

**Visuomotor Coupling Shapes a
Sensorimotor Circuit in Mouse Cortex for
Visual Flow Predictions**

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Alexander Attinger

aus Dübendorf (ZH), Schweiz

Basel, 2018

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel
edoc.unibas.ch



Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung 4.0
International Lizenz.

Genehmigt von der Philosophisch-Naturwissenschaftlichen
Fakultät auf Antrag von

Dr. Georg Keller (Dissertationsleiter)

Prof. Dr. Sonja Hofer (Fakultätsvertreterin)

Prof. Dr. Rainer W. Friedrich (Koreferent)

Basel, den 12. Dezember 2017

Prof. Dr. Martin Spiess (Dekan)

“No one ever seems quite satisfied with anyone else’s views.”

Kenneth Craik, Philosopher

Acknowledgements

I would like to thank my thesis advisor Georg Keller, for inspiration, guidance and continuous support throughout my PhD. It has been a great experience, both in the lab and in the mountains. I am grateful to Bo Wang for all the experiments we did together and all the night shifts he took over. This thesis would look different without him. I am grateful to Marcus Leinweber, together we implemented ivCRACM and he was always there for guidance and valuable discussions. Also, his chocolate ice cream is fantastic. Without Daniela Gerosa, we couldn't work, she produces all our AAVs and makes sure we have all the tools and toys to do our experiments. I would like to thank all the past and current members of the Keller Lab for the great times we had and a very enriching scientific environment. I would like to thank the members of my thesis committee, Sonja Hofer, Rainer Friedrich and Stefan Rotter for their support and valuable input.

A special thanks to all the people at FMI who made my time here very enjoyable. I am grateful to the excellent facilities, especially IT for providing us with an awesome infrastructure and the machine workshop for producing so many custom parts.

I would like to thank my family for all the support and encouragement to pursue my interests.

I am eternally grateful for Antonia Drinnenberg, my love, my partner, my friend, for all the adventures we have experienced and all the adventures to come.

Prologue	3
Predictive processing theory of brain function.....	3
Aim of this thesis.....	8
Chapter 1: Visuomotor Coupling Shapes the Functional Development of Mouse Visual Cortex	10
Summary	10
Introduction	11
Results.....	13
Discussion.....	24
Author Contributions	27
Acknowledgements.....	27
Supplementary Information	28
Methods.....	37
Chapter 2: A Sensorimotor Circuit in Mouse Cortex for Visual Flow Predictions	46
Summary	46
Introduction	46
Results.....	50
Discussion.....	59
Supplementary Information	64
Methods.....	70
Quantification and Statistical Analysis.....	74
Epilogue.....	77
Visuomotor experience shapes predictive processing in visual cortex	77
Motor-based prediction of visual flow from A24b to V1.....	79
Outlook	80
Final remarks.....	81
References	82

PROLOGUE

What we see, feel and hear is tightly coupled to our actions. Consequently, our sensory input is often dominated by self-generated sensory feedback. For example, as we run forwards, the world around us appears to move backwards. How can we distinguish this backward visual motion caused by our own movement from the visual motion caused by an object that is moving relative to us? Any agent that interacts with the environment through motor actions—animal, human, or robot—faces the problem how to disentangle the sensory input caused by its own actions from the sensory input generated external world.

Predictive processing theory of brain function

Our brain has to infer the state of the environment from noisy and possibly ambiguous sensory input (Craik, 1943; Dayan et al., 1995; Knill and Pouget, 2004). Predictive processing is a framework of how our brain might perform this inference in a probabilistic manner. Predictive processing also constitutes an elegant theoretical framework in which the problem of self-generated sensory feedback can be addressed. According to predictive processing theories, the brain is a statistical machine that aims to generate predictions about the causes for its sensory input. To predict the incoming sensory input, the brain maintains an internal model of the world (**Figure 1.1**). The difference between the prediction and the sensory input—the prediction error—is then used to refine the internal model and to adapt the behavior (Clark, 2013; Friston et al., 2006; Mumford, 1992). Internal models have also been proposed to play an important role in motor control, where internal models use the current state of the motor system and motor command to predict the next state of the motor system as well as the sensory consequences of the action (e.g. (Wolpert et al., 1995)). Similarly, dopamine neurons signal the difference between actual and predicted reward (Schultz et al., 1997). In the example of self-generated visual flow caused by forward locomotion, predictions of visual flow input could be compared to actual visual flow input. In this manner, predictions suppress self-generated sensory input, but objects moving relative to us would cause large prediction errors and are therefore more salient and easier to detect.

Hierarchical predictive processing

Distinct algorithms have been proposed for the implementation of predictive processing (Spratling, 2017). In a hierarchical predictive-processing model proposed by (Friston et al., 2006), feedback predictions based on prior knowledge are combined with feedforward sensory input in order to compute the most likely interpretation of sensory data (**Figure 1.2**). The model, which incorporates several levels, is hierarchical in the sense that the representations and the causal relations between them become more complex higher in the hierarchy (for example, elements in a low level might

represent small regions of visual space, while higher levels process larger regions). Specifically, the model consists of two distinct populations of units on every level: (1) *State units*, which combine feedforward input from lower levels, input from other state units on the same level, and input from error units. Their activity encodes the most likely cause for the feedforward input, which they send as a prediction of input to lower levels. (2) *Error units*, which compare the activity of the current level with a prediction from higher levels. The resulting prediction error is sent to higher areas, where state units alter their activity to minimize the prediction error. In this manner, prediction errors are iteratively minimized across all levels. Using this strategy, the predictive-processing model can infer the cause for the sensory input, i.e. the state of the environment that maximizes the posterior probability of the environmental states, given the current sensory input and past experience (for more details on the underlying mathematical framework, see (Friston, 2009, 2010; Friston et al., 2006).

