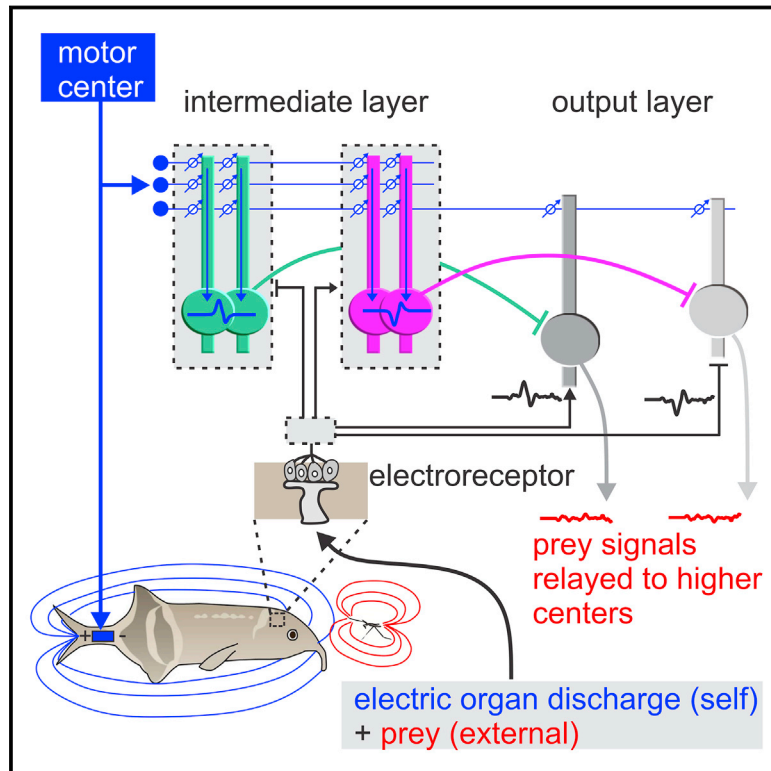


Continual Learning in a Multi-Layer Network of an Electric Fish

Graphical Abstract



Authors

Salomon Z. Muller, Abigail N. Zadina,
L.F. Abbott, Nathaniel B. Sawtell

Correspondence

ns2635@columbia.edu

In Brief

Using a cerebellum-like structure in an electric fish as a model system for investigating mechanisms of learning in multi-layer networks, Muller et al. observed that functional compartmentalization within individual neurons allows synaptic plasticity at an intermediate processing layer to adaptively shape network output.

Highlights

- Biological solutions to general problems in multi-layer learning are shown
- Intermediate layer function requires compartmentalization of learning and signaling
- Circuit organization based on learning solves the credit assignment problem



Continual Learning in a Multi-Layer Network of an Electric Fish

Salomon Z. Muller,^{1,2} Abigail N. Zadina,¹ L.F. Abbott,^{1,3} and Nathaniel B. Sawtell^{1,4,*}

¹Zuckerman Mind Brain Behavior Institute, Department of Neuroscience, Columbia University, New York, NY 10027, USA

²Department of Biological Sciences, Columbia University, New York, NY 10027, USA

³Department of Physiology and Cellular Biophysics, Columbia University, New York, NY 10027, USA

⁴Lead Contact

*Correspondence: ns2635@columbia.edu

<https://doi.org/10.1016/j.cell.2019.10.020>

SUMMARY

Distributing learning across multiple layers has proven extremely powerful in artificial neural networks. However, little is known about how multi-layer learning is implemented in the brain. Here, we provide an account of learning across multiple processing layers in the electrosensory lobe (ELL) of mormyrid fish and report how it solves problems well known from machine learning. Because the ELL operates and learns continuously, it must reconcile learning and signaling functions without switching its mode of operation. We show that this is accomplished through a functional compartmentalization within intermediate layer neurons in which inputs driving learning differentially affect dendritic and axonal spikes. We also find that connectivity based on learning rather than sensory response selectivity assures that plasticity at synapses onto intermediate-layer neurons is matched to the requirements of output neurons. The mechanisms we uncover have relevance to learning in the cerebellum, hippocampus, and cerebral cortex, as well as in artificial systems.

INTRODUCTION

Work on learning in neural systems has focused largely on the effects of plasticity at synapses that provide direct input to the neurons being studied (Bear and Malenka, 1994; Buonomano and Merzenich, 1998; Caporale and Dan, 2008; Knudsen, 1994). Learning a model of the environment or a complex skill, however, relies on plasticity that is widely distributed and may occur at synapses far from the neurons driving decisions or actions. As is well-known from multi-layer (or “deep”) artificial networks, distributing learning over multiple layers is substantially more powerful but also more difficult to implement than learning at a single layer (LeCun et al., 2015; Marblestone et al., 2016). In this study, we leverage a tractable system, the cerebellum-like circuitry of the mormyrid electrosensory lobe (ELL) (Bell et al., 2008), to identify neural mechanisms that implement multi-layer network learning.

Weakly electric mormyrid fish use passive electrosensing to detect electric fields in their environment, including the minute fields emitted by their prey. This task is made more difficult because mormyrids produce their own much larger electric fields by discharging an electric organ in their tail. Although vital for communication and active electrosensing, the fish’s electric organ discharge (EOD) induces large, long-lasting responses in passive electroreceptors that mask behaviorally relevant signals (Figure 1A; Bell and Russell, 1978). How do mormyrid fish distinguish self-generated from external sensory stimuli (Sperry, 1950; von Holst and Mittelstaedt, 1950)? Past studies suggest that this problem is solved in the ELL by the convergence of sensory input from electroreceptors on the body surface (Figures 1B and 1D, red) with corollary discharge signals triggered by the motor command nucleus that generates the EOD (Figures 1B and 1D, blue). Corollary discharge signals cancel self-generated sensory input from the EOD, allowing behaviorally relevant external signals to be detected and processed more effectively (Enikolopov et al., 2018). Cancellation is a continual dynamic learning process; the ELL can adapt to changes in the EOD signal within minutes and fine-tunes cancellation over a period of an hour or more (Bell, 1981, 1982).

To perform cancellation, the ELL must predict the sensory input produced by the EOD, subtract this prediction from the total sensory input, and transmit the difference (e.g., the prey signal) from its output layer (Figure 1B). The EOD cancelling signal, known as a negative image, is constructed from a temporally distributed set of corollary discharge inputs that are conveyed to the ELL through a granule cell-parallel fiber system (similar to that of cerebellum) and modified by anti-Hebbian synaptic plasticity (Bell et al., 1997b; Kennedy et al., 2014; Roberts and Bell, 2000).

Past studies have primarily treated sensory cancellation as cell autonomous, but ELL anatomy suggests a more complex, network-level process. The ELL is composed of two layers (Figures 1C and 1D) both of which receive electrosensory (red) and corollary discharge (blue) input. The first, the MG layer, consists of GABAergic neurons known as medium ganglion (MG) cells, and the second, the output layer, consists of output cells that project to higher processing stages (Bell et al., 1981). Output cells comprise two distinct classes known as E (excited) and I (inhibited) cells that respond with opposite polarity to sensory input. MG cells inhibit output cells and thus act as an intermediate or “hidden” processing layer. Both MG and output cells



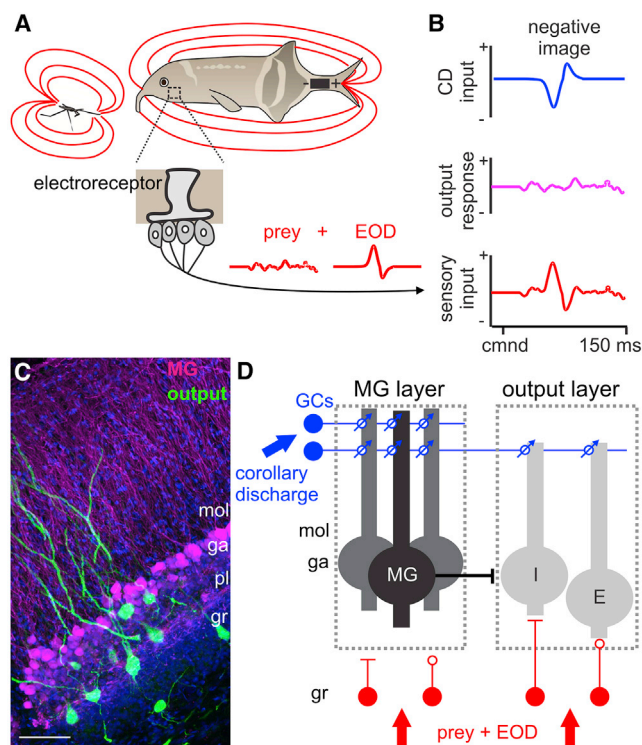


Figure 1. Function and Circuitry of the Mormyrid ELL

(A) Signals (red) conveyed by electroreceptor afferents contain both behaviorally relevant, prey-related and uninformative, EOD-related components that must be separated through learning in the ELL.

(B) Schematic of cancellation of the EOD-evoked component of the sensory input by a negative image constructed from corollary discharge (CD) inputs conveyed by granule cells (GCs; blue).

(C) Confocal tile scan of the ventrolateral zone (VLZ) of the ELL. Output cells were retrogradely labeled after a neurobiotin injection into the midbrain (green). MG cells were stained with a parvalbumin antibody (pseudo-color magenta). Scale bar, 100 μ m (mol, molecular layer; ga, ganglion layer; pl, plexiform layer; gr, granular layer).

(D) Schematic of ELL circuitry. Major inputs to the ELL are sensory input from electroreceptors on the skin (red) and CD inputs (blue) time-locked to the motor command (cmnd) that discharges the electric organ. Electrosensory input is relayed via excitatory and inhibitory interneurons, allowing some MG and output cells to be excited by sensory input (circles) and others to be inhibited (lines). CD inputs are relayed to the apical dendrites of MG and output cells via granule cells, the axons of which form the parallel fibers of the ELL molecular layer. Circles with arrows represent plastic synapses. Dashed boxes indicate the two distinct processing layers that are the focus of the present study.

receive the sensory and corollary discharge inputs needed for cancellation but, critically, ~90% of granule cell synapses are made onto MG cells (Bell et al., 2005; Meek et al., 1996). This suggests that the main site of plasticity in the ELL is at the intermediate MG layer.

Despite obvious differences between this biological system and artificial neural networks, processing, and learning at an intermediate layer raise a number of questions relevant to machine learning. The first concerns the nature of the processing that occurs at the intermediate layer of the ELL; what signals are the intermediate-layer MG cells conveying to the output layer and how do these sig-

nals contribute to ELL function? It has previously been proposed that MG cell output conveys EOD-subtracted information about relevant stimuli such as prey to the output cells (Mohr et al., 2003; Sugawara et al., 1999). Our results do not support this hypothesis. We argue, instead, that MG cells convey an estimate of the negative image output cells need for cancellation. This requires that MG cells simultaneously learn and transmit a negative image, raising a second basic question.

Learning and signaling can be conflicting requirements in networks because the activity needed to drive learning may not match the activity required for signaling. In machine learning, this problem is typically solved by cycling between separate phases of learning and signaling (Rumelhart et al., 1988). Such a solution may not be viable for brain structures like the ELL that must continuously transmit sensory information. In this study, we show that the ELL divides learning and signaling functions into separate neuronal compartments that operate simultaneously, obviating the need for temporal cycling.

The third question we address is related to the infamous credit assignment problem in machine learning: what assures that learning at intermediate layer synapses has a favorable impact on the performance of the output layer? We show that this issue is resolved in the ELL by an organization of synaptic connections between MG and output cells based on what MG cells learn through plasticity.

RESULTS

Examining MG Cell Output

In studying MG cell responses, it is important to note that MG cells fire two distinct types of action potentials: broad spikes and narrow spikes. Broad spikes are likely initiated in the apical dendrites, have a high threshold, are emitted at low spontaneous rates (~1 Hz), and drive plasticity. Narrow spikes are initiated in the axon, have a low threshold, are emitted at high rates (~50 Hz), and do not induce plasticity (Bell et al., 1997b; Han et al., 2000). The output of MG cells is dominated by narrow spikes because of their high rates. Importantly, MG cells do not receive a single powerful excitatory input that evokes broad spikes (like climbing fiber input to cerebellar Purkinje cells). Instead, *in vitro* studies indicate that both broad and narrow spikes can be evoked by granule cell and electrosensory inputs (Grant et al., 1998). The signals conveyed by broad and narrow spikes *in vivo* have not previously been examined in the region of the ELL involved in passive electrosensation.

