains smaller than for the BP hypothesis for the next components. The TL hypothesis is thus much closer to the minimum expected distance (Wilcoxon test, $p < 0.0001$ for all values of captured variance). These observations are consistent across mice and stimuli (**Fig. S1**).

3.4 Mechanisms for target-driven basal plasticity in PNs

Having established that cortical reactivations in the in vivo dataset support the TL hypothesis over BP, we now ask how the learning signals of these high-level algorithmic principles relate to synaptic plasticity at the level of PNs.

We start by isolating the TL learning signals in the vivo dataset. Models implementing TL (Song et al., 2024; Meulemans et al., 2021) often use a version of Eq. 1b, where the postsynaptic term driving the change in neuronal activity in the postsynaptic neuron is implemented as

$$\delta^{TL}_{\text{post}} \propto a^{\star}_{\text{post}} - a_{\text{post}} \tag{2}$$

where $a^{\star}_{\text{post}}$ is the target postsynaptic activity, and a_{post} the postsynaptic activity during inference. As our in vivo analysis suggests that reactivations set the target (**Fig. 2d&f**), we posit that $a^{\star}_{\text{post}}$ corresponds to the neuronal activity during reactivations, and a_{post} to the stimulus evoked responses. Thus, we expect that changes in somatic activity during reactivations correlate with the changes in somatic activity induced by plasticity (**Methods 2**). To test this, we calculate the difference of PN activities between two subsequent stimulus presentations – the change due to plasticity – and correlate that value with the difference between the stimulus-evoked responses and the reactivations. We find that the correlation is high (**Fig. 3a**, $R^2 = 0.80$).

As the neural activity during reactivations corresponds to the target activity, the inputs to PNs that trigger those reactivations should correspond to the neuron-level learning signal in TL. Experimental studies show that cortical reactivations are driven by apical inputs to PNs (Sirota et al., 2003; Khodagholy et al., 2017), suggesting that the learning signal arrives at the apical dendrite. This aligns with existing observations and theories that propose that PNs process learning signals through their apical dendrite (Spratling, 2002; Guerguiev et al., 2017), as well as modeling studies that build on that hypothesis (Sacramento et al., 2018; Meulemans et al., 2021; Aceituno et al., 2023). Furthermore, apical inputs also have a strong effect on the responses of PNs to stimuli (Larkum et al., 1999; Williams and Holtmaat, 2019), indicating that they could be used to both drive plasticity and enforce a target firing rate in PNs neurons, as would be expected from TL (see **1**).

Having isolated apical inputs as a biological correlate of algorithmic learning signals, we now investigate the subcellular processes by which they drive plasticity and neuronal activity. At the PN level, apical inputs that are strong enough to elicit a local dendritic event (e.g. an NMDA spike) are distant from the soma, and have only little effect on somatic activity Yuste et al. (1994). However, the combination of apical inputs and a backpropagating action potential (bAP) can engage voltage-gated conductance at the apical dendrite, triggering dendritic Ca^{2+} plateau events (BAC firing). These apical plateaus also elevate the somatic membrane potential, promoting bursting Larkum et al. (1999). Thus, apical inputs can determine whether a somatic action potential-elicited bAP will trigger an additional apical plateau potential. Both bAPs and plateau potentials can propagate along the dendrites to basal synapses, conveying information about somatic activity (Markram et al., 1997).

At basal synapses, the depolarization induced by single bAPs or plateaus can have strong effects on plasticity. Long-lasting depolarization – as would be induced by a plateau – can effectively relieve the Mg^{2+} block of NMDA receptors during presynaptic glutamate release, thereby increasing Ca^{2+} influx into the spine Clopath et al. (2009); Bock et al. (2022). In turn, the level of calcium determines long-term changes in synaptic strength (Lisman, 1989; Yang et al., 1999; Nevian and Sakmann, 2006; Graupner and Brunel, 2012). Thus, a plateau would generate enough membrane depolarization to induce long term potentiation (LTP) (Clopath et al., 2009), while single bAPs fail to induce LTP in physiological conditions (Inglebert et al., 2020), and can instead induce long term depression (LTD) (Chindemi et al., 2022).

To better understand the interactions between apical input, somatic bursting, and basal synaptic plasticity, we built a simplified neuron model capturing the dynamics of voltage and calcium in the apical tuft, the soma, and the synapse, (see **Fig. 3b** for a visualization and **Methods 3** for details about the model).

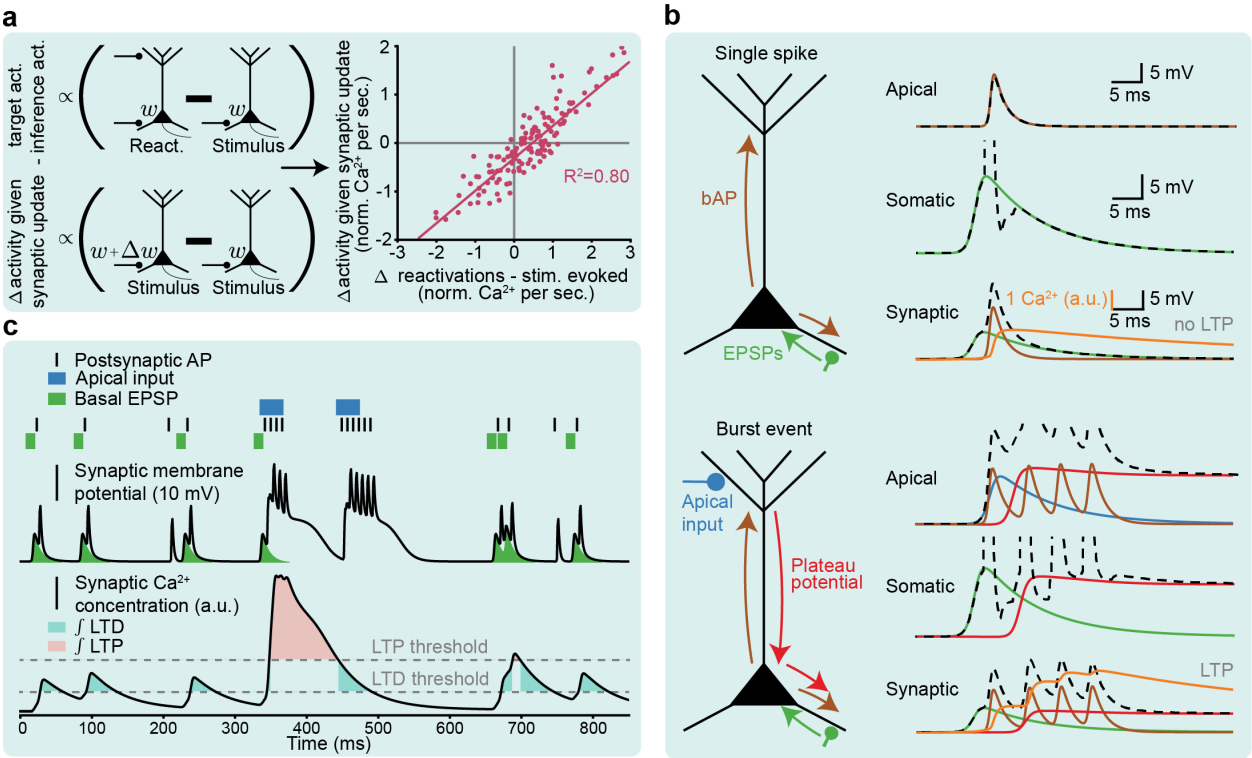

Figure 3: The mechanistic integration of basal and apical inputs in PNs. (**a**) Schema of the variables defining the postsynaptic learning signal (left, up) and the postsynaptic change in activity due to plasticity (left, down), and the correlation between both variables as measured in the in vivo dataset (right, $R^2 = 0.80$). (**b**) Neuron schematic (left) details synaptic depolarization events and their interactions in the PN. Responses of each compartment are shown for apical, somatic, and synaptic compartments for cases without (top) and with (bottom) top-down input to the apical dendrite. (**c**) Simulation of the effects of apical input on somatic firing rate, synaptic membrane potential, calcium influx, and plasticity. For a specific basal synapse, the causation of postsynaptic APs and apical input (top) depolarize the synaptic membrane potential (middle) and modulate the synaptic calcium concentrations leading to LTD and LTP (bottom).

Using our model we derive a learning rule similar to **Eq. 1**, where the postsynaptic term $\delta^{TL}_{\text{post}}$ resembles that of **Eq. 2** with the target activity being represented through the ratio of bursts and single action potentials, which is itself determined by the rate of apical inputs (see **Supp. Text 3**).

Our model aligns with theoretical studies using burst-driven learning (Payeur et al., 2021) or apical dendritic error mechanisms (Meulemans et al., 2020; Aceituno et al., 2023) that affect both activity and plasticity. However, the link between somatic bursts and basal plasticity is not a causal relationship, but rather a reflection of the fact that both are driven by plateau potentials (see **Fig. 3c**). In the next section we test the relationship between those three phenomena – basal synaptic plasticity, plateau potentials, and changes in somatic activity – by experimentally manipulating apical inputs.

3.5 Apical input directs synaptic plasticity in basal synapses

To test whether apical inputs facilitate plasticity in basal synapses and somatic activity changes experimentally, we conducted whole-cell patch clamp experiments on 12 L5 pyramidal neurons from the prefrontal cortex of six mice, and used extracellular stimulation of basal and apical afferents (**Fig. 4a**, **Methods 4**). Recordings were performed in physiological extracellular calcium concentrations (Inglebert et al., 2020).

We recorded baseline EPSPs for five minutes to assess initial synaptic responses. We increased basal stimulation intensity to induce one action potential and adjusted apical stimulation to trigger supra-threshold