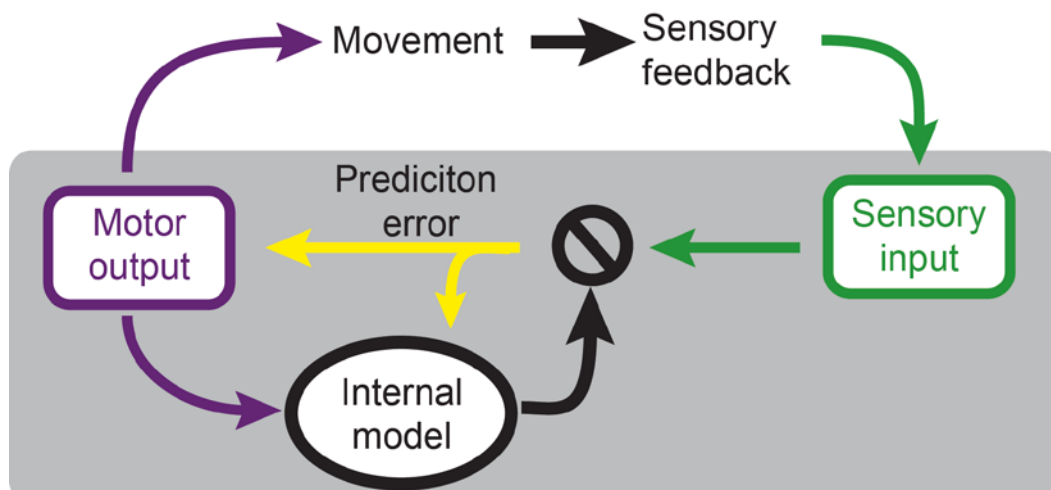


Figure 1.1. Motor output causes sensory feedback. In predictive processing, an internal model of the world generates a prediction of the expected sensory feedback based on the motor output. The prediction is then compared to the actual sensory input to compute a prediction error, or mismatch between the two.

In predictive processing terms, perception consists of finding the state of the internal model that minimizes the current prediction errors across all hierarchies and therefore best explains the sensory input. Perceptual learning can also be regarded as a form of minimizing prediction errors. It can be viewed as adjusting or extending the structure of the internal model, also with the goal of minimizing prediction errors. Predictive processing can also be extended to incorporate action planning and motor output: Since prediction errors also depend on the sensory input, prediction errors can also be minimized by changing the sensory input and since sensory input is coupled to motor output, it can be changed by action. Prediction errors can therefore be minimized by acting on the environment and

thereby change the sensory input. This framework formalized by Friston and offers the very intriguing approach of a unifying theory of brain function (Clark, 2013; Friston, 2010).

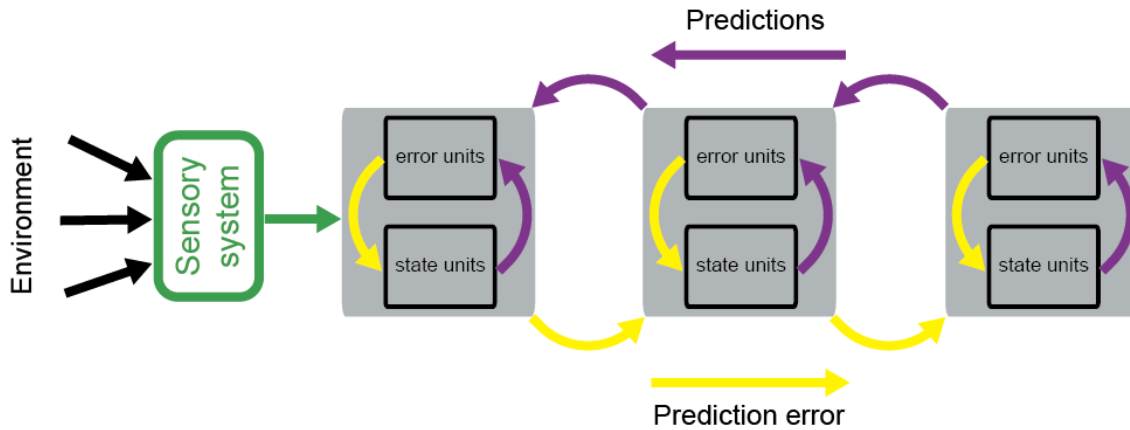


Figure 1.2. Schematic of the hierarchical predictive-processing model with three hierarchical layers. Input to the lowest level is generated by sensory systems. State units combine input from lower levels and the current level to infer and encode the causes and states of the environment. The error units combine top down input with input from the state units of the current level to calculate prediction errors between the state of the current level and a top-down prediction. In this scheme, feedforward signals encode prediction errors and feedback connections signal predictions. Prediction errors can be interpreted as input that still needs to be explained by the feedback predictions. By combining multiple feedforward error units, representations become more complex in higher levels. The state of the environment causing the sensory input is inferred by minimizing the prediction error across the hierarchical levels.

Cortical activity explained by the principles of predictive processing

Can predictive processing be used to understand cortical processing? Previous studies used predictive processing principles to explain neuronal activity in different cortical areas, including visual and auditory cortex. A computational, predictive-processing model with two hierarchical levels trained on natural images develops receptive fields similar to the receptive fields found in primary visual cortex (V1), including extra-classical receptive field effects, which were dependent on the predictive input from the higher level (Rao and Ballard, 1999). In auditory cortex, predictive processing principles can explain the large responses evoked by unexpected tones embedded in a series of familiar sounds, which simple habituation to the familiar stimulus can not account for (Garrido et al., 2009).

Strong evidence that sensory cortical areas employ predictive-processing principles comes from experiments where self-generated sensory feedback is controlled and manipulated by the experimenter. In these studies, neural activity is recorded while the subject is engaged in a sensorimotor task. The sensory feedback is briefly perturbed to induce a mismatch between the behavior of the subject and the sensory feedback. Using this approach, neurons signaling mismatch, or prediction error, have been found in mouse V1 (Keller et al., 2012; Saleem et al., 2013; Zmarz and

Keller, 2016). Similar feedback mismatch responses have been observed in primary auditory cortex of primates (Eliades and Wang, 2008) and in primary auditory pallium of songbirds (Keller and Hahnloser, 2008). These results are consistent with a predictive processing interpretation where sensory areas compare self-generated sensory input to a prediction based on motor output.

Predictive processing postulates the existence of state units as well as error units and makes predictions about the functional properties of feedforward and feedback connections. Based on the anatomical properties of excitatory neurons in superficial layers, they are proposed to be error units that compare the activity of the state units that might be located in the deeper layers to feedback predictions (Bastos et al., 2012; Shipp, 2016). Consistent with this idea, mismatch neurons are found in layer 2/3 of mouse visual cortex (Keller et al., 2012; Zmarz and Keller, 2016). How the computations proposed in predictive processing map to the connectivity between cortical regions, and to the laminar organization of excitatory and inhibitory neurons, is still largely unclear.

Predictive processing in mouse visual cortex

In my thesis work, I addressed how predictive processing can further our understanding of how the brain deals with the complex and ever changing environment. In particular, I am interested how visual information is integrated with information about body movement and if this integration can be understood in a predictive processing framework. The visual cortex of the mouse is an excellent model to investigate visuomotor integration, as it allows precise control over the visual input as well as access to defined population of neurons for the recording and manipulation of activity.