We used both intracellular and extracellular recordings to measure MG broad and narrow spike responses as well as output cell responses in awake, paralyzed fish (see STAR Methods; Bell, 1982; Enikolopov et al., 2018). Under these conditions, the fish continue to generate EOD motor commands at 2–5 Hz, but the EOD itself is blocked by the paralytic. This allows responses to corollary discharge inputs to be studied in isolation from electrosensory input (Figures 2A–2C, blue). In addition, the electric field normally produced by the electric organ can be replaced, in the paralyzed condition, by a laboratory-generated mimic. Delivery of the mimic can be temporally uncorrelated with the recorded EOD command (Figures 2A–2C, red) producing “unpaired” sensory input. This condition allows the effects of

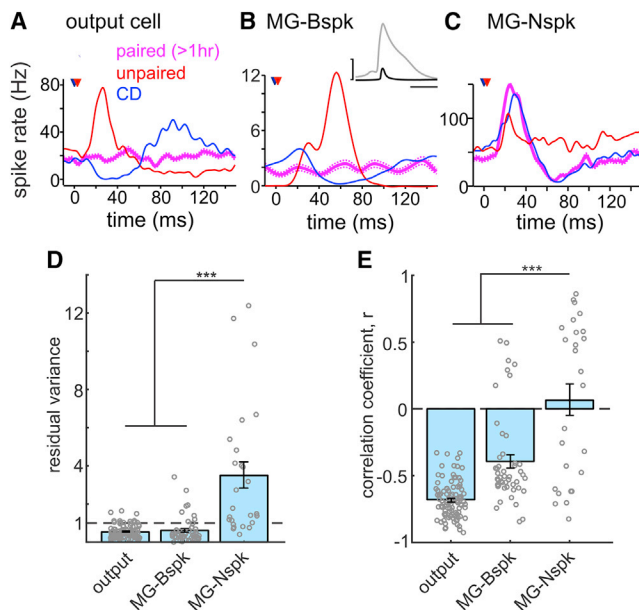


Figure 2. Cancellation of Output Cell and MG Cell Broad Spike, but Not MG Cell Narrow Spike, Responses

(A–C) Example output (A), MG broad spike (Bspk) (B), and MG narrow spike (Nspk) (C) firing rate responses after >1 h of paired sensory input (magenta), as well as responses to unpaired sensory input (red), and to CD input in the absence of sensory input (CD, blue). Dotted lines indicate SEM. Blue and red triangles indicate the times of the EOD command and of the sensory input, respectively. All recordings in this and subsequent figures were performed in the VLZ of the ELL which receives input from electroreceptors mediating passive electrosensation. Inset: example Bspk (gray) and Nspk (black) waveforms. Scale, 20 mV, 5 ms.

(D) Cancellation of paired sensory input (residual variance <1) was observed in output ($n = 95$) and Bspk ($n = 49$) but not Nspk ($n = 26$; $p < 0.0001$) responses. Error bars are SEM.

(E) Negative images, indicated by negative correlations between responses to CD and unpaired sensory input, were observed in output ($n = 95$) and Bspk ($n = 48$) but not Nspk responses ($n = 26$; $p < 0.001$).

See also Figures S1 and S2.

the sensory input to be measured in isolation from corollary discharge. Alternatively, the mimic can be delivered after a fixed (4.5 ms) delay that matches the normal interval between the motor command and the EOD (Figures 2A–2C, magenta). This condition, which produces “paired” sensory input, is used to induce learning of the negative image (the pairing-induced change in response to corollary discharge input alone) and to study the resulting sensory cancellation (the decrease in response to paired sensory and corollary discharge inputs).

In our studies, we refer to the EOD mimic as the sensory input. Because we do not include prey-like electric fields, the sensory input we discuss is entirely predictable on the basis of the EOD command signal and is therefore entirely uninformative to the fish. Thus, we consider a situation in which the ELL attempts to cancel all of its sensory input. It is important to appreciate that, in a natural setting, the mechanisms we analyze would only cancel the predictable self-generated component of the sensory input, leaving intact the unpredictable inputs of interest to the fish. Note that the ELL cannot construct the negative image directly

from sensory input because this would cancel the desired (prey) response along with the EOD response. Instead, sensory input is used as a teaching signal to construct a prediction of the EOD response based purely on motor corollary discharge.

Consistent with previous results (Bell, 1981, 1982), we found that output cell responses to paired sensory input exhibit cancellation (Figure 2A, magenta), and their responses to corollary discharge without any mimic resemble negative images (Figure 2A, blue) of their responses to unpaired sensory input (Figure 2A, red). We also found that MG cell broad spike responses to paired sensory input are canceled by negative images (Figure 2B) in a manner similar to the cancellation seen in output cells. In contrast, MG narrow spike responses remain strongly modulated by paired sensory input (Figure 2C), even after >1 h of pairing when most MG cell broad spike and output cell responses show full cancellation (Figures 2D and S1B–S1E). Consistent with this lack of cancellation, MG narrow spike responses to corollary discharge input alone do not resemble negative images of narrow spike sensory responses (Figures 2C and 2E). These results conflict with the idea that MG cell narrow spikes convey information about external signals (e.g., prey) to output cells because this would require cancellation. In addition, we found that MG cells have low sensitivity to prey-like sensory stimuli (Figure S2), prompting us to further investigate the nature of the signals carried by MG cell narrow spikes.

Sensory Input Affects Broad and Narrow Spikes Differently

The differences between cancelled broad spike and uncanceled narrow spike responses to paired sensory input could arise because they react differently to sensory input, to corollary discharge input, or to both. We began by examining sensory input. Interneurons convey both excitatory and inhibitory sensory input to the MG layer, as they do to the output layer. However, we found that the classification of MG cells into subtypes is more complex than for output cells. Surprisingly, broad and narrow spike responses to unpaired sensory input often have different polarities (Figures 3A–3C and S3). In fact, narrow and broad spike sensory responses are no more correlated in the same MG cell than they are in randomly chosen pairs of different cells (Figure 3C). For reasons that will become apparent, we chose to classify MG cells on the basis of their broad spike responses, defining BS⁺ (broad spikes excited by sensory input; Figure 3A) and BS[−] (broad spikes inhibited by sensory input; Figure 3B) subtypes.

In a subset of intracellular recordings from BS[−] MG cells, we also examined the effects of sensory input on broad spikes evoked by intracellular current injection. We found cells in which sensory input evoked membrane potential depolarization and increased narrow spike firing even though it dramatically reduced the probability of evoking a broad spike (Figure 3D). This provides further evidence that sensory input can affect broad and narrow spikes differently.

Granule to MG Cell Plasticity Affects Broad and Narrow Spikes Similarly

We next examined how learning-induced changes in the corollary discharge input conveyed by granule cells affect broad

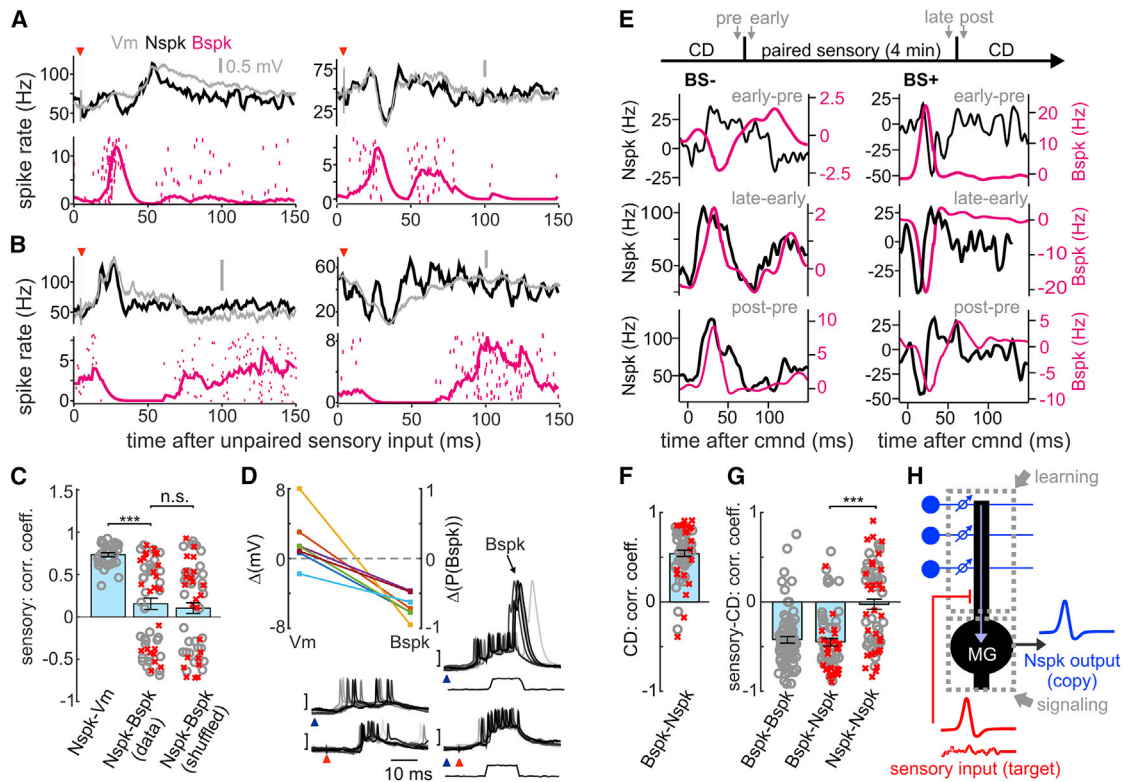


Figure 3. Effects of Sensory and CD Input on Bspk and Nspk in MG Cells

(A) Subthreshold membrane potential (Vm, gray), Nspk (black) and Bspk (magenta) rates, and Bspk raster for 2 example BS⁺ MG cells in response to unpaired sensory input (red triangles).

(B) Same as in (A) but for 2 example BS⁻ MG cells. BS⁺ cells are classified according to the polarity of the initial phase of their Bspk response.

(C) Left: MG cell Nspk and Vm sensory responses are correlated (Pearson $r = 0.72$, $n = 30$). Nspk and Bspk responses recorded in the same cells (middle) are no more correlated than pairs of responses selected randomly from different cells (right; $r = 0.12$ versus $r = 0.11$, $n = 37$, $p = 0.52$). Red crosses indicate Vm responses and gray circles indicate Nspk responses. Error bars are SEM.

(D) Overlaid traces from a BS⁻ cell showing responses to CD alone (blue triangle), unpaired sensory input (red triangle), and intracellular current injections (bottom traces). The sensory input is excitatory for Nspks (bottom panel, left column of traces), but it inhibits the generation of Bspks by current injection (compare two panels, right column of traces). Scale indicates range -50 to -60 mV. Inset (top left): effect of sensory input on the membrane potential (left points and axis) and on the probability of evoking a Bspk with current injection (right points and axis). Sensory input reduces the probability of evoking Bspks via current injection ($n = 7$ BS⁻ MG cells).

(E) Responses of BS⁻ (left) and BS⁺ (right) MG cells before, during, and after a 4 min delivery of paired sensory input.

(F) Pairing-induced changes in responses to the CD are positively correlated between Nspks and Bspks (circles; $r = 0.52$, $n = 38$) and between Vm and Bspks (crosses; $r = 0.56$, $n = 31$).

(G) Changes in Bspk CD responses are negatively correlated with Bspk sensory responses to the mimic (left, $r = -0.42$, $n = 73$). Changes in Nspk and Vm CD responses are also negatively correlated with Bspk sensory responses (middle; $r = -0.36$, $n = 38$ for Nspks and $r = -0.57$, $n = 28$ for Vm). However, changes in Nspk or Vm CD responses are uncorrelated, on average, with their sensory responses ($r = -0.003$, $n = 39$ for Nspk and $r = -0.06$, $n = 30$ for Vm; $p < 0.0001$).

(H) Schematic of a simplified two-compartment model of an MG cell. Sensory input selectively affects Bspks through local inhibition. Nspks transmit a learned copy (blue) of the predictable, self-generated component of the sensory input (red).

See also Figures S3 and S4.

and narrow spikes. A brief (4 min) pairing period was used so that we could compare responses in the same MG cell before, during, and after learning (Figure 3E). Plastic changes measured both during and after this period are highly correlated between narrow spikes and broad spikes (Figures 3F, S4A, and S4C), consistent with corollary discharge inputs affecting both responses similarly. Because plasticity is controlled by broad spikes, it constructs a negative image of the EOD sensory input affecting broad spikes. Narrow spikes are affected equally by this plasticity, so the broad-spike negative image is “inherited” by the narrow spikes (Figures 3G, S4B, and S4D). This inherited nega-

tive image does not cancel narrow spike responses (Figures 2D and S4G) due to the lack of correlation between the effects of sensory input on narrow and broad spikes (Figure 3C).