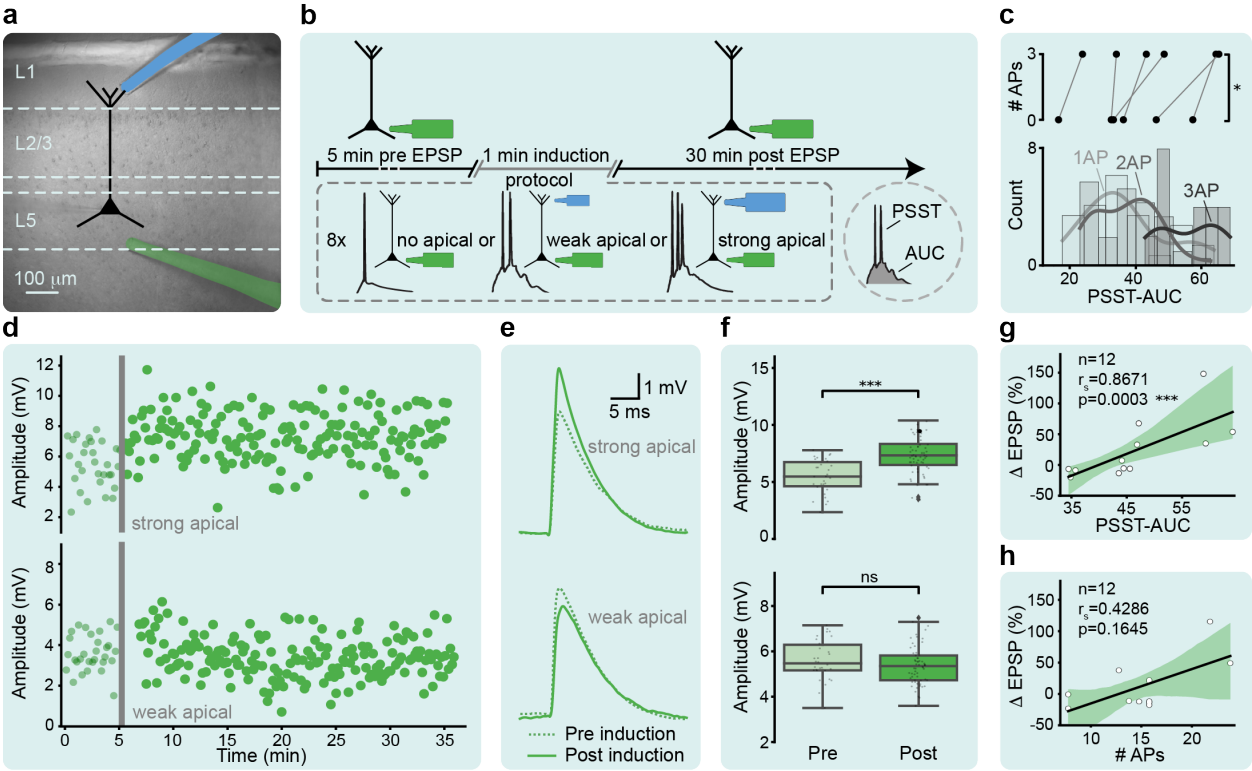

Figure 4: Apical inputs direct changes in synaptic strength in basal synapses and affect PN activity. (**a**) Wide-field image of mouse neocortex with an L5 pyramidal neuron sketch superimposed, and the placement of extracellular basal and apical stimulation electrodes. (**b**) Stimulation and recording protocol. (**c**) Pairwise comparison of PSST-AUC for the minimum and maximum number of somatic APs (top, Wilcoxon, $p = 0.031$). Histogram of the PSST-AUC for one, two, and three APs (bottom). (**d-e-f**) Example recordings of basal EPSP amplitudes before and after a plasticity induction event with significant increase (top) and no-significant difference in synaptic strength (bottom). (**d**) Evolution of the EPSP amplitude over time. (**e**) Mean EPSP for pre and post plasticity-induction event. (**f**) Quantification of ΔEPSP (T-test, top: $p < 0.0001$, bottom: $p = 0.229$), 12 neurons from 6 mice. (**g-h**) Correlation analysis between changes in basal EPSP amplitude and PSST-AUC and number of APs, respectively (Spearman correlation, n=12, **g**: $R_s = 0.6871$, $p = 0.0003$, **h**: $R_s = 0.4286$, $p = 0.1645$).

events with one to three action potentials and prolonged somatic depolarization, suggestive of a plateau potential. After eight paired stimulation-induced supra-threshold events (PSSTs), we recorded EPSPs for 30 minutes to measure long-term amplitude changes (**Fig. 4b**). We quantified somatic depolarization as a proxy for synaptic depolarization by measuring the area under the curve of somatic membrane depolarization during PSSTs (PSST-AUC).

Prolonged somatic depolarization from dendritic plateau potentials enables burst firing (Larkum et al., 1999), showing a dependency between the number of action potentials and PSST-AUC. Here, we find that the number of action potentials increases with PSST-AUC, and this change is significant even in single neurons (**Fig. 4c**, $p = 0.031$).

Applying the strong apical stimulation induction protocol facilitating burst firing increased basally-induced EPSP amplitudes indicating LTP (**Fig. 4d-f top**, $p < 0.0001$), while applying weaker apical stimulation showed no significant change (**Fig. 4d-f bottom**, $p = 0.229$). To assess basal synaptic plasticity predictors, we correlated the number of action potentials and PSST-AUC with EPSP amplitude changes (ΔEPSP). PSST-AUC showed a significant positive correlation with ΔEPSP ($R_{Spearman} = 0.8671$, $p = 0.0003$, **Fig. 4g**), whereas the correlation with the number of action potentials was not significant ($R_{Spearman} = 0.4286$, $p = 0.1645$, **Fig. 4h**). Thus, total somatic depolarization predicts synaptic strength

changes better than the number of action potentials.

In contrast, changes in apical synaptic strength do not correlate with basal synaptic plasticity (**Fig. S2**, **Supp. Text 4**, $r_s = 0.2308$, $p = 0.4705$), suggesting distinct plasticity rules for apical and basal synapses.

Our in vitro experiments show that basal synaptic plasticity in L5 PNs is driven by prolonged somatic depolarization modulated by apical inputs, consistent with modeling studies using PNs with apical inputs as learning signals (Sacramento et al., 2018; Meulemans et al., 2021). Plateau potentials and burst firing co-occur, aligning with burst-based learning rules (Payeur et al., 2021). However, the changes in somatic firing do not correlate significantly with plasticity, but are another effect of plateau potentials. Basal LTP correlates with somatic bursting, aligning with the TL prediction of feedback-driven learning changes, though inhibitory circuits or frequency multiplexing could potentially counteract these bursts (Sacramento et al., 2018; Payeur et al., 2021).

We note that our in vitro electrophysiology experiments targeted L5 PNs to avoid cross-excitation between apical and basal electrodes, but our in vivo results involved L2/3 neurons. Despite anatomical differences, both L5 and L2/3 PNs show prolonged depolarizations when bAPs coincide with distal synaptic inputs (Larkum et al., 2007; Gambino et al., 2014), indicating a shared plasticity gating mechanism.

4 Discussion

Our study combines theory as well as in vivo and in vitro experimental data, suggesting that cortical dynamics during learning align more closely with target learning (TL) rather than backpropagation (BP). This finding addresses a longstanding debate in neuroscience and machine learning about the type of learning algorithm the cortex employs (Grossberg, 1987; Crick, 1989; Körding and König, 2001; Richards et al., 2019; Lillicrap et al., 2020; Song et al., 2024). Furthermore, our in vitro experiments show that the plasticity of PNs can be modulated by apical inputs and expressed through the presence of somatic bursting, complementing existing bio-plausible models of hierarchical learning (Sacramento et al., 2018; Meulemans et al., 2021; Payeur et al., 2021) and providing a mechanistic explanation for their learning rule.

Our results complement previous theoretical studies which have proposed various circuits and potential architectures for learning in the neocortex – such as predictive coding (Rao and Ballard, 1999), equilibrium propagation (Scellier and Bengio, 2017), learning through dendritic errors (Sacramento et al., 2018), and deep feedback control (Meulemans et al., 2021). Importantly, we do not test the architectures or circuits themselves, but we focus instead on comparing the dynamics of PNs and cortical networks during learning. This paradigm shift is crucial, because BP and TL are mathematical frameworks which are not tied to any specific formalism, and can thus be implemented by similar network architectures or circuits (Scellier and Bengio, 2017; Whittington and Bogacz, 2017, 2019; Song et al., 2024). By considering only the neuronal activity patterns of TL and BP, we argue that our study is able to clarify the algorithmic principles driving learning in the neocortex.

However, by focusing solely on PNs, we do not address the specific circuit requirements needed to implement this family of algorithms. To address this gap, future work shall clarify the circuit-level role, origin, and generation of feedback signals in TL architectures (Rao and Ballard, 1999; Scellier and Bengio, 2017; Meulemans et al., 2022b). We provide a first step in doing so by measuring apical plasticity in L5 PNs, which has been hypothesized to mediate feedback signals (Williams and Holtmaat, 2019; Meulemans et al., 2021). However, we did not find any consistent relation between PN activity and apical plasticity, most likely due to the unspecific nature of our stimulation. As we stimulate inputs to apical dendrites using extracellular electrodes, we probably jointly activate axons from the contralateral prefrontal cortex (cPFC) as well as thalamic nuclei, specifically the mediodorsal (MD) and ventromedial (VM) thalamus. Those axonal feedback projections are likely mediated by different neuron types, which could explain the heterogeneous plasticity we observed. Future experimental research needs to disentangle the contributions of different afferent feedback projections to elicit apical inputs and drive plasticity, for example, using projection-specific optogenetic methods. Those methods could also bypass the problem of using extracellular electrodes to isolate basal and apical inputs, and could therefore be used in L2/3 PNs to validate our results in those layers.

In contrast to previous theoretical studies (Sacramento et al., 2018; Whittington and Bogacz, 2017; Scellier and Bengio, 2017; Meulemans et al., 2021; Song et al., 2024), the in vivo task that we consider here is unsupervised and simple, as in classical models (Cottrell et al., 1987; Rao and Ballard, 1999). More

complex tasks most likely would require more complex targets, making such tests more convoluted. For example, in a decision task with two stimuli with a specific movement associated to each stimulus, the final target activity might not be clear from the beginning, as the animal needs a neuronal activity for each stimulus, a neuronal activity for each movement, and eventually a single target for each contingency. Thus, we might have target activity that changes as the animal learns the task. This would require a more sophisticated analysis and possibly in vivo data collected from multiple areas. To disentangle the possibly flexible nature of targets, future studies should therefore investigate reactivations in other modalities such as motor learning, and multi-modal learning.

In summary, our study combines deep learning theory with in vivo cortical data and in vitro electrophysiology. Our investigation bridges multiple scales of neuronal learning suggesting that cortical hierarchical networks learn by enforcing targets rather than correcting errors, thus differing from their canonical artificial counterparts.

5 Author contributions

P.V.A. conceptualized the project and performed the mathematical analysis. P.V.A. and S.d.H. developed the in vivo data analysis. S.d.H. developed the neuron model and performed the data analysis. R.L. conceptualized and performed the electrophysiological experiments, and provided biology insights. L.B. aided the in vivo data analysis. S.d.H. created the figures. B.F.G. supervised the project. P.V.A., S.d.H., R.L., and B.F.G. wrote the paper.

6 Author information

The authors have declared no competing interest.

7 Data and code availability statement

Experimental data and analysis code will be made public upon acceptance

8 Acknowledgements

This work was supported by the Swiss National Science Foundation (B.F.G. CRSII5-173721 and 315230 189251, P.V.A. 182539), ETH project funding (B.F.G. ETH-20 19-01), and a Forschungskredit from the University of Zürich (P.V.A).

References

Aceituno, P. V., Farinha, M. T., Loidl, R., and Grewe, B. F. (2023). Learning Cortical Hierarchies with Temporal Hebbian Updates. *bioRxiv* doi:10.1101/2023.01.02.522459

Akrout, M., Wilson, C., Humphreys, P., Lillicrap, T., and Tweed, D. B. (2019). Deep learning without weight transport. *Advances in neural information processing systems* 32

Amit, Y. (2019). Deep learning with asymmetric connections and hebbian updates. *Frontiers in computational neuroscience* 13, 18

Bock, T., Negrean, A., and Siegelbaum, S. A. (2022). Somatic depolarization enhances hippocampal ca1 dendritic spike propagation and distal input-driven synaptic plasticity. *Journal of Neuroscience* 42, 3406–3425

Capone, C., Lupo, C., Muratore, P., and Paolucci, P. S. (2023). Beyond spiking networks: the computational advantages of dendritic amplification and input segregation. *Proceedings of the National Academy of Sciences* 120, e2220743120

Chen, Z. S. and Wilson, M. A. (2023). How our understanding of memory replay evolves. *Journal of Neurophysiology* 129, 552–580

Chindemi, G., Abdellah, M., Amsalem, O., Benavides-Piccione, R., Delattre, V., Doron, M., et al. (2022). A calcium-based plasticity model for predicting long-term potentiation and depression in the neocortex. *Nature Communications* 13, 3038. doi:10.1038/s41467-022-30214-w

Clopath, C., Büsing, L., Vasilaki, E., and Gerstner, W. (2009). Connectivity reflects coding: A model of voltage-based spike-timing-dependent-plasticity with homeostasis. *Nature Precedings* , 1–1

Cottrell, G. W., Munro, P., and Zipser, D. (1987). Learning internal representation from gray-scale images: An example of extensional programming. *Proceedings of the Annual Meeting of the Cognitive Science Society* 9