Feedforward and feedback inputs to V1

V1 receives feedforward visual input from the lateral geniculate nucleus of the thalamus (LGN), which in turn receives input from different retinal ganglion cell types that extract distinct features from the visual scene (Dhande et al., 2015). Feedback input to V1 mainly originate from higher visual cortical areas, in cats and primates, but also in the mouse (Wang and Burkhalter, 2007; Wang et al., 2012). While numerous studies investigated the role of feedback from higher visual areas onto visual processing in V1 in cats and primates, the role of feedback connections for processing in mouse V1 is less explored. Feedback projections are thought to be important for adapting or enhancing feedforward input via mechanisms like surround suppression, object binding, figure-ground segregation, attention, and visual awareness (Gilbert and Li, 2013; Zipser et al., 1996). Predictive processing provides a unifying framework to understand the role of feedback projections from higher visual areas for visual processing (Kanai et al., 2015; Spratling, 2010). V1 also receives prominent inputs from non-visual cortical areas. Activation of projections from anterior cingulate cortex (ACC) to

V1 modulates responses in V1 and improves performance on a visual discrimination task (Zhang et al., 2014). Similarly, stimulation of axons from the auditory cortex sharpens the tuning in V1 neurons (Ibrahim et al., 2016). Whether these feedback connections can be understood as predictions of visual input, as proposed by the predictive processing framework, remains to be seen.

One important question about feedback connections that has not been well explored is what kind of signal they provide to visual cortex. This is particularly interesting when investigating feedback projections to visual cortex from another sensory modality, like auditory cortex. In auditory cortex, neurons respond to sound and are organized tonotopically (Linden et al., 2003). How the dense projection from auditory cortex to visual cortex (Ibrahim et al., 2016), map to the retinotopic organization of visual cortex is not known. This coordinate transformation poses a central challenge for any projection conveying specific information. Interestingly, where the activity of V1 projecting axons was recorded, strong visual signals were recorded, both in axons from the retrosplenial cortex (Makino and Komiyama, 2015) as well as axons from ACC (Fiser et al., 2016). It will be interesting to see how inputs from higher areas are organized functionally, to what extent they represent information in a coordinate system that is similar to the retinotopic coordinate system of V1 and where the coordinate transfer is realized.

In addition to feedforward and feedback input, V1 receives neuromodulatory input, thought to be regulating brain states, attention and arousal (Lee and Dan, 2012). How different neuromodulatory input connections shape activity in visual cortex is not well understood. Activation of cholinergic axons in V1 improves behavioral performance in a visual discrimination task and enhances visual responses in mice (Pinto et al., 2013). Locomotion (Niell and Stryker, 2010), as well as general arousal (Vinck et al., 2015), have similar effects and lead to the release of acetylcholine and noradrenaline in V1 (Reimer et al., 2016). Consequently, locomotion is thought to change the gain of visual responses by increasing cholinergic and noradrenergic input to V1 (Fu et al., 2014; Polack et al., 2013). The impact of other neuromodulators like dopamine and serotonin on visual processing has not been explored in mouse V1.

Previous studies demonstrated that locomotion does not only modulate visual responses in V1, but can also drive activity, even in the absence of visual stimulation (Fu et al., 2014; Keller et al., 2012; Pakan et al., 2016; Saleem et al., 2013). Neurons in V1 encode locomotion speed and the difference—or mismatch—between locomotion speed and visual speed (Keller et al., 2012; Saleem et al., 2013; Zmarz and Keller, 2016). These effects are difficult to reconcile with the idea of gain modulation by locomotion. Predictive processing offers an alternative framework in which locomotion induced

activity in V1 can be explained. The functional role of motor related activity in V1, as well as its origin, are the major themes of my thesis.

Aim of this thesis

My project is based on three fundamental experimental observations. (1) In mouse V1, locomotion drives activity and a subset of neurons signal mismatch between visual flow speed and running speed (Keller et al., 2012; Saleem et al., 2013). (2) Responses to visual input are dependent on the experience of visual stimuli during a critical period in development (Blakemore and Cooper, 1970; Hirsch and Spinelli, 1970; Hubel and Wiesel, 1970). (3) From classical work in experimental psychology, we know that visual input alone is not sufficient to establish visuomotor behavior (Hein and Held, 1967; Held and Hein, 1963). Predictive processing offers a framework in which these effects can be interpreted. The aim of this thesis was to further expand our understanding of visuomotor integration within the framework of predictive processing.

In chapter 1, I demonstrate how visuomotor integration and mismatch responses depend on the coupling of visual feedback and locomotion during development. Additionally, I asked if the signatures of the core computation of predictive processing—the subtraction of sensory input and prediction—can be found in the activity of genetically identified neurons in V1. This work was done in collaboration with Bo Wang.

Chapter 2 then addresses the origin of prediction signals in V1. If mismatch responses in V1 are indeed the result of a comparison of the feedforward sensory input and the predictions thereof, where do these predictions originate? In hierarchical predictive-processing models, predictions are signaled by feedback connections from higher to lower levels. Consistent with this idea, in a project led by Marcus Leinweber, we found that prediction signals in V1 are delivered by feedback connections from a secondary motor area.

CHAPTER 1: VISUOMOTOR COUPLING SHAPES THE FUNCTIONAL DEVELOPMENT OF MOUSE VISUAL CORTEX

This chapter is based on a paper that has been published in Cell (Attinger et al., 2017). The text and figures of this chapter correspond largely to the submitted manuscript, with minor adaptations to formatting and numbering to conform to the style of this thesis.

Alexander Attinger^{*1,2}, Bo Wang^{*1,2} & Georg B. Keller^{1,2,3}

* These authors contributed equally to this work.

¹Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, CH-4058 Basel, Switzerland

²Faculty of Natural Sciences, University of Basel, Klingelbergstrasse 50/70, CH-4056 Basel, Switzerland

³Lead contact, corresponding author: georg.keller@fmi.ch

Summary

The emergence of sensory guided behavior depends on sensorimotor coupling during development. How sensorimotor experience shapes neural processing is unclear. Here we show that the coupling between motor output and visual feedback is necessary for the functional development of visual processing in layer 2/3 (L2/3) of primary visual cortex (V1) of the mouse. Using a virtual reality system, we reared mice in conditions of normal or random visuomotor coupling. We recorded the activity of identified excitatory and inhibitory L2/3 neurons in response to transient visuomotor mismatches in both groups of mice. Mismatch responses in excitatory neurons were strongly experience dependent and driven by a transient release from inhibition mediated by somatostatin-positive interneurons. These data are consistent with a model in which L2/3 of V1 computes a difference between an inhibitory visual input and an excitatory locomotion-related input, where the balance between these two inputs is finely tuned by visuomotor experience.