A Two-Layer, Two-Compartment Model of Learning in the ELL

Our results suggest that sensory input to an MG cell affects broad and narrow spikes differently. Past studies in zones of the ELL involved in active electrolocation have described a local inhibition of MG cell proximal apical dendrites that potently and selectively modulates broad spikes (Sawtell et al., 2007). We

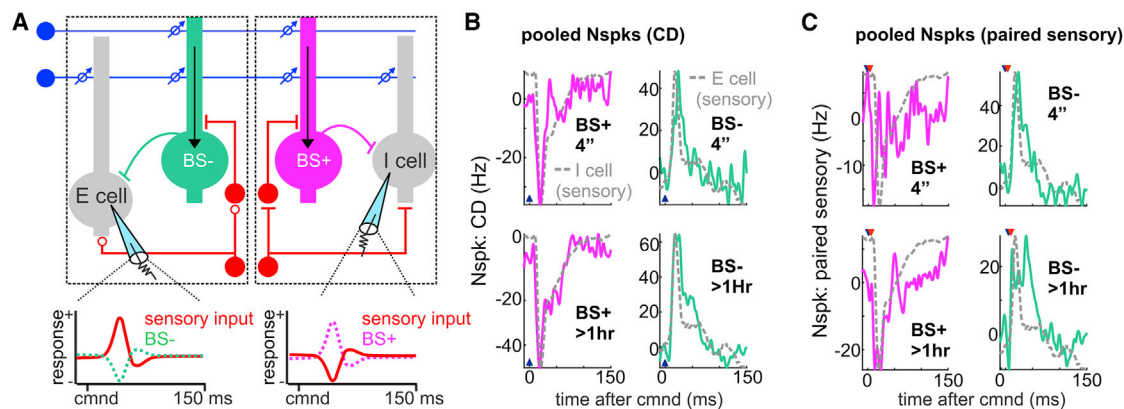


Figure 4. Grouping MG Cells According to Bspk Sensory Response Polarity Reveals Nspk Signals Appropriate for Output Cell Cancellation

(A) Hypothesized two-layer model of output cell cancellation. Negative images (dashed lines) appropriate for cancelling self-generated sensory input to output cells (solid lines) are transmitted from BS⁻ MG cells (green) to E output cells and from BS⁺ MG cells (magenta) to I output cells. Arrows indicate that MG cells transmit learned GC (as opposed to sensory) responses.

(B) Pooled change in Nspk CD responses induced by 4 min (top; BS⁺, $n = 15$; BS⁻, $n = 9$) or >1 h (bottom, BS⁺, $n = 5$; BS⁻, $n = 14$) of paired sensory input. Traces are averages across MG cells pooled according to the polarity of their Bspk sensory responses (BS⁺ or BS⁻). Dashed gray lines in (B) and (C) show the polarity and temporal profile of average sensory input recorded intracellularly in E and I output cells (arbitrary scale).

(C) Pooled change in paired Nspk responses (EOD mimic + command) after 4 min (top; BS⁺, $n = 13$; BS⁻, $n = 9$) or >1 h (bottom, BS⁺, $n = 5$; BS⁻, $n = 14$) of paired sensory input.

See also Figure S5.

therefore assume that there are two types of sensory input to MG cells, one, I_{both} , that affects both broad and narrow spikes (presumably due to synapses onto basal dendrites), and another, I_{BS} , that potentially affects broad spikes only (presumably due to synapses onto proximal apical dendrites). In addition, as shown above, broad and narrow spikes are affected equally by input, I_{gran} , from granule cells. Thus, the total broad spike input is $I_{\text{both}} + I_{\text{BS}} + I_{\text{gran}}$, and narrow spikes, being unaffected by I_{BS} , are driven by $I_{\text{both}} + I_{\text{gran}}$. We assume that I_{BS} is larger in magnitude than I_{both} .

Anti-Hebbian plasticity at granule cell synapses onto MG cells forces broad spikes rates to a constant value, such that after learning broad spikes are unmodulated by sensory input (Roberts and Bell, 2000). This implies that, after learning, the current driving broad spikes, $I_{\text{both}} + I_{\text{BS}} + I_{\text{gran}}$, is equal to a constant value C . As a result, the current driving narrow spikes, $I_{\text{both}} + I_{\text{gran}}$, is equal to $C - I_{\text{BS}}$. The first term in this expression generates tonic firing and the second modulates narrow spikes in a pattern that is a negative image of the broad-spike sensory input. Using a linear approximation for fluctuations in the narrow spike rate around its high baseline, the narrow spike modulation by learned granule cell input matches the shape and timing of EOD sensory input, making it a negative image that would be suitable for helping output cells cancel their sensory input.

Our analysis suggests that MG cells have two functional compartments: one that generates broad spikes and drives learning and the other that generates narrow spikes and transmits a negative image to other neurons (Figure 3H). This further suggests that the cancellation of self-generated sensory input in the ELL is a two-layer, target-based computation. The sensory input to broad spikes provides a target function, and broad spike mediated plasticity causes granule cell input to produce a nega-

tive image of this target that is transmitted to output cells by MG cell narrow spikes.

Because output cells come in E and I varieties, MG cells must generate and transmit two different types of negative images to the output layer. Taking into account that MG cells inhibit output cells, BS⁺ cells transmit negative images appropriate for output I cells and BS⁻ cell output is appropriate for output E cells (Figure 4A). Ten to twenty MG cells converge onto each output cell (Bell et al., 2005), which may allow imperfections in the negative images carried by individual MG cells to be averaged out. This two-compartment, two-layer model makes two key predictions: (1) summed narrow spike responses from BS⁻ MG cells must match the timing and shape of paired sensory input to output E cells, and narrow spike responses summed across BS⁺ MG cells should similarly match sensory input to output I cells; and (2) the system must be wired so that BS⁻ MG cells inhibit output E cells and BS⁺ MG cells inhibit output I cells. If true, these two predictions assure that learned MG narrow-spike-mediated inhibition is of the appropriate sign and shape to cancel unwanted sensory input to ELL output cells.

Physiological and Anatomical Evidence for the Two-Layer Model

To test the first model prediction, we summed narrow spike output from recorded MG cells, either BS⁻ and BS⁺, and compared the result to the subthreshold responses to unpaired sensory input in the appropriate type of output cell, either E or I. Learning-induced changes in narrow-spike responses pooled across BS⁻ cell have a polarity and temporal profile that matches the sensory response of E cells. Similarly, response changes pooled across BS⁺ cells match the sensory response of I cells (Figure 4B). This indicates that learning in each type of MG cell is appropriate for the requirements of the corresponding output cells. Importantly,

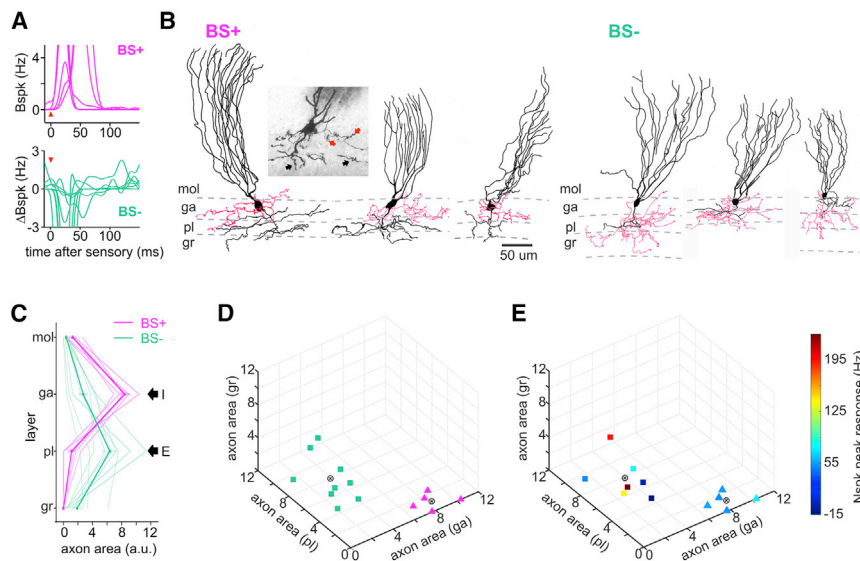


Figure 5. Laminar Distribution of MG Cell Axons Reflects Bspk Sensory Response Polarity and Is Appropriate for Sensory Cancellation in Output Cells

(A) Bspk sensory responses used to classify BS⁺ (n = 5) and BS[−] (n = 9) MG cells that were then morphologically reconstructed. Due to the low spontaneous firing rate of Bspks, reductions in current-evoked Bspk firing relative to baseline were used to classify some BS[−] cells. Bspk responses of some cells are truncated for clarity.

(B) Camera lucida reconstructions of physiologically identified BS⁺ (left, n = 5) and BS[−] cells (right, n = 9). Axon is drawn in red and ELL layer boundaries are indicated by dashed lines (mol, molecular layer; ga, ganglion layer; pl, plexiform layer; gr, granular layer). Inset: photomicrograph of a reconstructed BS⁺ cell. Red and black arrows indicate axonal and basal dendritic processes, respectively.

(C) Axon distribution across layers is different for BS⁺ and BS[−] cells (p < 0.0001, Skellings-Mack test). Arrows show laminar locations of the somata of E and I output cells.

(D and E) Unsupervised clustering of MG cells based on the laminar distribution of their axon areas by using k-means clustering with k = 2 (distance to centroid (circle with cross) was measured by squared Euclidean). The average silhouette value of the cluster denoted by triangles is 0.9038 and average silhouette value of the cluster denoted by squares is 0.6147. For visualization purposes we show values for 3 layers. (D) There is a complete overlap of cluster assignment with electrophysiologically defined Bspk response classes (BS⁺ in magenta and BS[−] in green). (E) In contrast, there is no clear relationship between cluster assignment and Nspk sensory responses.

See also Figure S6.

the narrow-spike outputs of BS[−] and BS⁺ cells to paired sensory input also matched negative images appropriate for E and I output cells (Figures 4C, S5A, and S5B). These results indicate that MG cells transmit appropriate cancellation signals to output cells provided that the connectivity we have assumed is correct. We therefore tested this essential prediction.

E and I output cells are located in different layers of the ELL, and MG cell synapses target their somata (Grant et al., 1996; Meek et al., 1996). These features allowed us to check connectivity by performing morphological reconstructions of physiologically identified BS⁺ and BS[−] MG cells recorded *in vivo* and filled with biocytin (Figures 5A, 5B, S6A, and S6B). Consistent with the two-layer model, we found that the axonal arbors of BS⁺ cells are largely restricted to the ganglion layer where output I cell somata are located, while the axonal arbors of BS[−] MG cells are mainly found below the ganglion layer where output E cell somata are located (Figure 5C). These results are consistent with past studies demonstrating two anatomical classes of MG cells (Han et al., 1999; Mohr et al., 2003) but are novel in showing that these classes correspond functionally to MG cells with different polarities of broad spike sensory input. Cells grouped by the polarity or strength of narrow spike sensory responses do not exhibit differences in axon location (Figures 5D, 5E, and S6C). Thus, as required by the two-layer model, ELL connectivity is organized on the basis of MG broad spike sensory responses or, by extension, on the basis of what the MG cells “learn” through broad spike evoked plasticity.

A Biophysical Model of MG Cell Function

Our model relies on an assumption that MG cell broad spikes are affected by an input (that we called I_{BS}) that has no appreciable

impact on narrow spikes. We constructed a biophysical model to see if such an input is plausible. Moreover, we wanted to test whether two types of sensory input can produce opposite responses, as we observe in the data, and whether granule cell input to the apical dendrite can propagate to affect narrow spikes despite the local dendritic inhibition that suppresses broad spikes. We built a conductance-based multi-compartment model with morphology taken from a *NeuroLucida* reconstruction of a biocytin-filled MG cell (see STAR Methods; Figure 6A). Voltage-gated sodium and potassium channels in axonal and apical dendritic compartments of the model cell generated dendritic and axonal spikes resembling broad and narrow spikes in MG cells (broad and narrow spikes are both blocked by tetrodotoxin and hence considered sodium-based) (Engelmann et al., 2008; Roberts and Leen, 2010). The densities of dendritic and axonal voltage-gated channels were adjusted to yield ~10 mV difference in threshold between broad and narrow spikes (Engelmann et al., 2008; Grant et al., 1998; Sugawara et al., 1999). Baseline excitatory and inhibitory synaptic inputs were adjusted to evoke broad and narrow spikes at rates consistent with those observed *in vivo*.

Sensory input was delivered to the model MG cell through excitatory inputs to basilar dendrites and inhibitory inputs to the proximal apical dendrites. Under these conditions, sensory input can robustly increase narrow spike firing while decreasing broad spike firing, consistent with experimental results from BS[−] cells (Figure 6B). An increase in broad spiking together with a decrease in narrow spikes, as in BS⁺ cells, can also be obtained if sensory input inhibits the basilar dendrites and diminishes inhibition in the proximal dendrites (Figure 6C). Thus, broad- and narrow-spike responses of opposite polarity are easily obtained.