Crick, F. (1989). The recent excitement about neural networks. *Nature* 337, 129–132. doi:10.1038/337129a0

Dayan, P. and Abbott, L. F. (2005). *Theoretical neuroscience: computational and mathematical modeling of neural systems* (MIT press)

Ernoult, M. M., Normandin, F., Moudgil, A., Spinney, S., Belilovsky, E., Rish, I., et al. (2022). Towards scaling difference target propagation by learning backprop targets. In *International Conference on Machine Learning* (PMLR), 5968–5987

Fernández-Ruiz, A., Oliva, A., Fermino de Oliveira, E., Rocha-Almeida, F., Tingley, D., and Buzsáki, G. (2019). Long-duration hippocampal sharp wave ripples improve memory. *Science* 364, 1082–1086

Frémaux, N., Sprekeler, H., and Gerstner, W. (2010). Functional requirements for reward-modulated spike-timing-dependent plasticity. *Journal of Neuroscience* 30, 13326–13337

Gambino, F., Pagès, S., Kehayas, V., Baptista, D., Tatti, R., Carleton, A., et al. (2014). Sensory-evoked ltp driven by dendritic plateau potentials in vivo. *Nature* 515, 116–119

Graupner, M. and Brunel, N. (2012). Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. *Proceedings of the National Academy of Sciences* 109, 3991–3996. doi:10.1073/pnas.1109359109

Grossberg, S. (1987). Competitive Learning: From Interactive Activation to Adaptive Resonance. *Cognitive Science* 11, 23–63. doi:10.1111/j.1551-6708.1987.tb00862.x

Guerguiev, J., Lillicrap, T. P., and Richards, B. A. (2017). Towards deep learning with segregated dendrites. *eLife* 6. doi:10.7554/eLife.22901

Gulledge, A. T. and Stuart, G. J. (2003). Action Potential Initiation and Propagation in Layer 5 Pyramidal Neurons of the Rat Prefrontal Cortex: Absence of Dopamine Modulation. *The Journal of Neuroscience* 23, 11363–11372. doi:10.1523/JNEUROSCI.23-36-11363.2003

Inglebert, Y., Aljadeff, J., Brunel, N., and Debanne, D. (2020). Synaptic plasticity rules with physiological calcium levels. *Proceedings of the National Academy of Sciences of the United States of America* 117, 33639–33648. doi:10.1073/pnas.2013663117

Inglebert, Y. and Debanne, D. (2021). Calcium and spike timing-dependent plasticity. *Frontiers in Cellular Neuroscience* 15, 727336

Jadhav, S. P., Kemere, C., German, P. W., and Frank, L. M. (2012). Awake hippocampal sharp-wave ripples support spatial memory. *Science* 336, 1454–1458

Kasper, E. M., Larkman, A. U., Lübke, J., and Blakemore, C. (1994). Pyramidal neurons in layer 5 of the rat visual cortex. I. Correlation among cell morphology, intrinsic electrophysiological properties, and axon targets. *The Journal of Comparative Neurology* 339, 459–474. doi:10.1002/cne.903390402

Khodagholy, D., Gelinas, J. N., and Buzsáki, G. (2017). Learning-enhanced coupling between ripple oscillations in association cortices and hippocampus. *Science* 358, 369–372

Körding, K. P. and König, P. (2001). Supervised and Unsupervised Learning with Two Sites of Synaptic Integration. *Journal of Computational Neuroscience* 11, 207–215. doi:10.1023/A:1013776130161

Larkum, M. E., Senn, W., and Lüscher, H.-R. (2004). Top-down dendritic input increases the gain of layer 5 pyramidal neurons. *Cerebral Cortex (New York, N.Y.: 1991)* 14, 1059–1070. doi:10.1093/cercor/bhh065

Larkum, M. E., Waters, J., Sakmann, B., and Helmchen, F. (2007). Dendritic Spikes in Apical Dendrites of Neocortical Layer 2/3 Pyramidal Neurons. *Journal of Neuroscience* 27, 8999–9008. doi: 10.1523/JNEUROSCI.1717-07.2007

Larkum, M. E., Zhu, J. J., and Sakmann, B. (1999). A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature* 398, 338–341. doi:10.1038/18686

[Dataset] LeCun, Y. (1986). Learning process in an asymmetric threshold network. disordered systems and biological and organizations, leshouches, france

Liao, Q., Leibo, J. Z., and Poggio, T. (2016). How important is weight symmetry in backpropagation? In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence* (Phoenix, Arizona: AAAI Press), AAAI'16, 1837–1844

Lillicrap, T. P., Cownden, D., Tweed, D. B., and Akerman, C. J. (2016). Random synaptic feedback weights support error backpropagation for deep learning. *Nature Communications* 7, 13276. doi: 10.1038/ncomms13276

Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., and Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience* 21, 335–346

Lisman, J. (1989). A mechanism for the Hebb and the anti-Hebb processes underlying learning and memory. *Proceedings of the National Academy of Sciences of the United States of America* 86, 9574–9578. doi: 10.1073/pnas.86.23.9574

Majewska, A., Brown, E., Ross, J., and Yuste, R. (2000). Mechanisms of Calcium Decay Kinetics in Hippocampal Spines: Role of Spine Calcium Pumps and Calcium Diffusion through the Spine Neck in Biochemical Compartmentalization. *The Journal of Neuroscience* 20, 1722–1734. doi: 10.1523/JNEUROSCI.20-05-01722.2000

Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science (New York, N.Y.)* 275, 213–215. doi: 10.1126/science.275.5297.213

McRory, J. E., Santi, C. M., Hamming, K. S. C., Mezeyova, J., Sutton, K. G., Baillie, D. L., et al. (2001). Molecular and Functional Characterization of a Family of Rat Brain T-type Calcium Channels. *Journal of Biological Chemistry* 276, 3999–4011. doi:10.1074/jbc.M008215200

Meulemans, A., Carzaniga, F. S., Suykens, J. A., Sacramento, J., and Grewe, B. F. (2020). A theoretical framework for target propagation. In *Proceedings of the 34th International Conference on Neural Information Processing Systems* (Red Hook, NY, USA: Curran Associates Inc.), NIPS'20, 20024–20036

Meulemans, A., Farinha, M. T., Cervera, M. R., Sacramento, J., and Grewe, B. F. (2022a). Minimizing control for credit assignment with strong feedback. In *International Conference on Machine Learning* (PMLR), 15458–15483

Meulemans, A., Tristany Farinha, M., Garcia Ordonez, J., Vilimelis Aceituno, P., Sacramento, J. a., and Grewe, B. F. (2021). Credit assignment in neural networks through deep feedback control. In *Advances in Neural Information Processing Systems* (Curran Associates, Inc.), vol. 34, 4674–4687

Meulemans, A., Zucchet, N., Kobayashi, S., Von Oswald, J., and Sacramento, J. (2022b). The least-control principle for local learning at equilibrium. *Advances in Neural Information Processing Systems* 35, 33603–33617

Nevian, T. and Sakmann, B. (2006). Spine Ca2+ Signaling in Spike-Timing-Dependent Plasticity. *The Journal of Neuroscience* 26, 11001–11013. doi:10.1523/JNEUROSCI.1749-06.2006

Nguyen, N. D., Lutas, A., Amsalem, O., Fernando, J., Ahn, A. Y.-E., Hakim, R., et al. (2024). Cortical reactivations predict future sensory responses. *Nature* 625, 110–118

Nøkland, A. (2016). Direct feedback alignment provides learning in deep neural networks. In *Advances in Neural Information Processing Systems* (Curran Associates, Inc.), vol. 29, 1037–1045

Payeur, A., Guerguiev, J., Zenke, F., Richards, B. A., and Naud, R. (2021). Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits. *Nature Neuroscience* 24, 1010–1019. doi:10.1038/s41593-021-00857-x

Rao, R. P. N. and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience* 2, 79–87. doi:10.1038/4580

Richards, B. A., Lillicrap, T. P., Beaudoin, P., Bengio, Y., Bogacz, R., Christensen, A., et al. (2019). A deep learning framework for neuroscience. *Nature neuroscience* 22, 1761–1770

Rothschild, G., Eban, E., and Frank, L. M. (2017). A cortical–hippocampal–cortical loop of information processing during memory consolidation. *Nature neuroscience* 20, 251–259

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature* 323, 533–536. doi:10.1038/323533a0

Sacramento, J., Ponte Costa, R., Bengio, Y., and Senn, W. (2018). Dendritic cortical microcircuits approximate the backpropagation algorithm. In *Advances in Neural Information Processing Systems* (Curran Associates, Inc.), vol. 31, 8735–8746

Saxe, A., Nelli, S., and Summerfield, C. (2021). If deep learning is the answer, what is the question? *Nature Reviews Neuroscience* 22, 55–67

Scellier, B. and Bengio, Y. (2017). Equilibrium propagation: Bridging the gap between energy-based models and backpropagation. *Frontiers in computational neuroscience* 11, 24

Sirota, A., Csicsvari, J., Buhl, D., and Buzsáki, G. (2003). Communication between neocortex and hippocampus during sleep in rodents. *Proceedings of the National Academy of Sciences* 100, 2065–2069

Slutsky, I., Sadeghpour, S., Li, B., and Liu, G. (2004). Enhancement of Synaptic Plasticity through Chronically Reduced Ca2+ Flux during Uncorrelated Activity. *Neuron* 44, 835–849. doi:10.1016/j.neuron.2004.11.013

Song, Y., Millidge, B., Salvatori, T., Lukasiewicz, T., Xu, Z., and Bogacz, R. (2024). Inferring neural activity before plasticity as a foundation for learning beyond backpropagation. *Nature Neuroscience* , 1–11

Spratling, M. W. (2002). Cortical region interactions and the functional role of apical dendrites. *Behavioral and Cognitive Neuroscience Reviews* 1, 219–228. doi:10.1177/1534582302001003003

Stuart, G., Schiller, J., and Sakmann, B. (1997). Action potential initiation and propagation in rat neocortical pyramidal neurons. *The Journal of Physiology* 505, 617–632

Whittington, J. C. and Bogacz, R. (2019). Theories of Error Back-Propagation in the Brain. *Trends in Cognitive Sciences* 23, 235–250. doi:10.1016/j.tics.2018.12.005. Publisher: Elsevier

Whittington, J. C. R. and Bogacz, R. (2017). An Approximation of the Error Backpropagation Algorithm in a Predictive Coding Network with Local Hebbian Synaptic Plasticity. *Neural Computation* 29, 1229–1262. doi:10.1162/NECO_a_00949

Williams, L. E. and Holtmaat, A. (2019). Higher-Order Thalamocortical Inputs Gate Synaptic Long-Term Potentiation via Disinhibition. *Neuron* 101, 91–102. doi:10.1016/j.neuron.2018.10.049

Yang, S. N., Tang, Y. G., and Zucker, R. S. (1999). Selective induction of LTP and LTD by postsynaptic [Ca2+]i elevation. *Journal of Neurophysiology* 81, 781–787. doi:10.1152/jn.1999.81.2.781

Yuste, R., Gutnick, M. J., Saar, D., Delaney, K. R., and Tank, D. W. (1994). Ca2+ accumulations in dendrites of neocortical pyramidal neurons: an apical band and evidence for two functional compartments. *Neuron* 13, 23–43. doi:10.1016/0896-6273(94)90457-x

Supplementary Materials for

Challenging Backpropagation: Evidence for Target Learning in the Neocortex

This PDF file includes:

Methods	2
Supplementary Text	10
Figs. S1 to S7	18

Methods

1 In vivo population data analysis

1.1 Data source and preprocessing

We measure the alignment between the activity during learning and during inference or stimulus presentation in the cortical in vivo data available from Nguyen et al. (2024). We modify the existing code to fit the analysis described in the main text and keep the pre-processing as in the original publication (Nguyen et al., 2024). For all cortical data included in our analysis, we used only the top thirty percent of neurons responding to the given stimulus, more than the ten percent as in Nguyen et al. (2024). We denote the early response centroid as the average neural activity of the first three stimulus-evoked responses, and late response centroid as the average neural activity of the last three stimulus-evoked responses. We will use $\mathbf{r}$ to denote the stimulus-evoked responses and ρ to denote reactivations.