Introduction

Sensory feedback is inherently coupled to movement, and sensorimotor coupling is necessary for both the development (Hein and Held, 1967; Held and Hein, 1963) and the maintenance (Leonardo and Konishi, 1999; Nordeen and Nordeen, 1992) of sensory guided behaviors. In classical experiments, Held and Hein demonstrated that cats reared with normal visual experience but without visuomotor coupling fail to perform simple visually guided behaviors (Held and Hein, 1963). This behavioral impairment is restricted to the movements that are decoupled from sensory feedback during development (Hein and Held, 1967). Thus, sensory guided behaviors rely on a mechanism to integrate sensory input and motor output that is instructed by experience. It is still unclear what the neural circuits are that underlie this type of sensorimotor integration and how they are shaped by sensorimotor experience during development.

Visual responses in primary visual cortex (V1) are known to depend on visual experience during development (Blakemore and Cooper, 1970; Hirsch and Spinelli, 1970; Hubel and Wiesel, 1970). In anesthetized or immobile animals, neural activity in V1 is known to closely reflect visual stimuli presented to the animal (Hubel and Wiesel, 1962; Niell and Stryker, 2008). Based on this, activity in V1 is classically interpreted in a representational framework (Marr, 1982), where neural responses are described in terms of receptive fields, and signal the presence of a specific visual stimulus in the environment. However, it is becoming increasingly clear that this interpretation of the function of visual cortex is incomplete. In monkeys freely moving their eyes, response patterns of neurons in V1 give surprisingly poor reflections of what an animal is viewing (Livingstone et al., 1996). One possible cause for this are motor-related signals. Self-generated locomotion has been shown to modulate visual responses (Fu et al., 2014; Niell and Stryker, 2010; Polack et al., 2013), and to even drive activity in V1 independent of visual input (Keller et al., 2012; Saleem et al., 2013). Thus, activity in V1 cannot be explained by visual input alone and is likely the result of an integration of sensory and motor-related signals.

An alternate framework within which the activity in visual cortex can be explained is that of predictive coding. It posits that the brain continuously predicts sensory feedback based on an internal model of the environment (Friston, 2005; Gregory, 1980; Rao and Ballard, 1999; Wolpert et al., 1995). Evidence for this interpretation comes from the finding that a subset of neurons in V1 selectively responds to a mismatch between predicted and actual visual feedback (Keller et al., 2012; Saleem et al., 2013). Similar feedback mismatch responses have also been described in primate primary auditory cortex (Eliades and Wang, 2008) and primary auditory pallium of songbirds (Keller and Hahnloser, 2008). If such feedback mismatch responses signal a deviation from a prediction that

is based on a learned relationship between motor output and sensory feedback, then they should depend on sensorimotor experience.

To test this, we reared mice in a virtual reality system either under coupled or non-coupled (yoked) visuomotor conditions and subsequently probed neural activity in layer 2/3 of V1. We found that responses to a mismatch between actual and expected visual input only occurred in mice that experienced normal visuomotor coupling. Using a simple model, in which an excitatory neuron computes a difference between an inhibitory visual input and an excitatory prediction of visual input, we show that mismatch responses can be explained by a relief from visually driven inhibition. By recording the activity of genetically identified interneurons in visual cortex, we show that this visual inhibition is likely mediated by somatostatin (SST) interneurons. Finally, we show that normal visuomotor experience restores sensorimotor integration. Together, our data are consistent with a predictive-coding interpretation of the function of visual cortex, where the balance between feed-forward and top-down input underlying the computation of visuomotor mismatch is finely tuned by visuomotor experience. In this way, visuomotor experience fundamentally shapes the functional development of visual processing in primary visual cortex.

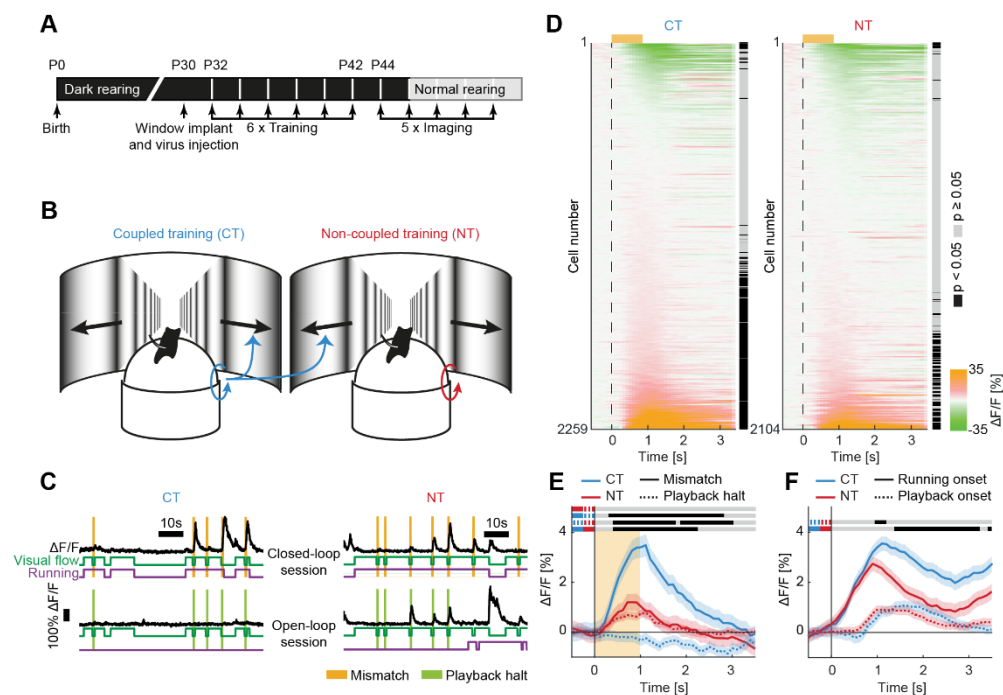


Figure 2.1. Mismatch responses in excitatory neurons depend on visuomotor experience.