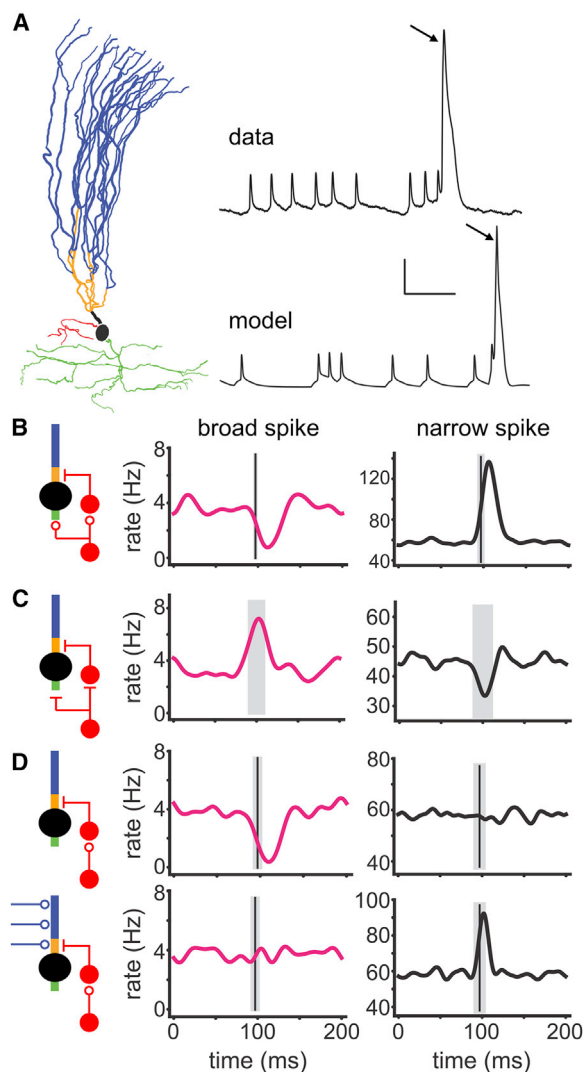


Figure 6. Biophysical Basis for Compartmentalization of Learning and Signaling in MG Cells

(A) Left: *NeuroLucida* reconstruction of the BS⁺ MG cell used to construct a multi-compartment model. Distal apical dendrites are shown in blue, proximal apical dendrites in orange, basal dendrites in green, and a partial reconstruction of the axon in red. Right: sample voltage traces from a recorded MG cell compared to the model. Arrows indicate Bspks. Scale, 10 mV, 20 ms. (B–D) Model Bspk (magenta) and Nspk (black) rates evoked by different combinations of excitatory (circles) and inhibitory (lines) input onto the compartments indicated in the schematics. The gray patch in (B) and (D) indicate mean (black line) and SD of inhibitory onset timing across proximal apical compartments. Gray patch in (C) indicates when tonic inhibitory input was removed. Model BS[−] cell with opposite Bspk and Nspk responses (B). Model BS⁺ cell with opposite Bspk and Nspk responses (C). Bspks, but not Nspks, are inhibited by local dendritic inhibition (I_{BS}) (D, top). GC input accompanied by dendritic inhibition cancels Bspk modulation but drives Nspk modulation, reflecting the negative image (D, bottom).

We next explored the effect of granule-cell input in the model. Granule-cell excitation was adjusted to cancel the inhibitory effect of sensory input on the broad spike rate, simulating the effects of plasticity in a BS[−] cell. Under these conditions, narrow

spikes are strongly modulated by granule cell excitation despite inhibitory input to the proximal dendrites, consistent with narrow spikes transmitting a negative image (Figure 6D). The selective effects of inhibition on broad spikes in the model are due, in part, to their higher threshold. Inhibition has a stronger effect at the higher broad spike threshold than at the lower narrow spike threshold due to its greater difference from the inhibitory reversal potential. These modeling results show that standard neuronal biophysics can support the compartmentalization of learning and signaling in MG cells.

Two Sites of Learning in the ELL

In our two-layer model, MG cells learn to extract a negative image from the large number of granule cell inputs they receive (~20,000) and transmit it to output cells. Granule cells also synapse directly onto output cells, although in smaller numbers (~5,000). *In vivo* studies suggest that these synapses are also plastic (Bell et al., 1997a), and although this plasticity has not been characterized *in vitro*, it appears to operate in a manner similar to the broad spike evoked plasticity in MG cells, except that it is driven by conventional spikes (output cells do not fire broad spikes). Thus, the negative images that cancel paired sensory input in output cells appear to come from two sources: MG cells and granule cells.

To isolate each of these contributions to the negative image in output cells, we performed manipulations that prevented plasticity from occurring either in the output cell being recorded (isolating the MG contribution to the negative image) or in the MG cells (isolating the direct granule cell contribution). These two conditions were compared with normal command-mimic pairings in the same cells (all pairing were 4 min in duration). In the first condition, intracellular current injections were used to counteract the effects of the paired sensory input on the output cell being recorded, thereby preventing anti-Hebbian plasticity from acting at its granule-cell synapses (Figure 7A, magenta). Because this manipulation only affected the output cell being recorded, plasticity acted normally in MG cells. For both E and I output cells, a substantial fraction of the negative image remained intact, suggesting that much of the negative image in output cells is transmitted from the MG layer, at least in the context of the brief pairings tested here (Figures 7B and 7C). In the second condition, we replaced the paired sensory input with a command-paired intracellular current injection into the recorded output cell (Figure 7B, blue). Because the MG cells received no sensory input in this case, plasticity was restricted to the recorded output cell (Bell et al., 1993, 1997a). This manipulation, in which the MG contribution to the negative image is absent, caused a significant reduction in the magnitude of the negative image in output cells compared to controls (Figures 7D and S5C). Both results indicate that a substantial fraction of the negative image in output cells is relayed from MG cells rather than being generated by direct granule cell input. This suggests that MG cells provide a “generic” negative image to output cells, while plasticity at granule cell to output cell synapses “fine-tunes” the negative image, cell by cell.

DISCUSSION

This study leverages a tractable model system, the mormyrid ELL, to address the general question of how synaptic

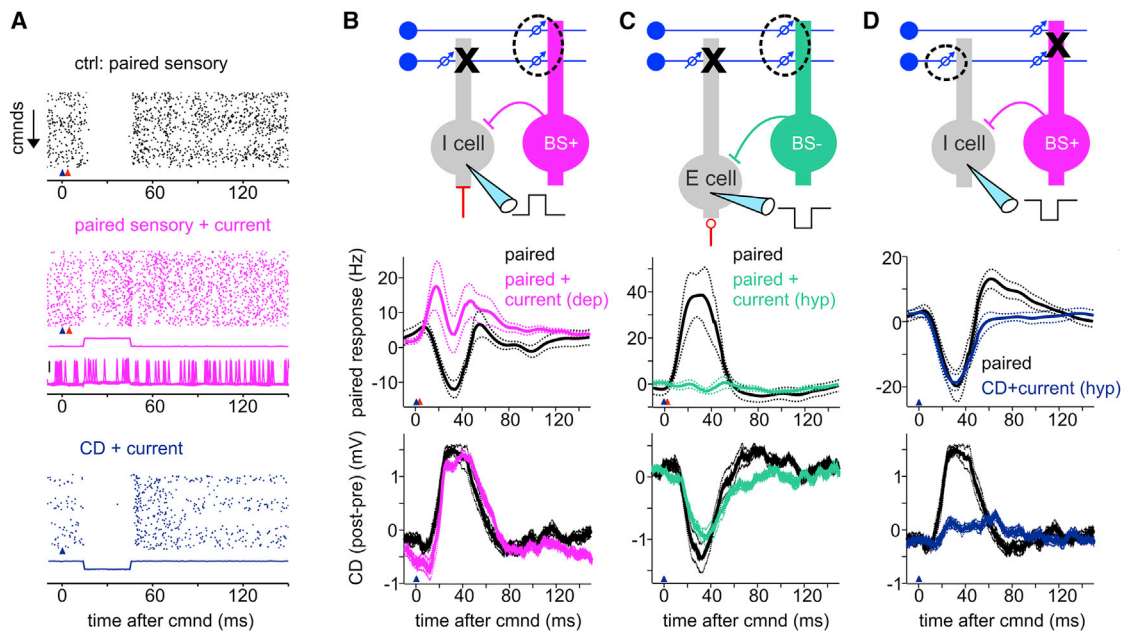


Figure 7. Isolating Single-Cell and Network Contributions to Negative Images in Output Cells

(A) Rasters from an I output cell illustrating responses during control (black) and experimental pairing conditions (magenta, blue). Timing of current injections are indicated by traces beneath the rasters. Bottom trace in the middle panel shows overlaid traces of the output cell membrane potential recorded during pairing. Scale bar, 25 mV.

(B) Average traces from output I cells ($n = 7$) showing responses during pairing (middle) and the resulting changes in CD responses (negative images; bottom). Black indicates control pairing and magenta indicates pairing in the same cell but with depolarizing current used to counteract effects of the paired sensory input. Dashed lines indicate SEM. The magnitude of negative images was not different in the two conditions ($p = 0.66$).

(C) Average traces from output E cells ($n = 14$) in which same manipulation and control as in (B) were performed (top). The magnitude of negative images was again not different in the two conditions ($p = 0.32$).

(D) Average traces from experiments ($n = 7$) in which control pairings of commands with an inhibitory sensory input were compared to pairings with hyperpolarizing current injections that evoked similar responses. Negative images were significantly larger under control conditions ($p < 0.001$). Pairings with depolarizing current were not performed because of the difficulty of matching current- and sensory-evoked spiking patterns.

plasticity at an intermediate layer of a network supports behaviorally relevant computation at the output layer. We show that intermediate layer plasticity sculpts corollary discharge responses that aid in the cancellation of unwanted self-generated sensory input at the critical output stage of the ELL network. Accomplishing this requires that intermediate layer MG cells solve two problems, both of which are broadly relevant to multi-layer learning in biological and artificial systems. First, we demonstrated that MG cells compartmentalize two functions: learning via dendritic broad spikes and transmitting cancellation signals via axonal narrow spikes. Without such separation, corollary discharge responses would cancel both narrow—as well as broad—spike responses, preventing MG cells from transmitting a cancellation signal to output cells. Second, we provide anatomical and electrophysiological evidence that opposite polarity signals learned by two distinct classes of MG cells are appropriately routed such that they contribute to cancellation in two opponent classes of output neurons. This implies that the ELL solves a version of the error credit assignment problem by organizing the connectivity between intermediate and output layer neurons on the basis of learning.

Mechanisms for Compartmentalization of Function and Credit Assignment in MG Cells

MG cells exhibit a striking separation of activity related to learning (broad spikes) and signaling (narrow spikes) despite lacking an obvious anatomical specialization for doing so, such as the climbing fiber input to Purkinje cells or an electrotonically remote site of dendritic spike initiation, as exists in cortical pyramidal neurons (Larkum et al., 1999; Schiller et al., 1997). Our compartmental model showed that the separation of learning and signaling observed in MG cells does not require highly specialized mechanisms or fine-tuning of biophysical parameters. In addition to the high sensitivity of broad spikes to inhibition due to the distance of their threshold from the inhibitory reversal potential, several other factors likely contribute to this separation, including inhibitory inputs spatially localized to the site of broad spike initiation and sensitivity of broad spike to suppression arising from the spatially and temporally distributed dynamics of broad (versus narrow) spike initiation (Engelmann et al., 2008). *In vitro* experiments are needed to reveal the relative importance and interplay between these factors.

We have argued that MG cell activity is driven by two classes of inputs, one that affects both broad and narrow spikes (I_{both})

and one that affects only broad spikes (I_{BS}). The second of these inputs, I_{BS} , drives learning and is the source of the negative image transmitted by MG cell narrow spikes. The shared input, I_{both} , is a curious feature according to our current understanding. After learning, I_{both} is cancelled by granule cell input and thus has no effect on either form of spiking. Its functional role, if any, is hence unclear.

Our proposed solution to the credit assignment problem in the ELL involves two functionally separate classes of MG cells defined by opposite responses to the sensory input that drives learning. How do these arise? It is known from past work that input from electroreceptor afferents on the skin is relayed to MG cells via a heterogeneous layer of small cells comprising distinct GABAergic and glutamatergic subpopulations and projecting to different layers within the ELL, including the ventral molecular layer where broad spikes are likely initiated (Bell et al., 2005; Hollmann et al., 2016; Zhang et al., 2007). The ventral molecular layer also contains GABAergic interneurons and excitatory feedback projections from the preeminential nucleus (Meek et al., 1999). Further delineating these circuits, for example by using connectomics, could provide an anatomical basis for the opposite and selective effects of sensory input on broad spikes in BS^+ and BS^- cells.

Learning in MG Cells without Output Error Backpropagation

Most multilayer error-correcting systems require some form of output-error feedback to guide learning. For example, output of the cerebellum that is correlated with errors has long been hypothesized to produce activity in the inferior olive, generating climbing fiber input to intermediate-layer Purkinje cells. This, in turn, drives complex-spike-mediated plasticity that is error correcting. In artificial networks, output errors are backpropagated to intermediate layers to control learning. In contrast, learning at the intermediate layer of the ELL, appears not to require feedback.

“Error,” in the ELL, corresponds to unwanted output neuron responses to the EOD. We have shown that MG cells provide a major contribution to error reduction at the output layer through the negative images they transmit. The error signal that drives learning in MG cells does not arise from feedback; there is no analog of the climbing fiber system in the ELL. Instead, MG cells generate their own error signal, internally, in the form of broad spikes that drive error-correcting plasticity. This avoids the need for error backpropagation, but it introduces a potential problem.