1.2 Reactivations predict learning at the single neuron level

To test whether reactivations play a causal role in shaping future stim-evoked activity, we predicted the sign of change (+1 for increase, -1 for decrease) of each neuron for consecutive stimulus-evoked responses. For each day and mouse, we trained numerous logistic regression models with a variety of combinations of predictors to predict this sign of change neuron per event. We used a five-fold cross-validation strategy. For each model type, we aggregated the model accuracies across all sessions and mice. Note that across all mice and days, there were fewer reactivations than stimulus-evoked responses. For balanced comparisons, we calculated a target vector, the direction of for the next stim-evoked response relative to the current response, only for trials that contained at least one reactivation in between. In all models, therefore, the test dataset across all five folds had dimensions of ($\#reactivations \times \#neurons$). In all models the target vector remained the same, whereas the set of predictors differed. For the set of predictors, we tried the following combinations: (1) Current stim-evoked response; (2) Current reactivation; (3) Current stim-evoked response minus previous stim-evoked response; (4) Current reactivation minus previous reactivation; (5) Current stim-evoked response minus current reactivation.

Granger causality was applied to assess whether reactivations enhance predictions of future stim-evoked plasticity, defined by a probability shift when reactivation data is included:

$$\mathbb{P}[Y(t+1) \in A \mid \mathcal{I}(t)] \neq \mathbb{P}[Y(t+1) \in A \mid \mathcal{I}-R(t)] \tag{3}$$

where $Y(t+1)$ is the future change, $\mathcal{I}(t)$ includes past stim-evoked and reactivation activity, and $\mathcal{I}-R(t)$ excludes reactivations.

1.3 Neural activity changes are not directly dependent on the initial neuronal activity

To assess whether credit assignment explains changes in neural patterns, we ranked neurons by activity of the first and last three stimulus-evoked responses per mouse, and computed the amount of overlapping neurons that were in the top 5% most active for both early and late trials. We show the average amount of overlapping neurons between early and late stimulus-evoked responses across mice in **Fig. 2c**. To analyze the changes in stimulus representations, we also computed the overlapping top 5% of neurons in early and late trials between the two stimulus types.

1.4 Measuring progression

To compare the reactivations with the stimulus-evoked responses we consider them as vectors in a space where every dimension corresponds to the activity of one neuron. In this vector space, we can define both the reactivations and the activities, but the baseline activity for single neurons during rest periods – which is when the reactivations take place – is not the same as during active periods (Nguyen et al., 2024).

The measure of progression is shown in **Fig. 2d-e** using data from sample mouse 'NN8' for stimulus 2 from the recording date 12-03-2021. As reported by us and in the original publication (Nguyen et al., 2024), the remaining data looks similar (**Fig. 2e bottom**).

1.5 Hypothesis comparison through distances

We can use projection onto the early-to-late vector and PC components orthogonal to it as a dimensionality reduction procedure that maintains the space where our hypotheses live (the early-to-late vector) as

well as other neural variability. In this space, we can measure the distance between the activity that would be expected in both our hypotheses and what we observe in the data.

Our approach is thus to project the neural activity of the reactivations into the space given by

$$\vec{u}_{\mathbf{d}} = \frac{\mathbf{d}}{\|\mathbf{d}\|},\ \vec{u}^1_{\perp\mathbf{d}} = PC_1(\rho_{\perp\mathbf{d}}),\ \vec{u}^2_{\perp\mathbf{d}} = PC_2(\rho_{\perp\mathbf{d}}), ... \tag{4}$$

where PC_i is the ith principal component operator that is applied to the data projected on the component orthogonal do $\mathbf{d}$. In this new basis, we can measure the noise-adjusted Euclidean (Mahalonobis) distances

$$D_{\mathrm{M}}(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=0}^{n} \frac{(x_i - y_i)^2}{\sigma_i^2}}, \tag{5}$$

where σ_i^2 is the variance captured by the component i (with $i = 0$ being the early-to-late direction). We compute the distances between the reactivations and the target activity (the late stimulus response, $\vec{u}_{\mathbf{d}} = \|d\|,\ \vec{u}^i_{\perp\mathbf{d}} = 0\ \forall i$), and the inference activity (as given by the stimulus-evoked responses), corresponding to the BP or TL hypotheses respectively. We can thus obtain a distance value for each reactivation for each hypotheses, and a Wilcoxon t-test to evaluate whether their difference is significant (**Fig. 2g**, up).

Since the intrinsic distances between points in a measurement with noise grow with the number of dimensions considered, the distance between our hypotheses and the reactivations is expected to grow even if one of the hypothesis is valid. To correct for this effect, we subtract the expected distance between two samples of the same point with different noises. In Mahalanobis space with normally distributed residuals and a covariance matrix normalizing variances, the squared distance between random points follows a chi-squared distribution with n degrees of freedom. With the covariance incorporated, the expected squared distance scales as n, and the expected distance (its square root) adjusts to $\sqrt{n}$. Note that there is a correction for finite dimensions, which reduces the distance to $\sqrt{n - 1/2}$), which we subtract from observed distances to isolate hypothesis-specific variance.

2 Testing alignment between reactivations and target learning rules

TL learning rules often follow an equation of the form

$$\Delta w = a_{\mathrm{pre}} \delta_{\mathrm{apical}} \tag{6}$$

where δ_{apical} is an apical input, which is similar in form to Eq. 1. In contrast to the algorithmic rules from Eq. 1, however, the apical input usually has an effect on the postsynaptic activity, which in the experimental literature is modulatory (Larkum et al., 1999), but is often formalized as additive in computational studies (Sacramento et al., 2018; Meulemans et al., 2021), giving

$$a_{\mathrm{post}}^{\mathrm{apical}} = \begin{cases} a_{\mathrm{post}}^{\mathrm{baseline}}(1 + \delta_{\mathrm{apical}}) \text{ if modulatory} \\ a_{\mathrm{post}}^{\mathrm{baseline}} + \delta_{\mathrm{apical}} \text{ if additive} \end{cases} \tag{7}$$

where $a_{\mathrm{post}}^{\mathrm{baseline}}$ the postsynaptic activity at baseline apical input. In the in vivo data we do not measure apical inputs directly, so we must use the somatic activity only. To isolate the apical input, we can rewrite the previous equation as

$$a_{\mathrm{post}}^{\mathrm{apical}} - a_{\mathrm{post}}^{\mathrm{baseline}} = \begin{cases} a_{\mathrm{post}}^{\mathrm{baseline}} \delta_{\mathrm{apical}} \text{ if modulatory} \\ \delta_{\mathrm{apical}} \text{ if additive} \end{cases} \tag{8}$$

Now we have isolated the change in activity due to the apical inputs, we can isolate the apical inputs for two subsequent presentations of the same stimulus with plasticity in-between. Assuming that the activity in response to a stimulus is done at baseline activity,

$$a_{\mathrm{post}}(s) = \phi\left(\sum_{i \in \mathrm{Pre}} w_i a_i(s)\right) \tag{9}$$

where s is the stimulus count (1 or 2), ϕ is the activation function of the neuron, and Pre is the set of presynaptic neurons. Assuming a similar activity of the presynaptic neurons, we can isolate the δ_{post} term in Eq. 6 as

$$\begin{aligned} a_{\text{post}}(2) - a_{\text{post}}(1) &= \phi\left(\sum_{i\in\text{Pre}} w_i a_i + \sum_{i\in\text{Pre}} \Delta w_i a_i\right) - \phi\left(\sum_{i\in\text{Pre}} w_i a_i\right) \\ &\approx \phi'(z)\left(\sum_{i\in\text{Pre}} \Delta w_i a_i\right) = \phi'(z)\left(\sum_{i\in\text{Pre}} a_i^2\right)\delta_{\text{post}}. \end{aligned} \tag{10}$$

Under the assumption that the activity during stimulus presentation is at baseline apical input, we can then combine **Eq. 10** and **Eq. 8** and recover

$$a_{\text{post}}^{\text{baseline}}(2) - a_{\text{post}}^{\text{baseline}}(1) \approx \beta\left(a_{\text{post}}^{\text{apical}}(1) - a_{\text{post}}^{\text{baseline}}(1)\right) \tag{11}$$

where the term β depends on unknown factors such as the activation function, the presynaptic activities, and the modulatory/additive nature of the apical input. In any case, we do not need to identify the form of β, only show that it is positive which we can do by correlating the left and right terms. To do so in our dataset, we need to assume that $a_{\text{post}}^{\text{apical}}(1)$ is the activtiy during reactivations.

3 Modeling basal synaptic plasticity and pyramidal neuron dynamics

We model basal synaptic plasticity in pyramidal neurons (PNs) based on synaptic calcium as a key driver, adopting a plasticity function P from Graupner and Brunel (2012) that integrates calcium dynamics to determine LTP or LTD. The synaptic strength change is computed as

$$\Delta w = \int P(c(t))dt \tag{12}$$

where Δw is the synaptic strength change and $c(t)$ is the calcium concentration.

Synaptic calcium influx is modeled via NMDA receptor (NMDAR) dynamics, regulated by presynaptic glutamate release and postsynaptic depolarization, using

$$\tau_c \dot{c} = -c + (1 - \frac{c}{c_{\max}}) I(v_{\text{synapse}}, g_{pre}) \tag{13}$$

where τ_c is the calcium time constant, $c_{\max}$ is the maximum calcium concentration, and I represents NMDAR-mediated influx (McRory et al., 2001; Slutsky et al., 2004), with g_{pre} as glutamate presence and v_{synapse} as depolarization.

Postsynaptic membrane depolarization is modeled across synaptic, somatic, and apical compartments, incorporating EPSPs, backpropagating action potentials (bAPs), and plateau potentials via

$$\tau_v \dot{v}_{\text{synapse}} = -v_{\text{synapse}} + \sum_{t\in\text{EPSPs}} \kappa_{\text{EPSP}}^{\text{synapse}} * \delta(t) + \sum_{t\in\text{bAPs}} \kappa_{\text{bAP}}^{\text{synapse}} * \delta(t) + \sum_{t\in\text{Plateaus}} \kappa_{\text{plateau}}^{\text{synapse}} * \delta(t) \tag{14}$$

$$\tau_v \dot{v}_{\text{apical}} = -v_{\text{apical}} + \sum_{t\in\text{EPSPs}} \kappa_{\text{EPSP}}^{\text{apical}} * \delta(t) + \sum_{t\in\text{Plateaus}} \kappa_{\text{plateau}}^{\text{apical}} * \delta(t) + \sum_{t\in\text{bAPs}} \kappa_{\text{bAP}}^{\text{apical}} * \delta(t) \tag{15}$$

$$\tau_v \dot{v}_{\text{soma}} = -v_{\text{soma}} + \sum_{t\in\text{Pre}} \kappa_{\text{pre}}^{\text{soma}} * \delta(t) + \sum_{t\in\text{Plateaus}} \kappa_{\text{plateau}}^{\text{soma}} * \delta(t) + \sum_{t\in\text{APs}} \kappa_{\text{AP}}^{\text{soma}} * \delta(t) \tag{16}$$

where τ_v is the membrane time constant, $\delta(t)$ is the Dirac delta, and κ denotes compartment-specific responses (see **Table 1** for hyperparameters).