(A) Experimental timeline. Mice were dark-reared from birth. AAV injection and imaging window implantation occurred on postnatal day 30 (P30). From P32 to P42, mice had 6 training sessions in either coupled (coupled trained: CT), non-coupled (non-coupled trained: NT), or dark (dark trained: DT) conditions, followed by 2 to 5 imaging sessions beginning at P44 and spaced by 2 days. Some of the mice were put on a normal 12 h/12 h light/dark cycle after the second imaging session. **(B)** Schematic of the training setup. Mice were trained in pairs; visual flow (black arrows) on both training setups was coupled

to the locomotion of the CT mouse (blue arrows). The NT mouse was free to run but had no influence on the visual flow it was seeing. **(C)** Sample fluorescence traces ($\Delta F/F$, black lines) of an excitatory neuron in a CT (left) and a NT (right) mouse, during a closed-loop (top traces) and an open-loop session (open-loop sessions consisted of a replay of the visual flow generated during the preceding closed-loop session, bottom traces). Vertical bars indicate mismatch (orange) and playback halt (green) events. Binarized visual flow (green) and running speed (purple) are indicated below the fluorescence traces. In CT mice, we found neurons that selectively respond to mismatch, whereas in NT mice, neurons that responded to mismatch also responded to corresponding playback halts in open-loop sessions. Note that all data presented in this and the following panels are from the first imaging day. **(D)** Average mismatch response ($\Delta F/F$) of all neurons in CT mice (left, 9 mice, 2259 neurons) and NT mice (right, 9 mice, 2104 neurons), sorted by amplitude of mismatch response. Black and grey shading to the right indicates significance of responses (grey: $p \geq 0.05$, black: $p < 0.05$, Mann-Whitney-U test; see Experimental Procedures). Orange bar marks the duration of mismatch. In CT mice, the fraction of neurons with a significant mismatch response was larger than in NT mice (CT: $40\% \pm 5\%$; NT: $26\% \pm 5\%$, $p = 0.03$, Mann-Whitney-U test; see Experimental Procedures). **(E)** The average population response ($\Delta F/F$) to mismatch (solid) was stronger in CT (blue) than in NT (red) mice. Population response to playback halt was negligible in CT mice, but was as large as the mismatch response in NT mice (dashed lines). Orange area indicates duration of mismatch; shading indicates s.e.m. The data in the different curves are compared bin-by-bin (100 ms bins) using a Student's *t* test. Bins with a significant difference ($p < 0.01$) are marked by a black line above the curves; those without are marked as light gray (see Experimental Procedures). Each comparison is marked by a pair of line segments to the left, corresponding in color and line style to the data plotted, indicating which two curves are being compared. **(F)** Same as in (E), but for running onset in closed-loop sessions (solid lines) and playback onset in open-loop sessions (dashed lines, see Experimental Procedures). Shading indicates s.e.m.

Results

To experimentally control the visuomotor experience of mice, they were dark-reared from birth and only exposed to visual stimulation in 6 separate 2-hour training sessions spaced by 48 hours over the course of 12 days, starting on postnatal day 32 (**Figure 2.1A**). During these sessions each mouse was trained either in a coupled visuomotor condition (coupled trained: CT), in which the visual flow feedback was coupled to the locomotion of the mouse in a virtual environment, or in a non-coupled condition (non-coupled trained: NT) in which visual flow was independent of the mouse's locomotion (**Figure 2.1B**). Mice were head-fixed on a spherical treadmill (Dombeck et al., 2007) surrounded by a toroidal screen that provided visual flow feedback in the form of full-field vertical gratings on the walls of a virtual corridor. To match the visual experience of both groups, mice were trained in pairs (one CT and one NT mouse) in two separate virtual environments such that the locomotion of the CT mouse was used to control the visual flow of both virtual environments. In this way, both CT and NT mice experienced identical visual flow. Both groups of mice were exposed to light only in this virtual reality environment during the 6 training sessions, every other day for 2 hours, and were otherwise fully dark-reared. A third group of mice was reared and trained in complete darkness (dark trained, DT). After the 6 training sessions, we recorded neural activity in V1 of all 3 groups of mice by two-photon imaging of a genetically encoded calcium indicator GCaMP5 (Akerboom et al., 2012) or GCaMP6f (Chen et al., 2013) during different visual flow feedback conditions in 2 imaging sessions spaced by 2 days, starting on postnatal day 44. Subsequently, mice were exposed to a normal dark-light cycle and imaged for an additional 3 sessions, again spaced by 2 days (**Figure 2.1A**). Imaging sessions for all groups of mice consisted of 1 or 2 repetitions of approximately 8 minutes of locomotion coupled to visual flow feedback (closed-loop session) and 2 replays of the same visual flow patterns during an open-loop session to quantify visual responses. To

probe for feedback mismatch responses, we briefly halted visual flow for 1 second at random times during the closed-loop session (referred to as mismatch). Open-loop sessions consisted of a playback of the visual flow that the mouse had generated during the closed-loop session including brief visual flow halts, which we refer to as playback halts. Note that analysis of playback halts was restricted to times when the mouse was not running (see Experimental Procedures). Mice were free to run during the entire experiment, including open-loop sessions and did so spontaneously. In early sessions, mice that exhibited low locomotor activity were prompted to run using air-puffs to the neck. CT and NT mice exhibited similar locomotion behavior during both training and imaging sessions (**Figures S2.1A and S2.1B**).

Mismatch responses in excitatory neurons depend on visuomotor experience

To test if mismatch responses in layer 2/3 excitatory neurons in V1 depend on coupled sensorimotor experience, we expressed GCaMP5 in C57BL/6 mice (3 CT and 3 NT) and GCaMP6f in vGAT-Cre (Vong et al., 2011) x Ai14 (Madisen et al., 2010) mice (6 CT and 6 NT) using an AAV2/1-EF1 α -GCaMP vector (see Experimental Procedures). In vGAT-Cre x Ai14 mice, inhibitory neurons express the red fluorescent protein tdTomato, which allowed us to restrict analysis to identified excitatory neurons. In these mice, we found that $96.8\% \pm 0.7\%$ (mean \pm s.e.m.) of GCaMP6f labelled neurons were excitatory (**Figures S2.1C and S2.1D**). Thus, for all following analysis we pooled putative excitatory neurons of the C57BL/6 mice and the identified excitatory neurons of the vGAT-Cre x Ai14 mice. In total we recorded from 2259 excitatory neurons in CT mice (996 putative excitatory and 1263 identified excitatory neurons) and 2104 excitatory neurons in NT mice (764 putative excitatory and 1340 identified excitatory neurons).

We found that in CT mice, a considerable fraction of excitatory neurons responded to mismatch (865 of 2259 neurons or 38.3%; **Figures 2.1C and 2.1D**) resulting in a large population mismatch response (**Figure 2.1E**). In CT mice, mismatch responses cannot be explained by visual input alone as there was no population response to playback halt (**Figure 2.1C and 2.1E**; note, mismatch and playback halt are identical visual stimuli). This is consistent with what we previously found in normally reared mice (Keller et al., 2012). In NT mice, the fraction of neurons that responded to mismatch was smaller (425 of 2104 neurons or 20.2%) and the population response to mismatch was weaker than in CT mice (**Figure 2.1E and S2.1E**). Interestingly, in NT mice the response to mismatch was of similar magnitude as the response to playback halt (**Figure 2.1E**) and individual neurons often responded to both mismatch and playback halt (**Figures S2.1F and S2.1G**). With increasing mismatch response, neurons in CT, but not NT, mice became increasingly selective for mismatch (**Figure S2.1H**). Thus, whereas in CT mice, mismatch responses were strongly dependent on motor-related inputs, mismatch responses in NT mice were only weakly modulated by motor-related signals. In both CT

and NT mice, the response reliability of mismatch responsive neurons increased with average amplitude of the mismatch response. On average mismatch neurons responded to 37.5% of mismatches in CT mice and to 33.8% in NT mice (**Figure S2.1I**). A subset of neurons responded with a decrease in activity to mismatch as well as playback halts (**Figure 2.1D and S2.1J**). This type of response possibly reflects a visual response driven by visual flow: upon cessation of the visual flow, these neurons decrease their response.