We have shown that the signals carried by the narrow-spike outputs of BS^+ and BS^- MG cells are of the right shape to provide negative images to output I and E cells, respectively. However, in the absence of an error signal from the output of the ELL, there is no guarantee that these transmitted negative images will be of the right amplitude to minimize the output response to the EOD, which is the output error. Nevertheless, our data strongly indicate that negative images conveyed by MG cells are, at least approximately, of the required strength. After 4 min of pairing, the putative MG cell component of the output cell negative image that we measured was ~ 1 mV (Figure 7) compared to an average 1.5 mV response evoked by sensory

input prior to cancellation (Figure S1F). Furthermore, an increase in the amplitude of sensory responses due to pairing was extremely rare in output cell responses (Figures S1B and S1D), arguing that cancellation signals are not too large. Plasticity at granule cell to output cell synapses could correct for moderately sized mismatches between the amplitude of MG inhibition and the required negative image. Alternatively, or in addition, there may be other active mechanisms for matching the scale of MG inhibition. Feedback connections from the preeminential nucleus to the ventral molecular layer of the ELL or plasticity (as of yet undescribed) at the synapses between MG cells and output cells are two possibilities for implementing such adjustments. Thus, although error correction at the MG layer does not directly rely on feedback, there may be some feedback-driven regulation of the strength of MG input onto the output layer. Whether or not such mechanisms exist, the absence of error backpropagation, either through a climbing fiber or any other mechanism, is an interesting feature given the difficulty of implementing more general backpropagation schemes in biological networks.

Broader Implications for Network Learning

The realization that negative images in the ELL are computed by both MG and output cells raises an interesting question: what are the advantages of having plasticity at two circuit layers rather than one? Definitive answers to this question will require additional work, but two possible advantages are suggested by features of the ELL. First, plasticity at granule cell synapses onto MG and output cells may operate at different rates, with MG plasticity providing a rapid source of approximate cancellation and output cell plasticity serving to fine-tune the cancellation over longer timescales. Second, estimates of the maximum amplitude of granule cell input onto output cells suggest that output cells could be stretched beyond the dynamic range of their plasticity if they were required to produce negative images without a contribution from MG cells.

Our results illuminate a number of key issues relevant to multilayer computation and learning beyond the ELL. First, the ELL reconciles learning with processing through a compartmental separation of learning and signaling functions in MG cells. Related two-compartment neuronal models have been proposed on theoretical grounds as biologically plausible implementations of machine-learning schemes in cerebral cortical circuits (Guerguiev et al., 2017; Körding and König, 2001; Marblestone et al., 2016; Schiess et al., 2016; Urbanczik and Senn, 2014). Hippocampal and neocortical neurons are increasingly implicated in complex and diverse forms of learning, including learning based on error or reward (Bittner et al., 2015; Hangya et al., 2015; Lacefield et al., 2019; Larkum et al., 2009). Our results may be relevant to understanding the plasticity that drives such learning.

Second, ELL wiring is organized on the basis of the effect of learning on MG cells, not according to their conventionally defined (narrow spike) response selectivities. This scheme ensures that plastic changes at synapses onto MG cells (at an intermediate layer of the ELL) are in the appropriate direction to aid cancellation at the output layer and, thus, represents a solution to the credit assignment problem in this circuit. More complex schemes are likely required to solve “deep” credit assignment

problems, such as those that may exist in cerebral cortical networks. Nevertheless, the “hard-wired” solution we describe may be sufficient for circuits like the cerebellum, where similar wiring-based solutions have been proposed (Apps and Garwicz, 2005; Herzfeld et al., 2015; Oscarsson, 1979).

Finally, it is notable that learning in the ELL appears to operate continuously without any mechanism for gating it on and off. This may be possible because anti-Hebbian learning assures stability. Like the ELL, hippocampal and cortical circuits may use internally generated error signals, such as dendritic plateau potentials, but these appear to drive Hebbian forms of plasticity. The complexity of inhibition in these circuits, much of it directed toward dendrites, may be required to gate learning (Basu et al., 2016; Lovett-Barron et al., 2014; Wilmes et al., 2016), assuring circuit stability despite the presence of unstable Hebbian forms of plasticity.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - Electrophysiology
 - Electrosensory stimulation
 - Intracellular pairing experiments (Figure 7)
 - Histology and morphological reconstructions
 - Biophysical model
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Software used and general statistical methods
 - Shuffled Responses
 - Smoothing
 - Subtracting pre-pairing response
 - Residual Variance
 - Cross Correlation
 - Analysis of magnitude of response related to Figures S3 and S4
- DATA AND CODE AVAILABILITY

ACKNOWLEDGMENTS

We thank R. Bruno for assistance with Neurolucida reconstruction, C. Dempsey for data analysis code contributions, H. Hou for Fiji scripts used for quantifying MG axon distributions, M. Hibi for the parvalbumin antibody, P. Roberts for advice on biophysical modeling, and H. Hou, K. Key, K. Perks, A. Wallach, and R. Warren for comments on the manuscript. This work was supported by grants from the NIH (NS075023) and the Irma T. Hirschl Trust to N.B.S. and by a grant from the National Science Foundation (NSF) (1025849) to N.B.S. and L.F.A. L.F.A. was further supported by the Simons and Gatsby Foundations and by the NSF NeuroNex Award DBI-1707398.

AUTHOR CONTRIBUTIONS

Conceptualization, N.B.S. and L.F.A.; Investigation, N.B.S. and A.N.Z.; Data Curation, S.Z.M., N.B.S., and A.N.Z.; Formal Analysis, S.Z.M. and L.F.A.; Software, S.Z.M.; Resources, S.Z.M.; Writing – Original Draft, N.B.S., L.F.A., S.Z.M., and A.N.Z.; Writing – Review & Editing, N.B.S., L.F.A., S.Z.M., and A.N.Z.; Funding Acquisition, N.B.S. and L.F.A.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 15, 2019

Revised: September 30, 2019

Accepted: October 21, 2019

Published: November 14, 2019

REFERENCES

- Apps, R., and Garwicz, M. (2005). Anatomical and physiological foundations of cerebellar information processing. *Nat. Rev. Neurosci.* 6, 297–311.
- Basu, J., Zaremba, J.D., Cheung, S.K., Hitti, F.L., Zemelman, B.V., Losonczy, A., and Siegelbaum, S.A. (2016). Gating of hippocampal activity, plasticity, and memory by entorhinal cortex long-range inhibition. *Science* 351, aaa5694.
- Bear, M.F., and Malenka, R.C. (1994). Synaptic plasticity: LTP and LTD. *Curr. Opin. Neurobiol.* 4, 389–399.
- Bell, C.C. (1981). An efference copy which is modified by reafferent input. *Science* 214, 450–453.
- Bell, C.C. (1982). Properties of a modifiable efference copy in an electric fish. *J. Neurophysiol.* 47, 1043–1056.
- Bell, C.C., and Russell, C.J. (1978). Effect of electric organ discharge on ampullary receptors in a mormyrid. *Brain Res.* 145, 85–96.
- Bell, C.C., Finger, T.E., and Russell, C.J. (1981). Central connections of the posterior lateral line lobe in mormyrid fish. *Exp. Brain Res.* 42, 9–22.
- Bell, C.C., Caputi, A., Grant, K., and Serrier, J. (1993). Storage of a sensory pattern by anti-Hebbian synaptic plasticity in an electric fish. *Proc. Natl. Acad. Sci. USA* 90, 4650–4654.
- Bell, C.C., Caputi, A., and Grant, K. (1997a). Physiology and plasticity of morphologically identified cells in the mormyrid electrosensory lobe. *J. Neurosci.* 17, 6409–6423.
- Bell, C.C., Han, V.Z., Sugawara, Y., and Grant, K. (1997b). Synaptic plasticity in a cerebellum-like structure depends on temporal order. *Nature* 387, 278–281.
- Bell, C.C., Meek, J., and Yang, J.Y. (2005). Immunocytochemical identification of cell types in the mormyrid electrosensory lobe. *J. Comp. Neurol.* 483, 124–142.
- Bell, C.C., Han, V., and Sawtell, N.B. (2008). Cerebellum-like structures and their implications for cerebellar function. *Annu. Rev. Neurosci.* 31, 1–24.
- Bittner, K.C., Grienberger, C., Vaidya, S.P., Milstein, A.D., Macklin, J.J., Suh, J., Tonegawa, S., and Magee, J.C. (2015). Conjunctive input processing drives feature selectivity in hippocampal CA1 neurons. *Nat. Neurosci.* 18, 1133–1142.
- Buonomano, D.V., and Merzenich, M.M. (1998). Cortical plasticity: from synapses to maps. *Annu. Rev. Neurosci.* 21, 149–186.
- Caporale, N., and Dan, Y. (2008). Spike timing-dependent plasticity: a Hebbian learning rule. *Annu. Rev. Neurosci.* 31, 25–46.
- Carnevale, N.T., and Hines, M.L. (2006). *The NEURON book* (Cambridge University Press).
- Destexhe, A., Mainen, Z.F., and Sejnowski, T.J. (1994). An Efficient Method for Computing Synaptic Conductances Based on a Kinetic-Model of Receptor-Binding. *Neural Comput.* 6, 14–18.
- Engelmann, J., van den Burg, E., Bacelo, J., de Ruijters, M., Kuwana, S., Sugawara, Y., and Grant, K. (2008). Dendritic backpropagation and synaptic plasticity in the mormyrid electrosensory lobe. *J. Physiol. Paris* 102, 233–245.
- Enikolopov, A.G., Abbott, L.F., and Sawtell, N.B. (2018). Internally Generated Predictions Enhance Neural and Behavioral Detection of Sensory Stimuli in an Electric Fish. *Neuron* 99, 135–146.
- Grant, K., Meek, J., Sugawara, Y., Veron, M., Denizot, J.P., Hafmans, T.G., Serrier, J., and Szabo, T. (1996). Projection neurons of the mormyrid electrosensory lateral line lobe: morphology, immunohistochemistry, and synaptology. *J. Comp. Neurol.* 375, 18–42.