Plateau potentials, triggered by coincident bAPs and apical inputs via voltage-gated calcium channels (VGCCs), propagate depolarization, facilitating bAP-activated Ca^{2+} spike (BAC) firing. This bidirectional coupling between compartments modulates synaptic plasticity through NMDAR-regulated calcium influx, with apical input enhancing firing rates and plasticity when paired with basal inputs, consistent with Larkum et al. (2004).

4 Electrophysiology

Parameter	Value	Reference(s)
Duration of AP	1 to 2ms	(Gulledge and Stuart, 2003)
Refractory period	6ms	(Gulledge and Stuart, 2003)
Resting membrane potential	-78 to -74 mV	Own experiments
Peak AP	55 to 65mV	(Kasper et al., 1994)
Number of spikes in burst	2 to 4 spikes	Own experiments
Duration of EPSP	20 to 30ms	Own experiments
bAP half-width	1 to 4 ms	(Stuart et al., 1997)
LTD/LTP threshold	$\theta_{LTP} > \theta_{LTD}$ $\theta_{LTP} \approx 2 \cdot \theta_{LTD}$	(Graupner and Brunel, 2012), (Nevian and Sakmann, 2006)
Calcium spike threshold	2.28 (± 0.14) nA	(Larkum et al., 1999)
Initiation of plateau potential	5 to 10 ms	(Larkum et al., 1999)
Duration of plateau potential	50 to 100 ms	Own experiments
Amplitude of plateau potential	15 mV	(Larkum et al., 1999), Own experiments
NMDAR influx vs voltage	$1/(1 + \exp(-0.22 * (x + 49)))$	(McRory et al., 2001), (Slutsky et al., 2004)
Calcium exp. time constant	81 ± 13 ms	(Majewska et al., 2000)
Maximal calcium in synapse	3 a.u. (see NMDAR influx)	n/a
Basal-apical coupling window	20 to 30 ms	(Larkum et al., 2004)

Table 1: Hyperparameters of the neuron model with their associated value and references.

Male B6/J-Crl1 mice of strain C57BL/6J (orig. Charles River) were bred and maintained at our in-house facility. Animals were kept in their home cages under normal day/night conditions (light: 7am – 7pm). Mouse experiments were approved by the institutional animal care and use committee of Anonymous Institution. All experiments were conducted with the approval of the Anonymous State Veterinary Office.

Mice (11-13 weeks old) were deeply anesthetized using Isoflurane and decapitated. The brains were dissected from the skull and submerged in ice-cold high-sucrose artificial cerebrospinal fluid (ACSF) during acute slice preparation. High sucrose ACSF containing (in mM): 75 Sucrose, 10 Glucose, 80 NaCl, 26 NaHCO3, 7 MgSO4, 2.5 KCl, 1 NaH2PO4, 0.5 CaCl2, fumigated with oxycarbon (95% O2, 5% CO2). Coronal slices (300 µm) were drawn from PFC (+1.4 to +2.6 from bregma) using a VT1200 S vibratome (Leica) and incubated for 30 minutes in oxygenated holding ACSF at 36°C. Holding ACSF containing (in mM): 15 Glucose, 120 NaCl, 26 NaHCO3, 0.9 MgSO4, 2.5 KCl, 1.25 NaH2PO4, 1.3 CaCl2, fumigated with oxycarbon (95% O2, 5% CO2). Magnesium/Calcium ratio was chosen according to Inglebert and Debanne (2021). Slices rested for at least 45 minutes in holding ACSF at room temperature prior to electrophysiological recordings.

During recordings, brain slices were continuously perfused with oxygenated holding ACSF at 36°C. Neurons were visualized with differential interference contrast (DIC) using an Olympus BX61 WI microscope with an Olympus 60x water immersion objective and a CellCam Kikker MT100 (Cairn) DCIM camera. L5 pyramidal neurons (PN) in the PFC were identified based on morphology and distance to the pia and patched with glass pipettes with resistances of 4.8-6.4 MΩ filled with intracellular solution containing (in mM): 10 HEPES, 20 KCl, 117 K-gluconate, 4 Mg-ATP, 0.3 GTP, 10 Na-P-creatine. For morphology reconstruction (**Supp. Fig. S3**), the intracellular solution also contained 0.2% Biocytin. For Ca^{2+} imaging experiments, the intracellular solution also contained 200 µM Oregon Green™ 488 BAPTA-5N. Whole-cell, somatic patch-clamp recordings were acquired in current-clamp mode using an Axon MultiClamp 700A amplifier and digitized using an Axon DigiData 1550 B digital converter. Data was sampled at 50 kHz and filtered at 3 kHz. Typical resting membrane potentials were between -64mV and -72mV, and access resistance ranged from 15 MΩ to 25 MΩ. Recordings in which neurons depolarized > 10% from resting potential or access resistance changed > 15% from initial value were discarded. Bridge potentials were compensated, and liquid-junction potentials were not corrected for. Depolarizing and hyperpolarizing currents were injected to characterize neurons further. Only neurons showing L5 PN-typical I_h-currents were considered. Theta-barrel glass-pipettes with openings of $\tilde{3}$0 µm were placed at 50 µm – 80 µm from the patched neuron

to stimulate afferents targeting basal dendrites as well as in cortical layer 1 to stimulate apical afferents with synapses onto apical tuft dendrites.

After stabilization of resting membrane potential, apical and basal stimulation strength was adjusted to elicit EPSPs with amplitudes in the range of 2-3 mV. This usually required 100 µs voltage pulses in the range of -1 to -3 V followed by immediate 100 µs voltage pulses of +1 to +3 V for basal stimulations. Apical stimulation had the same shape but ranged from -6 to -8 V and +6 to +8 V, respectively. Next, the required strength of basal stimulations to induce a single action potential was determined. For this, tripling basal stimulation amplitudes and stimulating 3 times at 100 Hz would sometimes elicited a single action potential. Due to focal stimulation (theta glass pipettes with small opening), extracellular electrical stimulation was often not sufficient to trigger action potentials. In these cases, additional brief current injections into the soma via the patch pipette were used to trigger an AP. Finally, apical stimulation strength was increased until multiple action potentials or plateau potentials were induced.

After all required stimulation amplitudes were set, baseline EPSP amplitudes were recorded for 5 minutes with alternating basal and apical stimulations every 5 seconds followed by 8 combined, strong basal and apical stimulations with a 10 second interval, each eliciting a plasticity induction event. Next, EPSPs were recorded for at least 30 minutes (same as baseline recordings).

Analysis of electrophysiology traces was performed using individual Python scripts and the pyabf library. The correlation analyses were performed with the seaborn library (regplot) and the statistical analysis with the scipy library (spearmanr).

Analysis of PSST-AUC with different numbers of action potentials was performed on a subset of neurons in which PSSTs occurred with different amounts of APs (1vs2, 1vs3, or 2vs3). PSST-AUCs were averaged for events with min. and max. number of APs respectively and plotted as 2 data points per cell.

In the case of cell morphology reconstruction (only performed in 8 of 12 neurons), slices were kept in 4% PFA for 24-32 hours and stained overnight with Streptavidin, Alexa Fluor 350 conjugate (invitrogen) and mounted on commercially available microscopy slides.

Calcium imaging was performed in a subset of neurons (n = 6 neurons, 4 mice) after post-induction EPSP recordings where completed by imaging pipette-filled OGB5-N using a Lumencor SPECTRA Light Engine (GFP/FITC 475/28 nm) for excitation and a ORCA-Lightning Digital CMOS camera (C14120-20P, Hamamatsu) for image acquisition at 50 Hz. Imaging data was pre-processed (background-subtraction, bleaching correction) using Fiji/ImageJ2. Extracted fluorescence traces were further processed and analysed using individual Python scripts. A correlation analysis between somatic calcium recordings (derived as ΔF/F AUC) and somatic membrane potential AUC showed a significant correlation (**Supp. Fig. S4**, $R_{Spearman} = 0.8545,\ p = 0.0016$) and was used to infer calcium levels during the actual plasticity induction events using linear regression.

Group data are expressed as mean ± s.e.m. unless otherwise stated. Statistical significance was calculated using paired or unpaired t-tests as well as Spearman correlation.

Supplementary Text

1 Formalism for TL and BP

In this section we describe the underlying mathematics of both TL and BP in general terms. Note that practical implementations of either make simplifications or approximations that are useful in practice, which we will not develop in detail.

A learning algorithm for a neuronal network is an optimization procedure that decreases a loss function

$$\mathcal{L}(\mathbf{w}) = \sum_s l(y_s, f(x_s, \mathbf{w})) \tag{17}$$

were $\mathbf{w}$ are the weights of the connections between neurons, y_s is a desired output or activity of the network, and $f(\cdot,\cdot)$ is the function implemented by the network parameterized by the weights and the sensory input, and $l(\cdot,\cdot)$ is an error function.

Note that both the error function l and the loss $\mathcal{L}$ are mathematical constructs, which can always be defined for any learning process, but do not need to be implemented by any observable biological quantity. The formalism is still useful to make our analysis understandable and to connect our results to the machine learning literature as well as computational neuroscience frameworks such as predictive coding or energy models.

1.1 Formalism for TL

In TL the loss is minimized by imposing a different activity through an input to each neuron, so that

$$l(y_s, f(x_s, \mathbf{w}, \mathbf{c}^*)) = 0 \tag{18}$$

where $\mathbf{c}^*$ is a vector encoding the signal that imposes the activity onto each neuron. Depending on the architecture, this signal is called a control signal, a top-down prediction or a feedback input, and is computed in various ways (Rao and Ballard, 1999; Meulemans et al., 2022a,b; Song et al., 2024; Scellier and Bengio, 2017). In general, the control is chosen so as to provide the smallest input that will fulfill the aforementioned condition. Energy models, for example, follow the direction of steepest descent of the loss (Rao and Ballard, 1999; Scellier and Bengio, 2017), although other solutions such as control theory approaches are also used (Meulemans et al., 2021). In the case of unsupervised learning, y_s is simply the activity of the network that is closest to the one imposed by x_s but fulfills a constraint on the activity such as sparsity.

The main point of TL is to provide a plasticity such that the control input will not be required, and thus the network will have learned the right computation without needing this control.

Now a subtle step is to connect the synaptic weight update to the reduction of the control signal $\mathbf{c}$. Since **Eq. 18** is an invariant, the change of loss is zero and so is the input, hence

$$l(y_s, f(x_s, \mathbf{w}{+}\Delta\mathbf{w}, \mathbf{c}^*{+}\Delta\mathbf{c})) = 0 \Rightarrow \frac{\partial l(y_s, f(x_s, \mathbf{w}, \mathbf{c}))}{\partial f(x_s, \mathbf{w}, \mathbf{c})} \left[\langle \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial \mathbf{c}}, \Delta\mathbf{c} \rangle + \langle \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial \mathbf{w}}, \Delta\mathbf{w} \rangle \right] = 0 \tag{19}$$

where $\langle \cdot,\cdot \rangle$ is a dot product, and the derivatives are taken at for a control signal $\mathbf{c} \in [\mathbf{c}^* + \Delta\mathbf{c}]$ that comes from the mean value theorem. Assuming that the change in control from one iteration to the next will be small (akin to the assumption in BP that the weight updates impose a small change), $\mathbf{c} \approx \mathbf{c}^*$. From this, we can obtain the constraint

$$\langle \frac{\partial l(y_s, f(x_s, \mathbf{w}, \mathbf{c}))}{\partial f(x_s, \mathbf{w}, \mathbf{c})} \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial \mathbf{c}}, \Delta\mathbf{c} \rangle = -\langle \frac{\partial l(y_s, f(x_s, \mathbf{w}, \mathbf{c}))}{\partial f(x_s, \mathbf{w}, \mathbf{c})} \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial \mathbf{w}}, \Delta\mathbf{w} \rangle \tag{20}$$

where we can already see how the weight update relates to the control signal, which we will now make more precise and in terms that are similar to the standard machine learning formalism.