The differences in mismatch responses between CT and NT mice could not be explained by differences in average visual or motor-related input to V1. Both the running-onset activity during the closed-loop session (referred to as running-onset response) as well as the visual flow onset responses during open-loop sessions (referred to as playback-onset response) were similar when comparing responses in CT and NT mice (**Figure 2.1F**). In dark trained mice, running-onset responses were normal, but mismatch and playback halt responses were smaller (**Figure S2.2**). This suggests that visual and motor-related inputs are maintained independently, and that visuomotor coupling is necessary for the development of normal integration of visual and motor-related inputs.

Mismatch responses can be explained as a difference between an excitatory prediction and an inhibitory visual input

Motor-related inputs have been shown to drive activity in mouse V1 (Keller et al., 2012; Saleem et al., 2013). One simple model to explain mismatch responses in a layer 2/3 excitatory neuron would be that such a neuron integrates an excitatory motor-related input, in this case a prediction of visual flow based on motor output, and an inhibitory input that conveys feed-forward visual flow input (**Figure 2.2A**). In this model, inhibitory and excitatory inputs are balanced when predictions match feed-forward input. At mismatch onset, a decrease in visual inhibition would then allow the excitatory motor-related input to activate the neuron. If this is correct, mismatch neurons should receive excitatory motor-related input and inhibitory visual input. To test this, we computed the correlation of the activity of each neuron with visual flow and with running speed during the open-loop sessions. As running and visual flow are independent in open-loop sessions, the activity of a neuron that receives net inhibitory visual input and net excitatory motor-related input would have a negative correlation with visual flow and a positive correlation with running speed and vice versa. Plotting the distribution of the correlations of all neurons revealed that neurons with a strong mismatch response, had a negative correlation with visual flow and a positive correlation with running speed, on average (**Figure 2.2B**). When comparing the entire population of neurons, we found that in CT mice, neurons with a positive correlation with running speed tended to have a negative correlation with visual flow, whereas in NT mice neurons with a positive correlation with running speed tended to also have a positive correlation with visual flow. We quantified this

interaction for every mouse as the angle (A) of the first principal component of the correlation scatter plot and found that, in CT mice, this angle was on average negative ($-41^\circ \pm 10^\circ$, mean \pm s.e.m.), whereas in NT mice it was on average positive ($9^\circ \pm 13^\circ$, mean \pm s.e.m.; **Figures 2.2B and 2.2C**). This suggests that visuomotor coupling establishes a balance between inhibition and excitation, such that those layer 2/3 excitatory neurons that are strongly activated by running also are also strongly inhibited by visual flow.

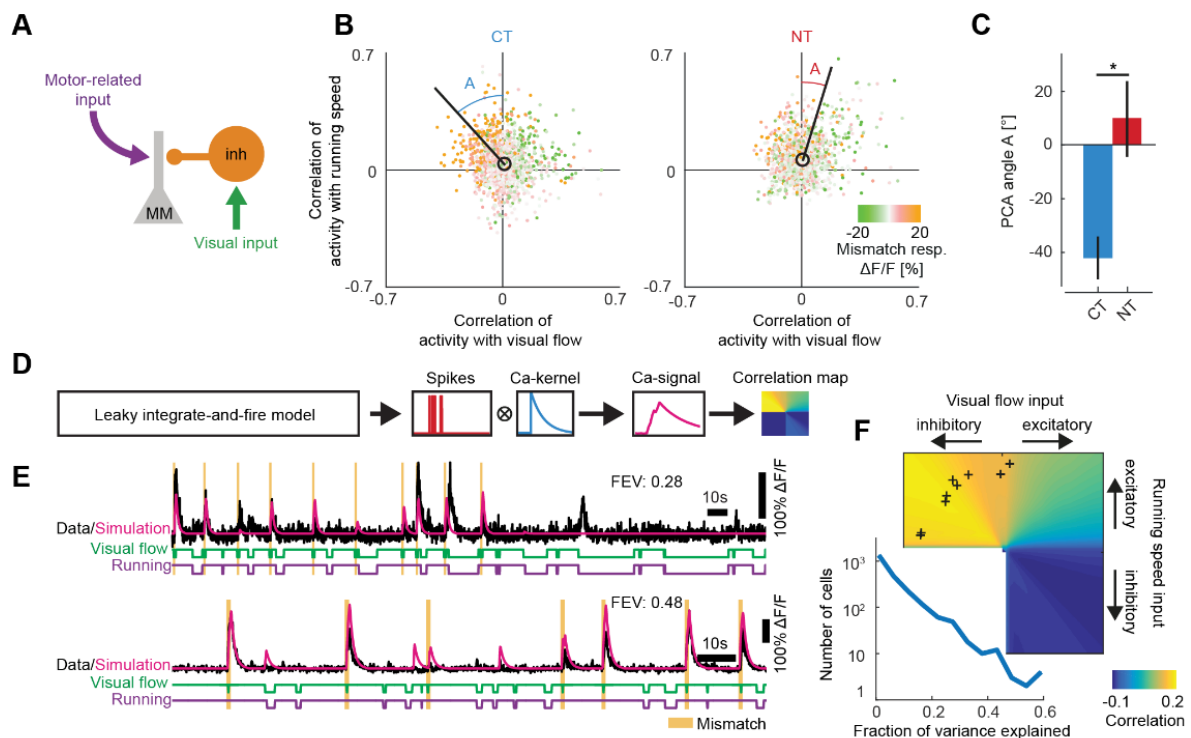


Figure 2.2. Mismatch responses can be explained as a difference between an excitatory motor-related input and an inhibitory visual input.