- Grant, K., Sugawara, Y., Gómez, L., Han, V.Z., and Bell, C.C. (1998). The mormyrid electrosensory lobe *in vitro*: physiology and pharmacology of cells and circuits. *J. Neurosci.* 18, 6009–6025.
- Guerguiev, J., Lillicrap, T.P., and Richards, B.A. (2017). Towards deep learning with segregated dendrites. *eLife* 6, e22901.
- Han, V.Z., Bell, C.C., Grant, K., and Sugawara, Y. (1999). Mormyrid electrosensory lobe *in vitro*: morphology of cells and circuits. *J. Comp. Neurol.* 404, 359–374.
- Han, V.Z., Grant, K., and Bell, C.C. (2000). Reversible associative depression and nonassociative potentiation at a parallel fiber synapse. *Neuron* 27, 611–622.
- Hangya, B., Ranade, S.P., Lorenc, M., and Kepecs, A. (2015). Central Cholinergic Neurons Are Rapidly Recruited by Reinforcement Feedback. *Cell* 162, 1155–1168.
- Herzfeld, D.J., Kojima, Y., Soetedjo, R., and Shadmehr, R. (2015). Encoding of action by the Purkinje cells of the cerebellum. *Nature* 526, 439–442.
- Hollmann, V., Engelmann, J., and Gómez-Sena, L. (2016). A quest for excitation: Theoretical arguments and immunohistochemical evidence of excitatory granular cells in the ELL of *Gnathonemus petersii*. *J. Physiol. Paris* 110 (3 Pt B), 190–199.
- Jaffe, D.B., Ross, W.N., Lisman, J.E., Lasser-Ross, N., Miyakawa, H., and Johnston, D. (1994). A model for dendritic Ca²⁺ accumulation in hippocampal pyramidal neurons based on fluorescence imaging measurements. *J. Neurophysiol.* 71, 1065–1077.
- Kennedy, A., Wayne, G., Kaifosh, P., Alviña, K., Abbott, L.F., and Sawtell, N.B. (2014). A temporal basis for predicting the sensory consequences of motor commands in an electric fish. *Nat. Neurosci.* 17, 416–422.
- Knudsen, E.I. (1994). Supervised learning in the brain. *J. Neurosci.* 14, 3985–3997.
- Körding, K.P., and König, P. (2001). Supervised and unsupervised learning with two sites of synaptic integration. *J. Comput. Neurosci.* 11, 207–215.
- Lacefield, C.O., Pnevmatikakis, E.A., Paninski, L., and Bruno, R.M. (2019). Reinforcement Learning Recruits Somata and Apical Dendrites across Layers of Primary Sensory Cortex. *Cell Rep.* 26, 2000–2008.
- 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.
- Larkum, M.E., Nevian, T., Sandler, M., Polsky, A., and Schiller, J. (2009). Synaptic integration in tuft dendrites of layer 5 pyramidal neurons: a new unifying principle. *Science* 325, 756–760.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444.
- Lovett-Barron, M., Kaifosh, P., Kheirbek, M.A., Danielson, N., Zaremba, J.D., Reardon, T.R., Turi, G.F., Hen, R., Zemel, B.V., and Losonczy, A. (2014). Dendritic inhibition in the hippocampus supports fear learning. *Science* 343, 857–863.
- Marblestone, A.H., Wayne, G., and Kording, K.P. (2016). Toward an Integration of Deep Learning and Neuroscience. *Front. Comput. Neurosci.* 10, 94.
- Meek, J., Grant, K., Sugawara, Y., Hafmans, T.G.M., Veron, M., and Denizot, J.P. (1996). Interneurons of the ganglionic layer in the mormyrid electrosensory lateral line lobe: morphology, immunohistochemistry, and synaptology. *J. Comp. Neurol.* 375, 43–65.
- Meek, J., Grant, K., and Bell, C. (1999). Structural organization of the mormyrid electrosensory lateral line lobe. *J. Exp. Biol.* 202, 1291–1300.
- Mohr, C., Roberts, P.D., and Bell, C.C. (2003). The mormyromast region of the mormyrid electrosensory lobe: I. Responses to the electric organ corollary discharge and to electrosensory stimuli. *J. Neurophysiol.* 90, 1193–1210.
- Oscarsson, O. (1979). Functional Units of the Cerebellum - Sagittal Zones and Microzones. *Trends Neurosci.* 2, 143–145.
- Roberts, P.D., and Bell, C.C. (2000). Computational consequences of temporally asymmetric learning rules: II. Sensory image cancellation. *J. Comput. Neurosci.* 9, 67–83.
- Roberts, P.D., and Leen, T.K. (2010). Anti-hebbian spike-timing-dependent plasticity and adaptive sensory processing. *Front. Comput. Neurosci.* 4, 156.
- Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1988). Learning representations by back-propagating errors. *Neurocomputing: Foundations of Research* (MIT Press Cambridge), pp. 696–699.
- Sawtell, N.B., Williams, A., and Bell, C.C. (2007). Central control of dendritic spikes shapes the responses of Purkinje-like cells through spike timing-dependent synaptic plasticity. *J. Neurosci.* 27, 1552–1565.
- Schiess, M., Urbanczik, R., and Senn, W. (2016). Somato-dendritic Synaptic Plasticity and Error-backpropagation in Active Dendrites. *PLoS Comput. Biol.* 12, e1004638.
- Schiller, J., Schiller, Y., Stuart, G., and Sakmann, B. (1997). Calcium action potentials restricted to distal apical dendrites of rat neocortical pyramidal neurons. *J. Physiol.* 505, 605–616.
- Sperry, R.W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *J. Comp. Physiol. Psychol.* 43, 482–489.
- Sugawara, Y., Grant, K., Han, V., and Bell, C.C. (1999). Physiology of electrosensory lateral line lobe neurons in *Gnathonemus petersii*. *J. Exp. Biol.* 202, 1301–1309.
- Urbanczik, R., and Senn, W. (2014). Learning by the dendritic prediction of somatic spiking. *Neuron* 81, 521–528.
- von Holst, E., and Mittelstaedt, H. (1950). The reafference principle. *Naturwissenschaften* 37, 464–476.
- Wilmes, K.A., Sprekeler, H., and Schreiber, S. (2016). Inhibition as a Binary Switch for Excitatory Plasticity in Pyramidal Neurons. *PLoS Comput. Biol.* 12, e1004768.
- Zhang, J., Han, V.Z., Meek, J., and Bell, C.C. (2007). Granular cells of the mormyrid electrosensory lobe and postsynaptic control over presynaptic spike occurrence and amplitude through an electrical synapse. *J. Neurophysiol.* 97, 2191–2203.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
mouse parvalbumin7	Masahiko Hibi, Nagoya University,	N/A
Deposited Data		
Raw and analyzed data	This paper; Mendeley Data	https://doi.org/10.17632/4pxbpc6d5c.1
MG Cell morphology, simulation and analysis code of biophysical model	This paper	https://modeldb.yale.edu/259261 , ModelDB: accession number: 259261
Experimental Models: Organisms/Strains		
Gnathonemus petersii	Wild-born fish caught in Africa	N/A
Software and Algorithms		
Spike2	Cambridge Electronic Design	http://ced.co.uk/
MATLAB	MathWorks	https://www.mathworks.com
NEURON	Carnevale and Hines, 2006	https://www.neuron.yale.edu/neuron/
Python3	Python	https://www.python.org/

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for data should be directed to and will be fulfilled by the Lead Contact, Nate Sawtell (ns2635@columbia.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Male and female Mormyrid fish (7–12 cm in length) of the species *Gnathonemus petersii* were used in these experiments. Fish were housed in 60 gallon tanks in groups of 5–20. Water conductivity was maintained between 40–65 microsiemens both in the fish's home tanks and during experiments. All experiments performed in this study adhere to the American Physiological Society's *Guiding Principles in the Care and Use of Animals* and were approved by the Institutional Animal Care and Use Committee of Columbia University.

For surgery to expose the brain for recording, fish were anesthetized (MS:222, 1:25,000) and held against a foam pad. Skin on the dorsal surface of the head was removed and a long-lasting local anesthetic (0.75% Bupivacaine) was applied to the wound margins. A plastic rod was cemented to the anterior portion of the skull to secure the head. The posterior portion of the skull overlying the ELL was removed. In a subset of experiments, the valvula cerebelli was reflected laterally to expose the eminentia granularis posterior (EGp) and the molecular layer of the ELL, facilitating whole-cell recordings. Gallamine triethiodide (Flaxedil) was given at the end of the surgery (~20 µg/cm of body length) and the anesthetic was removed. Aerated water was passed over the fish's gills for respiration. Paralysis blocks the effect of electromotoneurons on the electric organ, preventing the EOD, but the motor command signal that would normally elicit an EOD continues to be emitted at a rate of 2 to 5 Hz.

METHOD DETAILS

Electrophysiology

The EOD motor command signal was recorded with a Ag-AgCl electrode placed over the electric organ. The command signal is the synchronized volley of electromotoneurons that would normally elicit an EOD in the absence of neuromuscular blockade. The command signal lasts about 3 ms and consists of a small negative wave followed by three larger biphasic waves. Onset of EOD command was defined as the negative peak of the first large biphasic wave in the command signal. For pairing experiments, the EOD mimic was presented 4.5 ms following EOD command onset. Recordings were started ~1 hour after paralysis.

Extracellular single-unit recordings were made using glass microelectrodes (2–10 MΩ) filled with 2M NaCl. The location of the ventrolateral zone (VLZ) was established using characteristic field potentials evoked by the EOD command. The precise location of the recording pipette with respect to the VLZ somatotopic map was subsequently determined by finding the skin region for which low-frequency electrosensory stimulation delivered via a local dipole electrode evoked multi-unit responses. Ampullary electoreceptor afferents, E cells and I cells are located in different layers of ELL and have distinctive electrophysiological characteristics

(Bell, 1981, 1982; Enikolopov et al., 2018). Ampullary afferents terminate in the deep layers of the ELL, exhibit highly regular spontaneous firing at ~50 Hz, and increase firing rate in response to an electrosensory stimulus that makes the pore of the receptor positive with respect to the basal face within the body. E cells are located in the plexiform layer and I cell in the ganglion layer. Spontaneous firing in E and I output cells is much more irregular and lower rate (~15 Hz) than in afferents. E cells are excited by the same stimulus polarity as ampullary afferents, while I cells are excited by the opposite polarity. Previous studies using intracellular recording and biocytin labeling and antidromic stimulation from the midbrain have shown that E and I cells correspond to two morphologically distinct types of ELL efferent cells known as large fusiform and large ganglion cells (Bell et al., 1993, Bell et al., 1997a). In addition to efferent cells, the other major large cells of the VLZ are the MG cells, also located in the ganglion layer. As in previous studies, recordings from MG cells were identified based on the presence of two distinct spike types (broad and narrow spikes) differing in both their waveform and frequency of occurrence (Bell et al., 1993, Bell et al., 1997a; Mohr et al., 2003; Sawtell et al., 2007). MG broad spikes with characteristics identical to those described above, could also be recorded in the molecular layer in the absence of narrow spikes, consistent with their separate initiation sites and the failure of narrow spikes to actively propagate to the soma or apical dendrites of MG cells (Engelmann et al., 2008; Grant et al., 1998).

For *in vivo* whole-cell recordings electrodes (8–15 M Ω) were filled with an internal solution containing, in mM: K-gluconate (122); KCl (7); HEPES (10); Na₂GTP (0.4); MgATP (4); EGTA (0.5), and 0.5%–1% biocytin (pH 7.2, 280–290 mOsm). No correction was made for liquid junction potentials. Only cells with stable membrane potentials more hyperpolarized than –45 mV and spike amplitudes > 40 mV were analyzed. Membrane potentials were recorded and filtered at 3–10 kHz (Axoclamp 2B amplifier, Axon Instruments) and digitized at 20 kHz (CED micro1401 hardware and Spike2 software; Cambridge Electronics Design, Cambridge, UK).

Electrosensory stimulation

The EOD mimic was a 0.2 ms duration square pulse delivered between an electrode in the stomach and another positioned near the electric organ in the tail. The amplitude was 25 μ A at the output of the stimulus isolation unit (stomach electrode negative). Recordings from ampullary afferents showed that firing rate modulations evoked by this mimic are within the range of those induced by the fish's natural EOD (Bell and Russell, 1978). We use the terms sensory input or sensory response to refer to the effect of the mimicked electric field on the ELL. Because we do not include prey-like electric fields (except for in Figure S2) the sensory input we discuss is entirely predictable on the basis of the EOD command signal and is therefore entirely uninformative to and 'unwanted' by the fish. Thus, we consider a situation where the ELL attempts to cancel *all* of its sensory input. It is important to appreciate that, in a natural setting, the mechanisms we analyze would only cancel the predictable self-generated component of the sensory input, leaving the unpredictable inputs of interest to the fish intact. For the experiments in Figure S2 low-frequency stimuli (white noise to which a 5–20 Hz band-pass Butterworth filter was applied) was delivered between electrodes attached to the edges of the recording chamber in front of and behind the fish on the side ipsilateral to the recording. For all of the cells included in the analysis of prey-like responses we confirmed the presence of typical responses to the EOD mimic. Stimulus amplitudes were chosen such that electroreceptor afferents were driven through their entire dynamic range.

Intracellular pairing experiments (Figure 7)

Effects of pairing on corollary discharge responses of output neurons were evaluated based on averages of 30 s of membrane potential data taken immediately before and after a 4 minute pairing period. Spikes were removed prior to averaging using a median filter. Recordings in which spike height changed by > 10% over the course of the pairing or in which apparent access resistance changed abruptly were excluded from the analysis. The amplitude and timing of current injection needed to counteract the effects of the EOD mimic was adjusted during the experiment based on online spike rasters. When recording time allowed, several different current strengths were tested sequentially. In such cases, we selected for analysis the case in which the current injection most effectively counteracted the sensory response or averaged the results in cases in which sensory responses were equivalent. Data was only analyzed for cells in which at least one control and one experimental pairing were obtained. As in previous studies of the ELL, we observed that washout of plasticity is rapid (occurring over roughly the same time course as the pairing) and no evidence for order effects, such as savings, was observed. Nevertheless, to minimize possible order effects, control and experimental pairing were performed in pseudo-random order across cells. Paired responses shown in the middle panel of Figures 7B–7D were computed by subtracting the pre pairing response.

Histology and morphological reconstructions

After recording, fish were deeply anesthetized with a concentrated solution of MS:222 (1:10,000) and perfused through the heart with a teleost Ringer solution followed by a fixative, consisting of 4% paraformaldehyde and 0.2% glutaraldehyde in 0.1 M phosphate buffer. The brains were postfixed for 12–24 hours, cryoprotected with 30% sucrose, and sectioned at 50 μ m on a cryostat. Sections were subsequently processed with the nickel-intensified avidin-biotin-peroxidase method (Vectastain Elite ABC kit) to reveal the biocytin filled cells, dehydrated, cleared in xylenes, and counterstained with neutral red to visualize the layers of ELL. Morphologically recovered neurons were inspected and reconstructed using a 100x oil immersion objective and a camera lucida system. Only well-filled cells in which fine processes, including apical dendritic spines, were clearly visible were selected for reconstruction. As described previously in a combined light and electron microscopy study (Meek et al., 1996), MG axons can be distinguished from

basal dendrites based on their initial lateral direction from the soma and looping course. Axons also were generally thinner than basal dendrites and exhibited numerous swellings along their length suggestive of terminal boutons. Processes were only considered to be axon (as opposed to basal dendrite) if they could be traced continuously back to their origination at the soma. The laminar distribution of axon was quantified as pixel area using a custom routine written in FIJI. Laminar boundaries were determined by superimposing photomicrographs of the neutral red counterstained section containing the soma of the labeled neuron onto the digitized camera lucida drawing.