The first step is to recognize that both terms in **Eq. 20** are single scalars, but are computed as dot products of high dimensional quantities. Thus, the constraint is on one degree of freedom, but we want to turn the constraint into an optimization process that applies to many neurons and weights. For this we select the weight update that reduces the control signal in the most efficient way with the minimum weight change.

Starting with the fastest reduction of the control signal. If we want to pick $\Delta\mathbf{c}$ that reduces $\mathbf{c}$ the fastest, we simply set

$$\Delta\mathbf{c} = -\eta\mathbf{c} \tag{21}$$

where the proportionality constant corresponds to the learning rate. Thus, we can rephrase **Eq. 20** for single neurons,

$$\eta\frac{\partial l(y_s, f(x_s, \mathbf{w}, \mathbf{c}))}{\partial f(x_s, \mathbf{w}, \mathbf{c})}\frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial c_{\text{post}}}c_{\text{post}} = \langle\frac{\partial l(y_s, f(x_s, \mathbf{w}, \mathbf{c}))}{\partial f(x_s, \mathbf{w}, \mathbf{c})}\frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial \mathbf{w}_{\text{pre}}}, \Delta\mathbf{w}_{\text{pre}}\rangle, \tag{22}$$

and apply the chain rule to the partial derivatives and make them pass through the postsynaptic activity of the neuron,

$$\frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial c_{\text{post}}} = \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial a_{\text{post}}}\frac{\partial a_{\text{post}}}{\partial c_{\text{post}}}, \quad \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial \mathbf{w}_{\text{pre}}} = \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial a_{\text{post}}}\frac{\partial a_{\text{post}}}{\partial \mathbf{w}_{\text{pre}}}. \tag{23}$$

Since there is a derivative that relates the postsynaptic activity to the loss (a scalar to a scalar) on both sides of **Eq. 22**, we can simplify to

$$\eta\frac{\partial a_{\text{post}}}{\partial c_{\text{post}}}c_{\text{post}} = \langle\frac{\partial a_{\text{post}}}{\partial \mathbf{w}_{\text{pre}}}, \Delta\mathbf{w}_{\text{pre}}\rangle. \tag{24}$$

Most TL implementations use a control signal is additive (or computed as additive) to the hidden state of the neuron, the feed forward input in standard ANNs. Technically, other choices would work, such as using an additive control to the output, or a multiplicative effect, but to keep our derivations simple we will assume that the additive to the internal state is maintained, and note that the changes would be minimal.

In this line of thought, we introduce an internal neuron variable that conveys the weighted sum of the neuron inputs

$$v_{\text{post}} = \langle\mathbf{a}_{\text{pre}}, \mathbf{w}_{\text{pre}}\rangle \tag{25}$$

which gives the firing rate of the neuron through the relationship $a_{\text{post}} = \phi(v_{\text{post}})$, which we can then use to simplify **Eq. 24** to

$$\frac{\partial\phi(v_{\text{post}})}{\partial v_{\text{post}}}\frac{\partial v_{\text{post}}}{\partial c_{\text{post}}}c_{\text{post}} = \frac{\partial\phi(v_{\text{post}})}{\partial v_{\text{post}}}\langle\frac{\partial v_{\text{post}}}{\partial \mathbf{w}_{\text{pre}}}, \Delta\mathbf{w}_{\text{pre}}\rangle \Rightarrow \ \eta c_{\text{post}} = \langle\frac{\partial v_{\text{post}}}{\partial \mathbf{w}_{\text{pre}}}, \Delta\mathbf{w}_{\text{pre}}\rangle. \tag{26}$$

where

$$\frac{\partial v_{\text{post}}}{\partial \mathbf{w}_{\text{pre}}} = \mathbf{a}^{\star}_{\text{pre}} \Rightarrow \eta c_{\text{post}} = \langle\mathbf{a}^{\star}_{\text{pre}}, \Delta\mathbf{w}_{\text{pre}}\rangle. \tag{27}$$

where $\mathbf{a}^{\star}_{\text{pre}}$ is the target presynaptic activity. Note that a general assumption is that learning has to be done in the full network simultaneously, most TL architectures implicitly use a presynaptic activity that is controlled to be in the target state when learning happens (Rao and Ballard, 1999; Meulemans et al., 2022b; Song et al., 2024). To clarify the contrast with BP, we will denote $a^{\star}$ the neuronal activity in the controlled state.

Just as with the control signal, there are many possible changes of the presynaptic weight that would reduce the amount of control signal for the simple reason that there are usually many presynaptic neurons. Thus, we assume that the learning weight should be efficient, in the sense that the learning should provide the maximum decrease of c_{post} for the minimum amount of weight changes $\Delta\mathbf{w}_{\text{pre}}$. Geometrically, the natural way to maximize the effects of a presynaptic weight update is to align the update with the presynaptic inputs, thus setting the angle of the vectors in the dot product to zero.

Taken together, the update rule for a single synapse ends up being

$$\Delta w = \eta a^{\star}_{\text{pre}} c_{\text{post}} \tag{28}$$

which we can easily recast in the language of machine learning, $c_{\text{post}} = \delta^{TL}_{\text{post}}$,

$$\Delta w \propto a^{\star}_{\text{pre}}\delta^{TL}_{\text{post}}. \tag{29}$$

A crucial component of TL is that $a^{\text{target}}_{\text{pre}}$ is taken when the network is controlled, thus the presynaptic activity will be at the target (when the network has already learned).

1.2 Formalism for BP

In classical BP, the change of weights is computed by gradient descent, where the derivatives are computed over the neuron activity. This last technicality is crucial, because this makes BP computationally efficient. Explicitly, this is framed as

$$\Delta \mathbf{w}_{BP} = -\eta \frac{\partial \mathcal{L}(\mathbf{w})}{\partial \mathbf{w}} \tag{30}$$

where η is the learning rate. The change for a given weight for one given sample is then expressed as

$$\Delta_{BP} w = -\eta \frac{\partial a_{\text{post}}}{\partial w} \frac{\partial l(y_s, f(x_s, \mathbf{w}))}{\partial a_{\text{post}}} \propto -a^b_{\text{pre}} \frac{\partial l(y_s, f(x_s, \mathbf{w}))}{\partial a_{\text{post}}} \tag{31}$$

where the notation is commonly simplified as $\delta_{BP} = -\frac{\partial l(y_s, f(x_s, \mathbf{w}))}{\partial a^b_{\text{post}}}$, giving us

$$\Delta_{BP} w \propto a_{\text{pre}} \delta^{BP}_{\text{post}} \tag{32}$$

where a to refers to the activities during inference.

2 Links between BP and TL

The two families of algorithms that we study are not separated classes, but rather two extremes of a spectrum. While both algorithms adapt the synaptic weights to modify the activity so that a desired output is provided, BP uses the neuronal responses of the network to a stimulus, while TL disregards the current responses and uses instead the desired neuronal responses. This difference can emerge from the effect of the teaching signal on the neural activity or not (Meulemans et al., 2022b), and indeed various architectures originally designed for TL can be turned into BP in such way (Scellier and Bengio, 2017; Whittington and Bogacz, 2017; Rao and Ballard, 1999).

We make this relationship more precise by showing the possibility to use the machinery of BP to implement TL and vice versa:

- We can use BP to implement TL by using the BP gradients as a control signal (Ernoult et al., 2022). However, this requires adding a parameter to specify the strength of the gradient signal, or multiple iterations of the gradient. In either case, the parameter or iterated gradient needs be re-computed through the learning process, as the nonlinearities in a neural network would imply that the error does not scale linearly with the control signal.
- We can modify TL to approximate BP, by creating ad-hoc targets that are very close to the current activity (LeCun, 1986; Whittington and Bogacz, 2017). However, this requires creating a target for each stimulus presentation, rather than a global target that is maintained through learning (which is how TL is defined).

We can now formalize this statements. The gist of our argument relies on the fact that the transpose of the Jacobian of a system (which is computed in BP) and the inverse of that Jacobian (which provides optimal control) are closely related.

2.1 BP with specific strong learning signals is a close approximation of TL

Consider a neural network with a random initialization (as is often done in both deep neural networks and computational models). Through this initialization, we can make the following assumptions:

- All entries of the Jacobian matrix are sampled from a random distribution
- The mean of that random distribution is zero (the network is balanced), and the standard deviation σ_J is finite.
- The network is large, and contains many more neurons than outputs.

Our goal is to show that the first order control matrix Q and the transpose of the Jacobian are very similar. Starting with the control signal, a control matrix in a linearized system would guarantee that sending every neuron a control signal given by

$$Q(\mathbf{y} - \hat{\mathbf{y}}) \tag{33}$$

where $\mathbf{y} - \hat{\mathbf{y}}$ are the desired and actual system output respectively. A well conditioned control matrix for that network has the property that

$$JQ \approx I \tag{34}$$

where I is the identity matrix (Meulemans et al., 2021), where the rank of I is small compared to the number of neurons. Using the transpose of the Jacobian with a scaling as Q, we get that the control signal will work if

$$sJJ^\top \approx I. \tag{35}$$

Consider first the product of the Jacobian with its transpose. We can consider the value of entries in the diagonal and off diagonal of the product $P = sJJ^\top$

$$\mathbf{E}_i \left[P_{ii}\right] = \mathbf{E}_i \left[s \sum_k J_{ik} J_{ik}\right] = ns\mathbf{E}_{i,k} \left[J_{ik}^2\right] \tag{36}$$

$$\mathbf{E}_{i\neq j} \left[P_{ij}\right] = \mathbf{E}_{i\neq j} \left[s \sum_k J_{ik} J_{jk}\right] = ns\mathbf{E}_{i,j,k,l} \left[J_{ij} J_{kl}\right] \tag{37}$$

and by using the standard rules to compute the variance and expectation of the product of independent variables,

$$\mathbf{E}_i \left[P_{ii}\right] = sn\sigma_J^2 \tag{38}$$

$$\mathbf{E}_{i\neq j} \left[P_{ij}\right] = sn\mu_J^2 = 0 \tag{39}$$

$$\mathbf{Var}_i \left[P_{ii}\right] = sn\sigma_J^2 \tag{40}$$

$$\mathbf{Var}_{i\neq j} \left[P_{ij}\right] = sn\sigma_J^2 \tag{41}$$

Thus, the matrix P is effectively equivalent to

$$P = s\left(n\sigma_J^2 I + \sqrt{n}\sigma_J \mathcal{N}(0, I^n)\right) \tag{42}$$

where $\mathcal{N}(0, I^n)$ is a random matrix with gaussian entries. Since $n\sigma_J^2 \gg \sqrt{n}\sigma_J$, the random component will be small, and we can ensure that $P \approx I$, by setting

$$s = \left(n\sigma_J\right)^{-1}. \tag{43}$$

For example, if we set the Jacobian initialization ensuring that $\sigma_J = n^{-1/2}$, then $s = 1$.

Thus, for $s = \left(n\sigma_J\right)^{-1}$ the gradient signal can be used as a control signal. Note, however, that this requires a specific value of s. In other words, for BP to be similar to TL, the two following properties have to be fulfilled:

- The error signal acts as a control signal to the firing of the neuron.
- The error signal is scaled as $s = \left(n\sigma_J\right)^{-1}$.