(A) Circuit model in which an excitatory mismatch neuron (MM, grey) integrates excitatory motor-related input and inhibitory visual input relayed by a local inhibitory interneuron (orange) to compute the difference between predicted and actual visual flow. **(B)** Correlation coefficients between neural activity ($\Delta F/F$) of layer 2/3 excitatory neurons with running speed and with visual flow in CT (left; 9 mice) and NT (right; 9 mice) mice during open-loop sessions. Each dot represents a single neuron (CT: 2259 neurons; NT: 2104 neurons). Dot color indicates the amplitude of the mismatch response. Black circles indicate the mean correlation values. The angle A indicated by the solid black line is the average angle between the first principle component of the distribution and the y-axis (see Experimental Procedures). Note that all data presented in this and the following panels are from the first imaging day. **(C)** Mean angle of the first principle component relative to the y-axis of the distribution of correlation coefficients as in **(B)** for CT ($n = 9$) and NT mice ($n = 9$). Error bars indicate s.e.m., Mann-Whitney-U test, $p = 0.04$. **(D)** Spiking output of a simple conductance-based leaky integrate-and-fire neuron (cLIF) was convolved with a unitary calcium-kernel to simulate neuronal activity during closed-loop and open-loop sessions. Excitatory and inhibitory inputs were approximated by running speed (αR) and visual flow (βV), e.g. for scaling factors $\alpha > 0$ and $\beta < 0$, excitatory input is proportional to running speed and inhibitory input is proportional to visual flow. By varying α and β systematically, we calculated correlation maps with data from open-loop sessions. The scaling factors maximizing the correlation map were used to simulate activity during closed-loop sessions (**Figure 2.2E**). **(E)** Sample fluorescence ($\Delta F/F$) traces of two neurons responding to mismatch from 2 mice during a closed-loop session (black traces) and the corresponding simulated traces (pink traces). Note that the simulation parameters are based on optimization during open-loop sessions. Running speed, visual flow and mismatch are labeled as in (**Figure 1.1C**). Also shown is the FEV: fraction of explained variance (see Experimental Procedures) for each example neuron. **(F)** Lower left: Distribution of the fraction of variance explained (2259 neurons), estimated as the squared correlation (R^2) coefficient between model output and calcium activity during the closed-loop session. Top right: Average correlation map and average location of maxima (black crosses) for neurons with significant positive responses to mismatch, averaged over mice (9 mice, top 50% of significant neurons per mouse).

To test this model further, we implemented a conductance based leaky-integrate-and-fire (LIF) model (Salinas and Sejnowski, 2001) with 2 free parameters: A scaling factor for the running-related input (α) and a scaling factor for the visual input (β), which were used to modulate the excitatory and inhibitory conductances. The spiking output of the LIF model was convolved with a calcium kernel to generate a simulated calcium response (**Figures 2.2D and 2.2E**; see Experimental Procedures). Using data from open-loop sessions, we optimized the correlation between the model output and neural activity with a grid search over α and β for every excitatory neuron. We then predicted the activity of each excitatory neuron during the closed-loop session by using visual flow and running speed of that session as inputs to the LIF model optimized for the particular neuron (**Figure 2.2E**). We found that the average fraction of explained variance, estimated by a cross validation approach on the open-loop session data (see Experimental Procedures), was twice as large when using a model based on visual flow and running speed as when using a model based on just visual flow or just running speed (full model $R^2 = 0.06$; just visual flow $R^2 = 0.02$; just running speed $R^2 = 0.03$; $p < 0.01$ for both comparisons, Mann-Whitney U test; **Figure 2.2F**). We then averaged the correlation maps generated by the grid search (see Experimental Procedures) for excitatory neurons with a significantly positive response to mismatch and found that activity of these neurons could be best approximated when the motor-related conductance is positive ($\alpha > 0$) and the visual conductance is negative ($\beta < 0$) (**Figure 2.2F**). This shows that mismatch responses in excitatory neurons can be explained by a combination of an excitatory motor-related input and inhibition by visual flow. Consistent with a visually driven inhibition of mismatch neurons, mismatch responsive neurons exhibited a decrease of activity in response to the onset of visual flow in open-loop conditions (**Figures S2.3A – S2.3C**).

Somatostatin interneurons decrease activity during mismatch

As most long-range inputs to V1 are excitatory, feed-forward visual inhibition would need to be relayed by local inhibitory neurons. These neurons would have to be strongly driven by visual flow and, as a consequence, decrease activity in response to a brief stop in visual flow during mismatch and playback halt. To probe the responses of different inhibitory neuron subtypes, we repeated the training and imaging protocol using 4 different Cre driver lines to selectively express GCaMP6f (AAV2/1-EF1 α -DIO-GCaMP6f-WPRE) in somatostatin (SST) (Taniguchi et al., 2011), vasoactive intestinal polypeptide (VIP) (Taniguchi et al., 2011), parvalbumin (PV) (Hippenmeyer et al., 2005), or neuropeptide-Y (NPY) (Gong et al., 2007) interneurons. The SST-Cre, VIP-Cre and PV-Cre lines collectively target approximately 80% of interneurons in mouse V1 and the labelled populations are largely non-overlapping (Pfeffer et al., 2013).

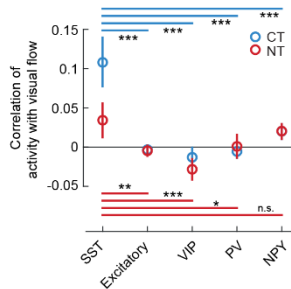


Figure 2.3. SST interneurons are strongly driven by visual flow.

Average correlation of neural activity with visual flow during open-loop sessions for excitatory neurons (average correlations: CT: 0.00, NT: -0.01) and SST (CT: 0.13, NT: 0.04), VIP (CT: -0.01, NT: -0.03), PV (CT: -0.01, NT: 0.00) and NPY (CT: 0.02, NT: 0.02) interneurons in CT and NT mice. Average correlation of activity with visual flow was highest for SST interneurons. Error bars indicates s.e.m. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, n.s., not significant, $p \geq 0.05$, Student's t test.

We found SST interneurons exhibited a higher correlation with visual flow than other interneuron subtypes or excitatory neurons (**Figure 2.3**). Moreover, of the 4 interneuron subtypes, only SST interneurons responded, on average, with a drop in activity to a brief stop in visual flow both during mismatch and playback halt (**Figure 2.4A**). Notably, this decrease in average activity on visual flow halt was independent of visuomotor experience, as it was present in both CT (5 mice, 118 neurons) and NT mice (5 mice, 157 neurons), indicating that the visual input onto SST neurons is established independently of motor-related input. Locomotion strongly increased visual responses in SST interneurons (**Figure 2.4B**), but running-onset responses were almost completely absent in darkness (**Figure S2.3D**), consistent with a predominantly visual drive to SST interneurons. Overall, the responses of SST interneurons to mismatch were diverse (**Figure S2.3E**), indicating that SST expression does not mark one homogenous functional class of interneurons.