Biophysical model

The morphology of the model MG cell was taken from a reconstruction of a biocytin-stained BS+ cell made using NeuroLucida software. Care was taken to accurately represent the thickness of dendritic processes. No attempt was made to reconstruct dendritic spines. The reconstructed cell consists of 78 compartments, and further divided to 230 segments. Simulation of cell activity was done using the simulation software NEURON (Carnevale and Hines, 2006), along with a Python 3 wrapper. Voltage gated Na^+ and K^+ channels inserted in the apical dendrites and axon are Hodgkin-Huxley type channels. Temperature was set to 20°C. To broaden the dendritic spikes, we added L-type calcium channels (Jaffe et al., 1994), consistent with experimental evidence (Engelmann et al., 2008; Sugawara et al., 1999). Narrow spikes were attenuated at the soma due to passive decay from their site of initiation at the thin axon initial segment. Next, we adjusted the conductance of these channels as described in the following table to achieve higher spike threshold for broad spikes. We matched distance between narrow and broad thresholds to values measured experimentally (Sugawara et al., 1999; Engelmann et al., 2008).

	g_l (S/cm ²)	leakage reversal potential (mV)	axial resistance (Ωcm)	Capacitance ($\mu\text{F}/\text{cm}^2$)	\bar{g}_{Na} (S/cm ²)	\bar{g}_{K} (S/cm ²)	\bar{g}_{Ca} (S/cm ²)
axon	0.0003	−70	100	1	4	0.5	0
apical	0.0003	−65	100	1	0.1	0.008	0.073
rest	0.0003	−65	100	1	0	0	0

In the table, ‘rest’ includes the soma, the somatic-connected axonal and apical compartments and all basal dendrite compartments. g_l is leakage conductance. \bar{g}_{Na} , \bar{g}_{K} , and \bar{g}_{Ca} are the maximal conductances of the sodium, potassium and L-type calcium channels, respectively.

To drive activity in the model cell, we inserted AMPA and GABA_A receptor channels (Destexhe et al., 1994). Reversal potential of the AMPA and GABA_A channel are 0 mV and −65 mV, respectively. Each reported simulation was run for 100 s and averaged. Broad and narrow spike rate traces were smoothed using a Gaussian filter with 8 ms and 5 ms standard deviation, respectively. Proximal compartments were defined as those whose center is within 100 μm of the center of the soma. For the rest of the section, units of time are in *ms* and units of maximal conductance, \bar{g} , are in μS .

Simulation of BS- cell (Figures 6B and 6D)

To achieve baseline firing of ~ 60 Hz for narrow spikes, and firing of ~ 3.5 Hz for broad spikes (we set firing of broad spikes ~ 2 Hz higher than observed in data in order to visualize inhibitory effects clearly) we activated AMPA channels in a compartment in the basal dendrites with timing onset $\sim \mathcal{N}(t_i, 6^2)$ where $t_i = 18 \cdot i$; $\bar{g} = 0.004 + 0.005 \cdot \text{Poisson}(0.074)$.

To simulate inhibitory inputs (Figure 6D, top), we activated a GABA_A channel in each proximal compartment with timing onset $\sim \mathcal{N}(t_i, 7^2)$ where $t_i = 200 \cdot i + 97$; $\bar{g} = 0.03$.

To simulate cancellation of inhibitory input by parallel fiber input (Figure 6D, bottom), we added activation of an AMPA channel in each apical dendrite with timing onset $\sim \mathcal{N}(t_i, 3^2)$, where $t_i = 200 \cdot i + 97.3$; $\bar{g} = 0.00021$. To simulate opposite responses to sensory inputs (Figure 6B) we activated 1) a GABA_A channel in each proximal compartment with timing onset $\sim \mathcal{N}(t_i, 2^2)$, where $t_i = 200 \cdot i + 97$; $\bar{g} = 0.2$ and 2) AMPA channel in a compartment in the basal dendrites with timing onset $\sim \mathcal{N}(t_i, 4^2)$, where $t_i = 200 \cdot i + 100$; $\bar{g} = 0.021$.

Simulation of BS+ cell (Figure 6C)

Baseline firing in BS+ was achieved by inserting 3 inputs; 1) AMPA input in a basal dendrite compartment with timing onset $\sim \mathcal{N}(t_i, 10^2)$, where $t_i = 36 \cdot i + 10$; $\bar{g} = 0.004 + 0.0053 \cdot \text{Poisson}(0.3)$. 2) AMPA input in an apical dendrite compartment (the apical compartment connected to the soma) with timing onset $\sim \mathcal{N}(t_i, 10^2)$, where $t_i = 36 \cdot i + 28$; $\bar{g} = 0.002 + 0.0055 \cdot \text{Poisson}(0.3)$. And 3) GABA_A input in each proximal apical dendrite compartment with timing onset $\sim \mathcal{N}(t_i, 5^2)$, where $t_i = 10 \cdot i + 10$; $\bar{g} \sim \mathcal{N}(0.009, 0.001^2)$. This tonic inhibition only affected broad spike rate (not shown). To simulate opposite responses to sensory inputs we removed the tonic inhibition from $t_i - 10$ to $t_i + 10$ where $t_i = 200 \cdot i + 100$, and added GABA_A input in a basal dendrite compartment with timing onset $\sim \mathcal{N}(t_i, 7^2)$, where $t_i = 200 \cdot i + 97$; $\bar{g} = 0.1$.

QUANTIFICATION AND STATISTICAL ANALYSIS

Software used and general statistical methods

Data were analyzed offline using Spike2 (Cambridge Electronic Design) and MATLAB (MathWorks) software. Biophysical model was simulated using NEURON module in Python 3 and analyzed using Python 3. No statistical methods were used to predetermine sample size. The experimenters were not blinded to the condition during data collection or analysis. Non-parametric tests were used for testing statistical significance. Unless otherwise indicated, we used the two-sided Wilcoxon rank sum test for unpaired samples and the Wilcoxon signed ranks test for paired samples. Differences were considered significant at $p < 0.05$. 3 stars indicate that $p < 0.001$.

Shuffled Responses

In Figure 3C we used bootstrap analysis with 1,000 repetitions to measure the mean and SEM of the shuffled response.

Smoothing

Output neuron spike trains and MG cell narrow spike trains were smoothed using a 10 ms Gaussian kernel. MG cell broad spikes (which have a much lower baseline rate of ~ 1 Hz) were smoothed using a 25 ms Gaussian kernel.

Subtracting pre-pairing response

MG cell narrow spikes and membrane potential tend to have an excitatory response to the command alone prior to any pairing. Therefore, to isolate negative image responses (or sensory responses from experiments where the mimic was paired with EOD commands) we subtracted the pre-pairing response. For pooling MG narrow spike responses after > 1 hour pairing (Figures 4B and 4C, > 1 hr) we subtracted the average pre-pairing response taken from a different set of MG cells for which data was obtained prior to pairing, i.e., those from the 4 minutes pairing experiments.

Residual Variance

(Figures 2D and S1A)

$$\text{Residual Variance} = \frac{\langle \text{pair}^2 \rangle - \langle \text{pair} \rangle^2}{\langle \text{EOD}^2 \rangle - \langle \text{EOD} \rangle^2}$$

$\langle x \rangle \equiv E[x]$. “Pair” is the response after some period of pairing/learning (cmd+ mimic) and “EOD” is the response to the EOD mimic.

Cross Correlation

We first define,

$$R_{xy}(m) = \sum_{n=1}^{n=N-m} x_{n+m} y_n \quad \text{for } m \geq 0.$$

$$R_{xy}(m) = R_{yx}(-m) \quad \text{for } m < 0.$$

m is the lag, N is sample size.

Cross correlation for lag m was measured as $xcorr(m) = \frac{R_{xy}(m)}{\sqrt{R_{xx}(0)R_{yy}(0)}}$. The normalization limits the range of values to be $[-1, 1]$ akin to Pearson Coefficient. The traces were generally mean subtracted. However, for processed traces (i.e., when pre pairing was subtracted) the median was subtracted. Cross correlation was measured across lag of 8 ms, and the value reported is $\max_{x \in m} (|xcorr(x)|)$. However, for Figure 3C showing lack of relationship in the mimic response between broad spikes and narrow spikes we calculated lag up to 15 ms.

Analysis of magnitude of response related to Figures S3 and S4

To calculate the magnitude of responses we focused our analysis on the area around the peak response, restricted to a range 0.5–45 ms following the command. We report the average change from pre pairing in a window of 16 ms around the peak response. Peak response time is the index of $\max(|\text{EOD}|, |\text{CD}|)$. For broad spikes and output cells, the traces were mean subtracted. For narrow spikes and membrane potential of MG cells we first subtracted pre-pairing from the trace and then the trace was baseline subtracted (baseline was approximated by the median). In all the plots using this method, the time of peak response found for the data plotted in the x axis was used for the data plotted in the y axis.

DATA AND CODE AVAILABILITY

Data and data analysis code will be available at Mendeley: <https://doi.org/10.17632/4pxbpc6d5c.1>

Morphology of the MG cell as well as code of biophysical model will be available at ModelDB: accession number: 259261, <https://modeldb.yale.edu/259261>

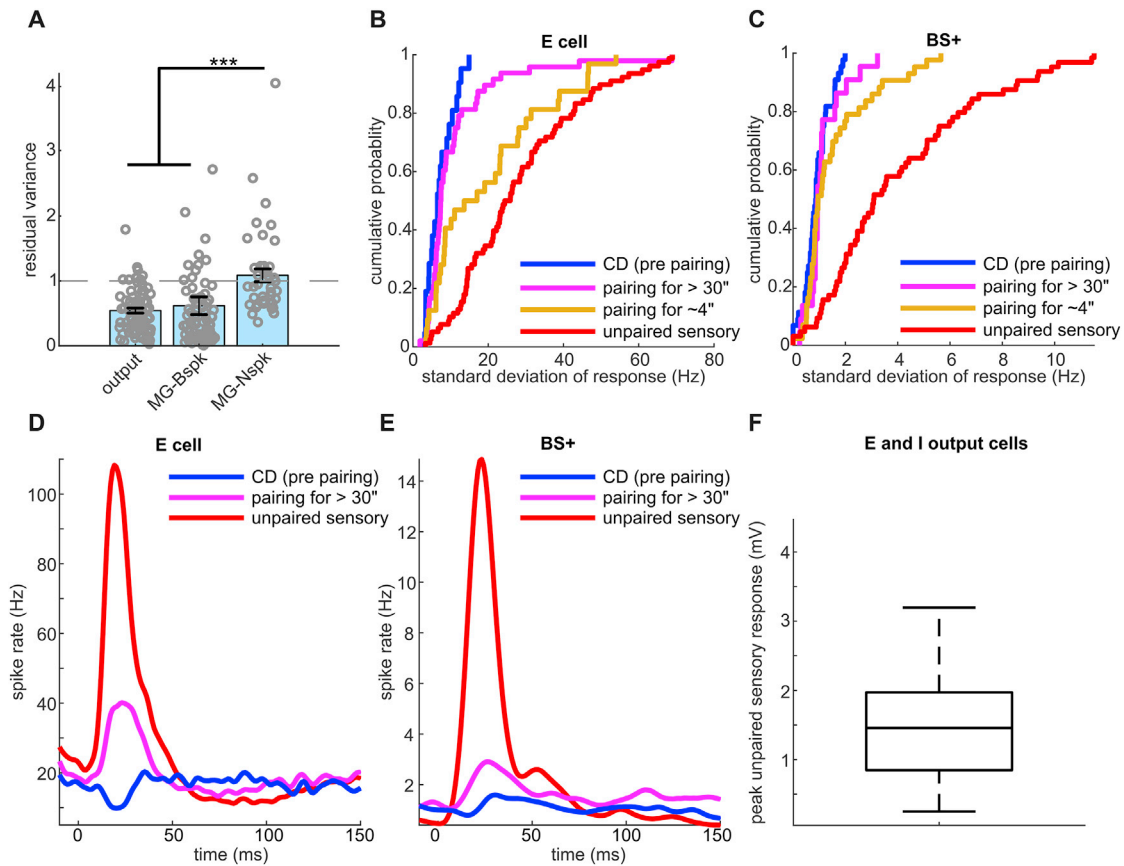


Figure S1. Time Course of Sensory Cancellation, Related to Figure 2

(A) Similar to Figure 2D, but from brief (~4 min) periods of sensory pairing. In this dataset we also recorded corollary discharge responses before pairing (pre pairing) and subtracted them from responses after pairing in order to reveal the negative image. Cancellation of the effects of the mimic (residual variance < 1), was observed in output ($n = 78$) and MG-Bspk ($n = 78$) but not in MG-Nspk responses ($n = 44$; $p < 0.0001$). (B) Cancellation of the effects of the mimic in E output neurons measured as the standard deviation of the firing rate response during pairing. After > 30 minutes ($n = 51$) and after 4 minutes ($n = 32$) of pairing, responses are significantly reduced relative to the standard deviation of the response to sensory input alone ($n = 78$; $p < 0.0001$ and $p < 0.001$, respectively). After > 30 minutes of pairing most cells are within the range of pre-pairing responses ($n = 21$). Note, the rate and accuracy of cancellation is qualitatively similar in I cells. However, quantifying cancellation in I cells is hindered by firing rate rectification. Hence data from I cells are not shown here. (C) Cancellation of the effects of the mimic in broad spikes of BS+ MG cells. After > 30 minutes of pairing ($n = 22$) and after 4 minutes ($n = 43$), paired responses to the mimic are significantly reduced relative to the standard deviation of the response to the sensory input alone ($n = 65$; $p < 0.0001$). Most cells are within the range of pre-pairing responses ($n = 44$). (D) Average responses of E cells under same conditions as described above. Note, that responses to the paired stimulus are markedly reduced after pairing but are still larger than pre-pairing responses, consistent with under- rather than over-cancellation. (E) Same as in D but for BS+ cells. (F) Distribution of peak amplitudes of subthreshold sensory responses on membrane voltage measured intracellularly in E and I output cells. Spikes were blocked by hyperpolarization or removed offline using a median filter.