Thus, BP can be interpreted as TL under those assumptions. However, there are two key differences to keep in mind:

- TL is equivalent to BP only if the parameter s (the learning signal strength) is set to a precise value, while in TL the value of s emerges naturally from first principles. Thus, for such observation TL is more parsimonious and thus the preferred theory.
- The previous assumptions work under linearized dynamics. However, when we have strong nonlinearities the value of s would need to be fine-tuned.

2.2 TL with small nudging targets is an approximation of BP

Having shown that BP with strong feedback is close to TL, we now prove the complementary direction and argue that TL with weak feedback is close to BP. For simplicity and to align with our implementation of TL, we will use the nomenclature of control theory, although a similar argument has been build for other specific implementations such as predictive coding (Whittington and Bogacz, 2017) or Equilibrium Propagation (Scellier and Bengio, 2017).

In TL, the learning signal δ_{TL} from **Eq. 29** is computed as the input necessary to enforce a target activity at the output $\hat{\mathbf{y}}$. Using a Taylor expansion approximation,

$$\delta_{TL} = J^{\dagger}(\mathbf{y} - \hat{\mathbf{y}}) + O\left((\mathbf{y} - \hat{\mathbf{y}})^{\top} H(\mathbf{y} - \hat{\mathbf{y}})\right) \tag{44}$$

where $J^{\dagger}$ is a pseudo-inverse of the Jacobian, which accounts for second-order interactions between neurons and H a matrix capturing nonlinearities and second-order interactions in the network. Matrix inverses and second order interactions are notably hard to compute, hence to consider the limit $\mathbf{y} - \hat{\mathbf{y}} \to 0$ is often more practical. Then the second order components are ignored and

$$\delta_{TL} = J^{\dagger}(\mathbf{y} - \hat{\mathbf{y}}). \tag{45}$$

Furthermore, in because of the small error we can simply consider the effect of a small perturbation in the neurons and compute how they would affect the output. This yields the simple computation

$$J^{\dagger} = \left(\frac{\partial \mathbf{r}_L}{\partial \mathbf{r}}\right)^{\top}. \tag{46}$$

Hence, if we set up a target that is very close to the current output,

$$\delta_{TL} \approx J^{\top}(\mathbf{y} - \hat{\mathbf{y}}), \tag{47}$$

which is the learning signal given by BP.

The natural question is then whether targets that are always closer to the current output are feasible to have. A natural approach is to define intermediate small targets, or "weak nudging" (Meulemans et al., 2021), where the control signal continuously adapt the target to be always close to the current output, but in the direction of the desired output activity.

Using such architectures we can obtain a valid approximation of BP through TL, but then we need to add an extra layer of complexity: on top of the target defined by standard TL, we need to add a circuit that computes a small target that can then be used to compute the learning signal. Thus, considering BP as an approximation of TL is less parsimonious than simply considering BP.

3 Derivation of the target learning rule from our PN model

To render our mechanistic PN comparable to the average activities observed in neuronal recordings in vivo, we reformulate our PN model as rate-based with a three-factor learning rule.

The inputs that arrive at the soma and the inputs that arrive at the apical dendrites are assumed to follow a Poisson distribution with a relatively low firing rate, in order to ensure that consecutive presynaptic spikes or feedback spikes are temporally separated with high probability, allowing us to consider spikes and bursts as independent events.

We divide the somatic output into bursts and single spikes. The number of events where somatic presynaptic input drove the soma over its firing threshold is represented by N. The events will sometimes generate a burst N_{burst} and sometimes a single spike N_{spike}, thus $N = N_{\text{burst}} + N_{\text{spike}}$. A burst is generated only if feedback input arrives at the apical dendrites within a short time interval ΔT after a somatic postsynaptic spike. The aforementioned assumption of low firing rates for apical and somatic inputs, combined with poisson events, implies that somatic and apical spikes rarely arrive very close in time, and thus two events do not lead to the same burst. As a consequence, the probability of turing a single spike into a burst grows linearly with the apical input rate. We can thus calculate the expected number of bursts and single APs as,

$$N_{\text{burst}} = \frac{\Delta T}{T}\alpha N, \quad N_{\text{spike}} = (1 - \frac{\Delta T}{T}\alpha)N \tag{48}$$

where T represents the total time of the simulation, and α is the rate of apical input. The number of bursts is simply the fraction of time at which the apical input can generate a burst times the probability of receiving suprathreshold somatic activity.

Lastly, the firing rate of the postsynaptic neuron increases when burst firing occurs. Therefore, considering the activity as the number of spikes, the activity can be expressed as,

$$a_{\text{post}} = N_{\text{spike}} + \overline{m} N_{\text{burst}} = N\left(\overline{m} + \frac{\Delta T}{T}\alpha(1-\overline{m})\right). \tag{49}$$

where $\overline{m} > 1$ is the expected number of spikes per burst. As denoted here, the expected firing rate of the neuron grows monotonically with the amount of apical feedback.

We can now derive plasticity from the calcium model, understanding that a plateau potential generates LTP, and a burst and a single spike without the corresponding apical input generates LTD when there is a presynaptic spike

$$\Delta w = [\Delta_{\text{burst}} w N_{\text{burst}} - \Delta_{\text{spike}} w N_{\text{spike}}]\,\frac{N_{\text{pre}}}{T} \tag{50}$$

where $\Delta_{\text{burst}} w > 0$ and $\Delta_{\text{spike}} w > 0$ denote the synaptic change per burst or spike, respectively, and N_{pre} correspond to the number of presynaptic spikes. Substituting **Eq. 48** into **Eq. 50** produces,

$$\Delta w = \big(\frac{\Delta T}{T} N \frac{N_{\text{pre}}}{T}\big)(\alpha(\Delta_{\text{burst}} w + \Delta_{\text{spike}} w) - \frac{T}{\Delta T}\Delta_{\text{spike}} w). \tag{51}$$

Now there exist a α^{baseline} such that $\Delta w = 0$, which corresponds to

$$\alpha^{\text{baseline}} = \frac{\frac{T}{\Delta T}\Delta_{\text{spike}} w}{\Delta_{\text{burst}} w + \Delta_{\text{spike}} w}. \tag{52}$$

Using α^{baseline}, we can write the synaptic weight change as

$$\Delta w = \eta \frac{N_{\text{pre}}}{T}\frac{N}{T}(\alpha - \alpha^{\text{baseline}}), \tag{53}$$

where $\eta = \Delta T(\Delta_{\text{burst}} w - \Delta_{\text{spike}} w)$. To simplify the notation further, we will denote the presynaptic firing rate as $r_{\text{pre}} = N_{\text{pre}}/T$ and the postsynaptic firing rate at baseline apical input $r_{\text{post}}^{\text{baseline}}$, which leaves us with the synaptic update rule

$$\Delta w = \eta a_{\text{pre}} a_{\text{post}}^{\text{baseline}}(\alpha - \alpha^{\text{baseline}}). \tag{54}$$

and if we set $\alpha^{\text{baseline}} = 1$ for simplicity, the postsynaptic firing rate for a given apical input is then

$$a_{\text{post}} = a_{\text{post}}^{\text{baseline}}\alpha. \tag{55}$$

We can then plug this postsynaptic activity into **Eq.** , and obtain

$$\Delta w = \eta a_{\text{pre}}(a_{\text{post}}^{\star} - a_{\text{post}}), \tag{56}$$

where $a_{\text{post}}^{\star}$ is the target postsynaptic activity which is driven by the apical input. In this formulation the error term matches the one used in existing architectures Rao and Ballard (1999); Whittington and Bogacz (2017); Meulemans et al. (2021), and the learning rule as we tested in **Eq. 2**.

4 Apical plasticity

In contrast to our results on basal synapses, we found that for apical synapses neither the plasticity induction AUC ($R_{Spearman} = 0.1049$, $p = 0.7456$, **Fig. S2D**) nor the number of APs ($R_{Spearman} = 0.3996$, $p = 0.1982$, **Fig. S2E**) are good predictors for changes in synaptic strength. This difference in the predictability of basal versus apical plasticity is also reflected in the finding, that overall changes in basal EPSP amplitudes did not correlate with changes in apical EPSPs ($R_{Spearman} = 0.2308$, $p = 0.4705$, **Fig. S2F**).

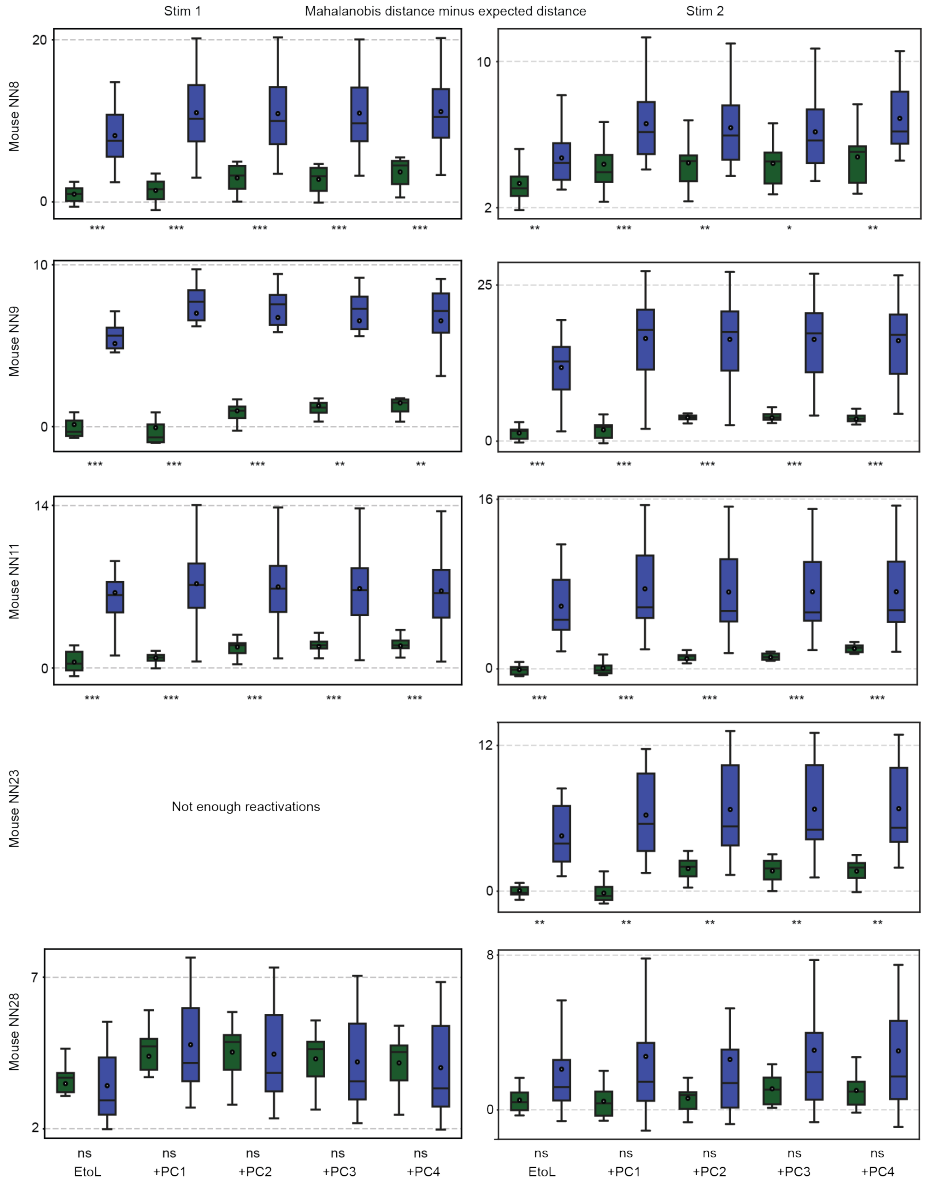

Figure S1: The Mahalanobis distances for the TL (green) and BP (blue) hypotheses for both stimuli and all mice. The statistical significance is denoted beneath the x-axis for each pair of distributions. For each sub-figure, every subsequent pair of distributions denote an additional dimension added to the component space.