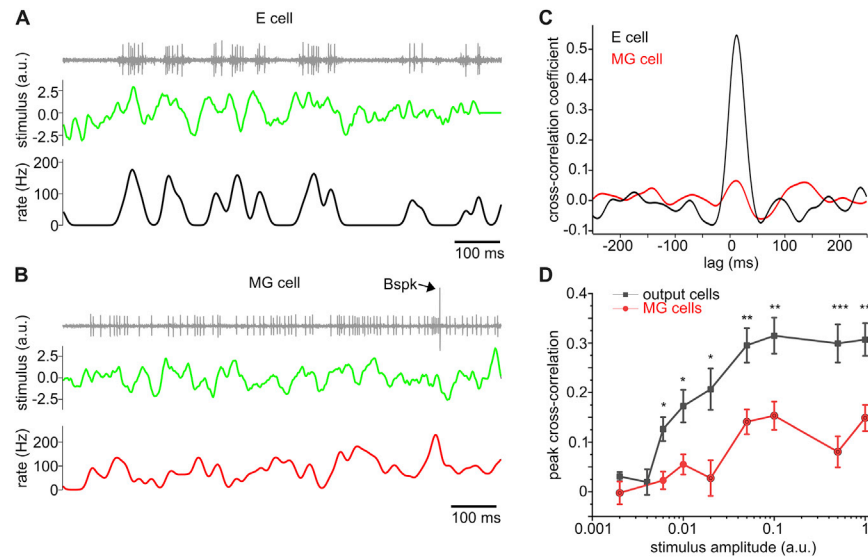


Figure S2. Responses of Output and MG Cells to Prey-like Stimuli, Related to Figure 2

(A) Raw voltage trace (top), stimulus waveform (middle, 5-20 Hz band-passed noise stimulus), and smoothed firing rate (bottom, 35 ms Gaussian kernel) for an example E cell. (B) Same display for an example MG cell. Arrow indicates the occurrence of a broad spike. Firing rate analysis was performed only on the narrow spikes. (C) Cross-correlation between the firing rate and the stimulus waveform for the example output cell (black) and MG cell narrow spikes (red) shown in A and B. (D) Summary plot of the magnitude of the peak cross-correlation value for output cells (E cells: $n = 7-16$ cell per data point; I cells: $n = 6-17$ cell per data point) versus MG cells ($n = 4-14$ cell per data point) across a range of stimulus amplitudes. *** indicates $p < 0.001$, ** indicates $p < 0.01$, * indicates $p < 0.05$, Student's t test with Bonferroni correction for multiple comparisons.

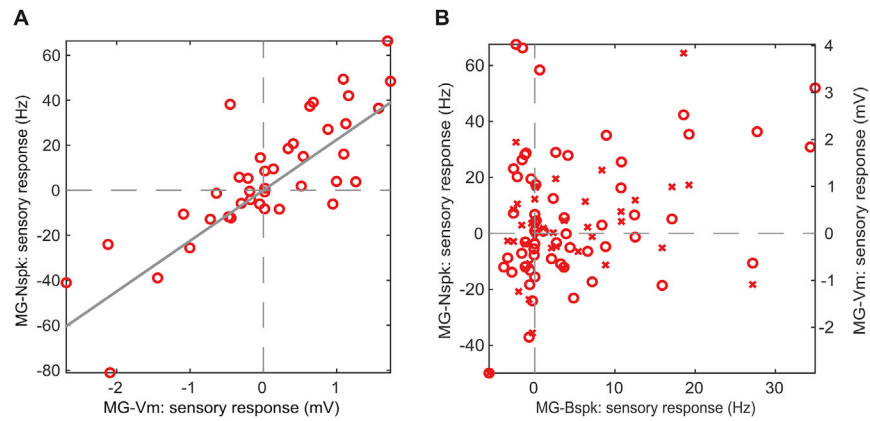


Figure S3. EOD Responses of MG Cell Broad and Narrow Spikes Are Uncorrelated, Related to Figure 3

Additional analysis related to Figure 3C. Data points indicate the average change in response relative to the pre-pairing condition calculated in a 16 ms window centered on the peak response (see STAR Methods). (A) Sensory responses of narrow spikes (Nspks) (circles) and subthreshold membrane potential (Vm; crosses) recorded in the same cell exhibit an approximately linear relationship (gray line, linear regression; $n = 41$, $r^2 = 0.64$). (B) In contrast, sensory responses of Nspk (circles) and Vm (crosses) exhibit no clear relationship with the sensory responses of broad spikes (Bspks) recorded in the same cell (Nspk: $n = 60$, Vm: $n = 42$).

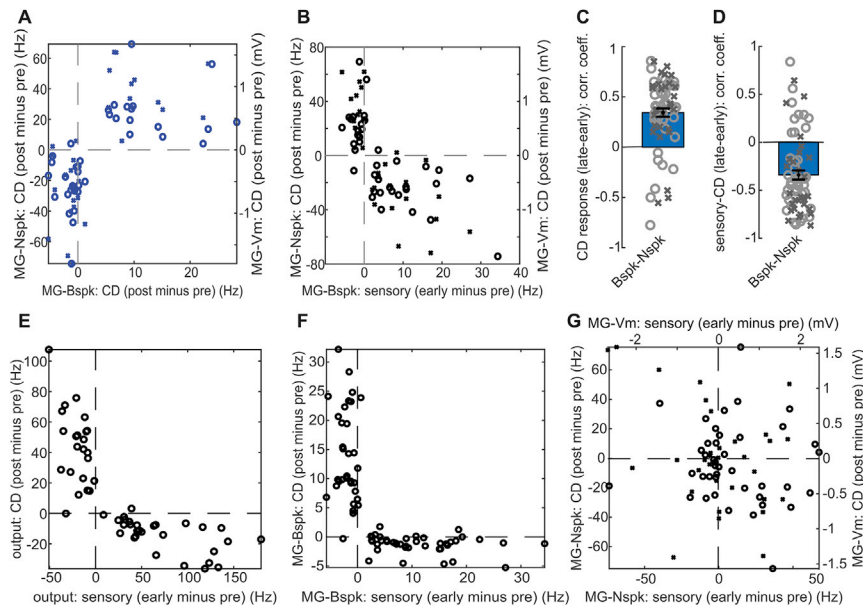


Figure S4. MG Cells Transmit a Negative Image of the Response of Broad Spikes to the EOD Mimic, Related to Figure 3

All data are from experiments in which paired sensory input was delivered for 4 minutes. Data points in the scatterplots indicate the average change in response relative to the pre-pairing condition calculated in a 16 ms window centered on the peak response (see [STAR Methods](#)). (A) Changes in corollary discharge (CD) response of narrow spikes (Nspks; circles; $n = 30$) and Vm (crosses; $n = 38$) have the same polarity as changes in the CD response of broad spikes (Bspks) recorded in the same cell and (B) opposite polarity to the sensory response of Bspks recorded in the same cell. (C) Related to [Figure 3F](#), but here responses are changes *during* the pairing period, i.e., subtracting early (initial ~ 30 s) from late (last ~ 30 s) pairing responses, rather than subtracting pre from post pairing responses. Changes in CD responses during pairing were positively correlated between Nspks and Bspks (circles, $r = 0.33$, $n = 41$) and between Vm and Bspks, (crosses, $r = 0.37$, $n = 30$). (D) Related to the middle bar in [Figure 3G](#), changes in Nspk and Vm CD responses during pairing were negatively correlated with Bspk sensory responses ($r = -0.30$, $n = 41$ for Nspks and $r = -0.40$, $n = 30$ for Vm). (E) Changes in CD responses are opposite to sensory responses in output neurons ($n = 52$). (F) Same as in (E) but for Bspks ($n = 74$). (G) Changes in CD responses of Nspks (circles, $n = 41$) and Vm (crosses, $n = 33$) are uncorrelated with the Nspk response to sensory input. Note, sensory response can be excitatory or inhibitory. Sensory-evoked firing rate response of I cells, BS- cells as well as CD response of E cells and BS+, are rectified at 0.

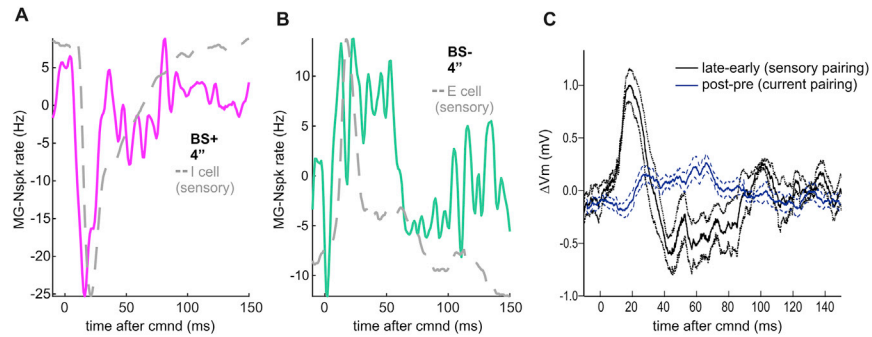


Figure S5. Negative Images during Pairing with Sensory Input, Related to Figure 4

(A–B) Analysis related to Figure 4B,C, but showing pooled changes in corollary discharge (CD) responses during (as opposed to after) pairing measured by subtracting early (initial ~30 s) from late (last ~30 s) pairing responses (BS+, $n = 13$; BS-, $n = 9$). BS+ and BS- cells show approximately opposite changes, similar to those observed after pairing (Figure 4B). For BS- the average change in the CD response during pairing is smaller than the change relative to the pre pairing condition (Figure 4C). This is likely due to the fact that the initial narrow spike (Nspk) response to the mimic is, on average, in the same direction as the negative image. Also, for BS- cells the average change in the Nspk CD response during pairing is smaller than the average change in the CD response observed after pairing (Figure 4B). This may be due to the fact that during pairing there is some shunting of the CD input by the sensory input. (C) Related to Figure 7D. Average traces from experiments ($n = 7$) in which control pairings with an inhibitory stimulus were compared with pairing with hyperpolarizing current injections which evoked similar responses. Negative images measured under control conditions during pairing (measured by subtracting the early response from the late response during pairing; black trace) were larger than changes induced by current injections (blue trace). The latter was measured after pairing because the current injection obscures the membrane potential trace during pairing.

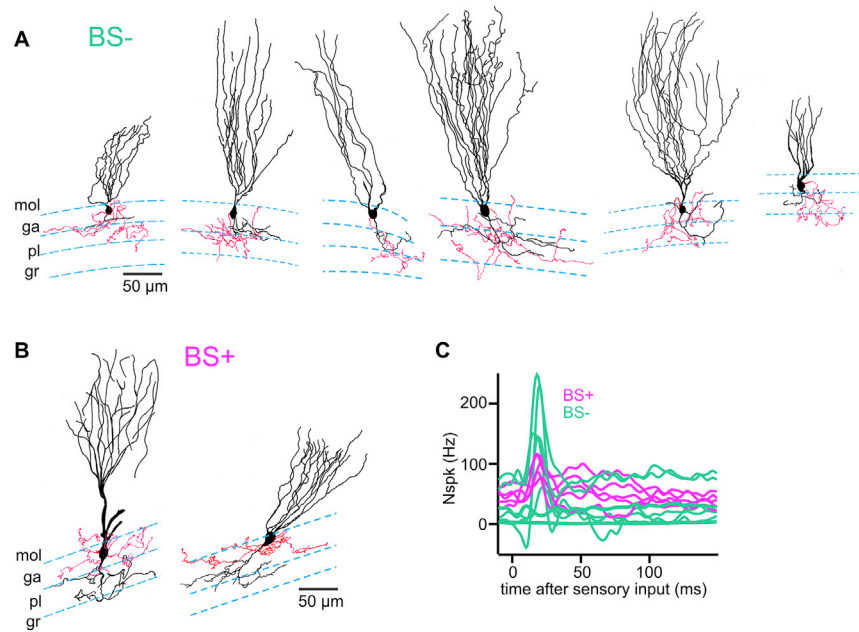


Figure S6. BS⁺ and BS⁻ Cells Correspond to Distinct Anatomical Classes, Related to Figure 5

(A) Camera lucida reconstructions of remaining BS⁻ cells not shown in Figure 5. Axon is drawn in red and ELL layer boundaries indicated by dashed lines. mol, molecular layer, ga, ganglion layer, pl, plexiform layer, gr, granular layer. (B) Camera lucida reconstructions of remaining BS⁺ cells not shown in Figure 5. (C) Narrow spike (Nspk) firing evoked by the EOD in morphologically reconstructed BS⁺ ($n = 5$) and BS⁻ ($n = 7$) cells. Nspk responses of two BS⁻ cells recorded in the dorsolateral zone of ELL were excluded from the clustering analysis in Figure 5E because the absolute magnitude of Nspk responses is expected to differ systematically across the zones. However, the results are in qualitative agreement with those from VLZ MG cells. A few cells show negative firing rates because we computed the Nspk response to sensory input by subtracting pre pairing from early pairing responses.