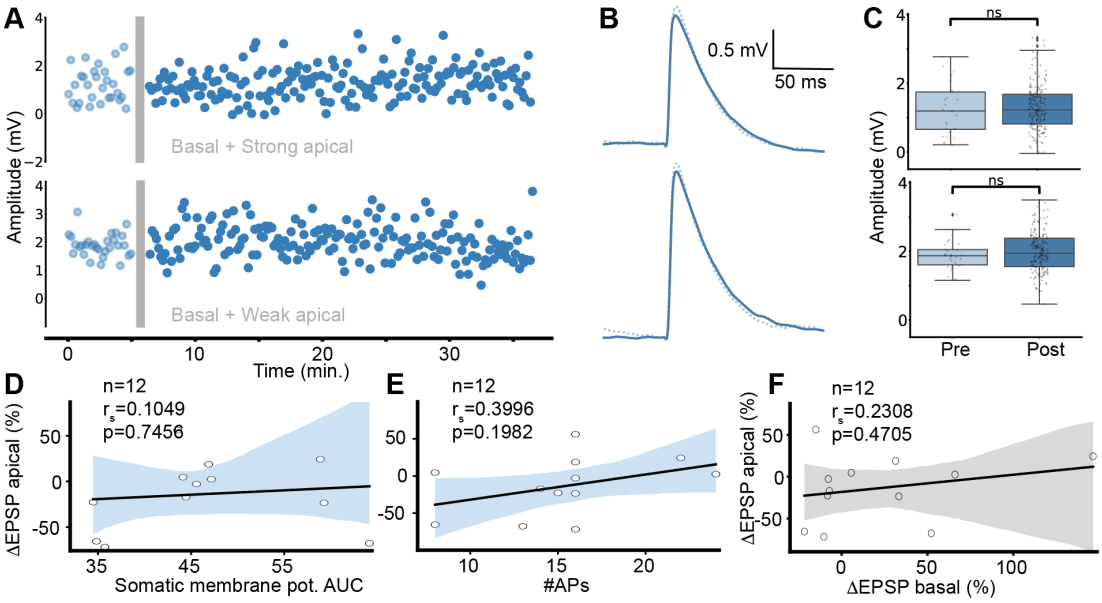

Figure S2: Representative example recordings of EPSP amplitudes during apical stimulation before and after the plasticity induction event (**A**) EPSP amplitude evolution over time; each dot represents one EPSP pre (light blue) and post (dark blue) plasticity induction event (gray bar). (**B**) The quantification of EPSPs for two sample cells shows no significant change in EPSP amplitudes. (**C, D**) Linear correlation analysis between changes in apical EPSP amplitude and somatic membrane potential AUC ((C) n=12, r_S=0.1049, p=0.7456) and number of APs ((D) n=12, r_S=0.3996, p=0.1982), respectively. (**E**) Correlation analysis of basal and apical ΔEPSP (n=12, r_S=0.3669, p=0.2408). (**F**) Correlation between basal and apical EPSP amplitude changes $r_s = 0.2308,\ p = 0.4705$.

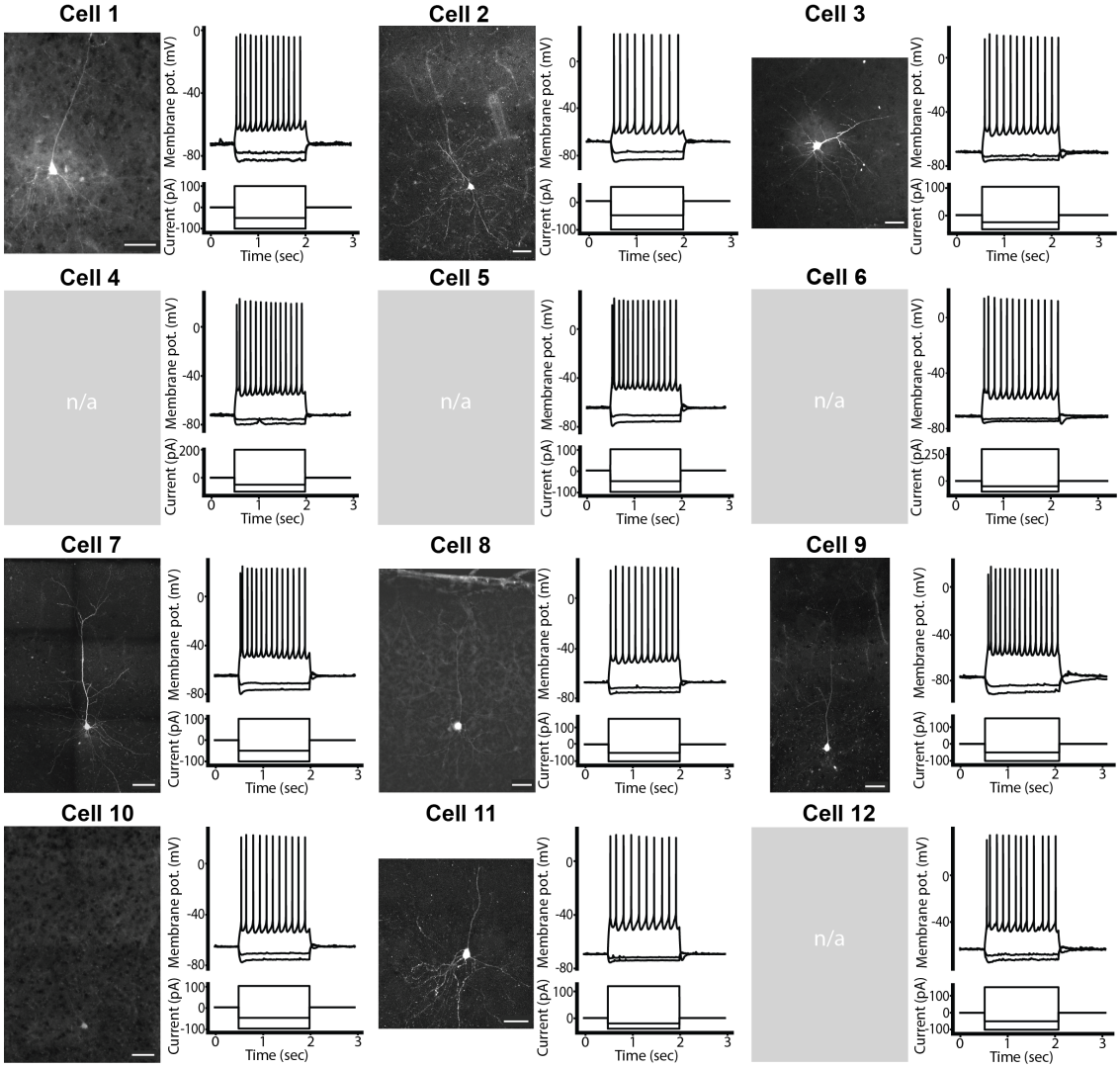

Figure S3: Biocytin-Streptavidin stained neurons (left) and typical hyperpolarizations and action potential trains (right top) induced by negative and positive current injection steps (right bottom) from recorded neurons. For cells 3, 4, 5, and 11, no Biocytin-Streptavidin staining/morphology reconstruction was performed.

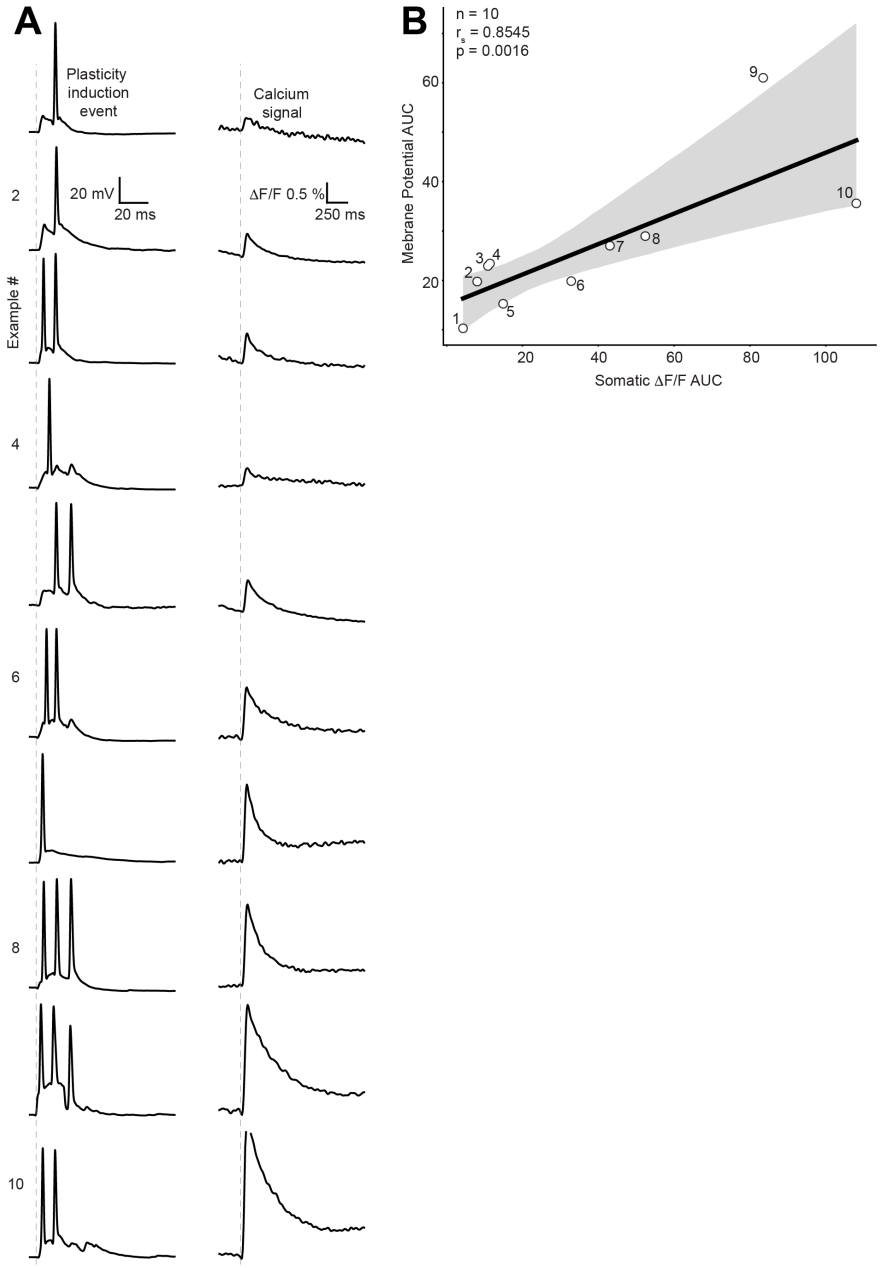

Figure S4: (**A**) Somatic membrane potential recording during plasticity induction (left) and somatic calcium ΔF/F during the same event (right). In total, showing events from 10 recordings in 6 neurons (4 mice) from **Fig. S3**. (**B**) Linear fit of induction event AUCs and somatic calcium signal AUCs showing significant correlation (n=6, 10 events total, R_S=0.8303, p=0.0029)