The responses of VIP interneurons were independent of visuomotor experience. In both CT and NT mice, they responded with an increase of activity to mismatch but not to playback halt (**Figure 2.4C**; CT: 3 mice, 189 neurons; NT: 3 mice, 137 neurons). Given that VIP interneurons receive direct inhibitory input from SST interneurons (Pfeffer et al., 2013), mismatch responses may result from the combination of a running-related excitatory input to VIP interneurons (Fu et al., 2014) and a relief from SST interneuron mediated inhibition. Interestingly, running-related input to VIP interneurons was strongly experience dependent. VIP interneurons were only driven by running onset during closed-loop sessions in CT but not in NT mice (**Figure 2.4D**). Consistent with the strong reduction of running-onset responses in SST interneurons in darkness, a running-related input to VIP interneurons in NT mice was unmasked in darkness (**Figure S2.3F**). Taken together, our findings suggest that the inhibitory connection from SST interneurons onto VIP interneurons is stronger in absence of visuomotor experience.

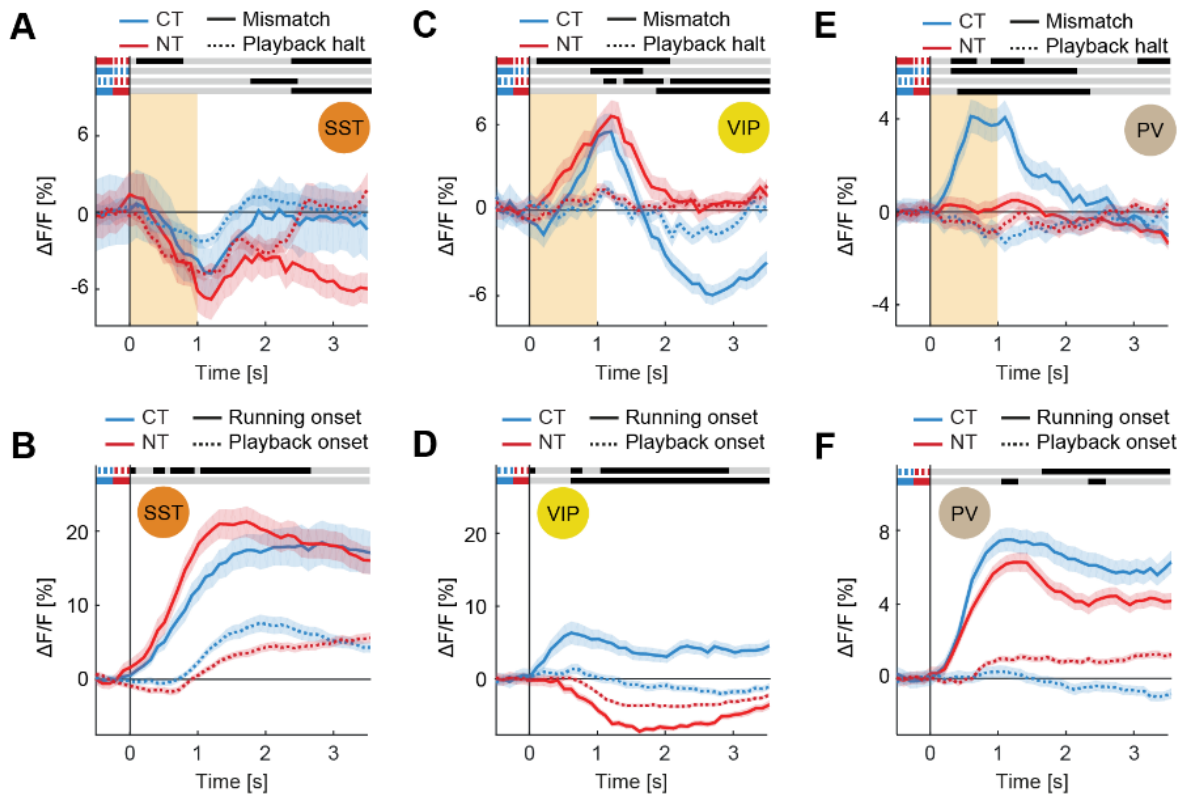


Figure 2.4. Experience dependent visuomotor integration in inhibitory interneurons.

(A) Average population responses to mismatch (solid line) and playback halt (dashed line) for SST interneurons from CT (blue, 5 mice, 118 neurons) and NT (red, 5 mice, 157 neurons) mice. For both CT and NT mice, SST interneurons responded with a decrease in activity to mismatch and playback halt. Orange area indicates duration of mismatch; shading indicates s.e.m. Note that all data presented in panels **A-F** are from the first imaging day. The data in the different curves are compared bin-by-bin (100 ms bins) using a Student's t test. Bins with a significant difference ($p < 0.01$) are marked by a black line above the curves; those without are marked as light gray (see Experimental Procedures). Each of the four comparisons is marked by a pair of line segments to the right, corresponding in color and line style to the data plotted, indicating which two curves are being compared. **(B)** Same as in **(A)**, but for running onset in closed-loop sessions (solid lines) and playback onset in open-loop sessions (dashed lines). **(C and D)**, Same as in **(A and B)**, but for VIP interneurons (CT: 3 mice, 189 neurons; NT: 3 mice, 137 neurons). VIP interneurons responded with an increase in activity independent of experience but did not respond to playback halt. **(E and F)** Same as in **(A and B)** but for PV interneurons (CT: 5 mice, 498 neurons; NT: 6 mice, 344 neurons). The mismatch response in PV interneurons was strongly experience dependent.

Lastly, responses in both PV interneurons (CT: 5 mice, 498 neurons; NT: 6 mice, 344 neurons) and NPY interneurons (CT: 3 mice, 189 neurons; NT: 3 mice, 137 neurons) were behavioral state and visuomotor-experience dependent. These two interneuron subtypes were activated by mismatch in CT mice, but unresponsive to mismatch in NT mice and unresponsive to playback halt in both CT and NT mice (**Figures 2.4E, S2.3G – S2.3I**). This highly selective response to mismatch in CT mice could be a direct consequence of the stronger activation of the excitatory neuron population in CT mice in response to mismatch (**Figure 2.1E**). Either excitatory neurons recruit PV and NPY interneurons only above a given activity level or the calcium dynamics in PV and NPY interneurons are such that we are unable to measure activity changes below a given threshold. Note however, that such a simple measurement threshold cannot account for the observation that in CT mice the population response

of excitatory neurons to running onset is smaller than that to mismatch (**Figures 2.1E, 2.1F**), but the running-onset response of PV interneurons is larger than that to mismatch (**Figures 2.4E, 2.4F**). One potential consequence of a selective activation of PV interneurons in CT mice is that the PV activation could lead to a response normalization in excitatory neurons (Wilson et al., 2012) that narrows the population response to mismatch. Consistent with this, we found that the distribution of mismatch responses is narrower in CT mice (**Figure S3J**). This narrowing of the distribution of mismatch responses could function to make mismatch responses more selective to one particular type of mismatch.

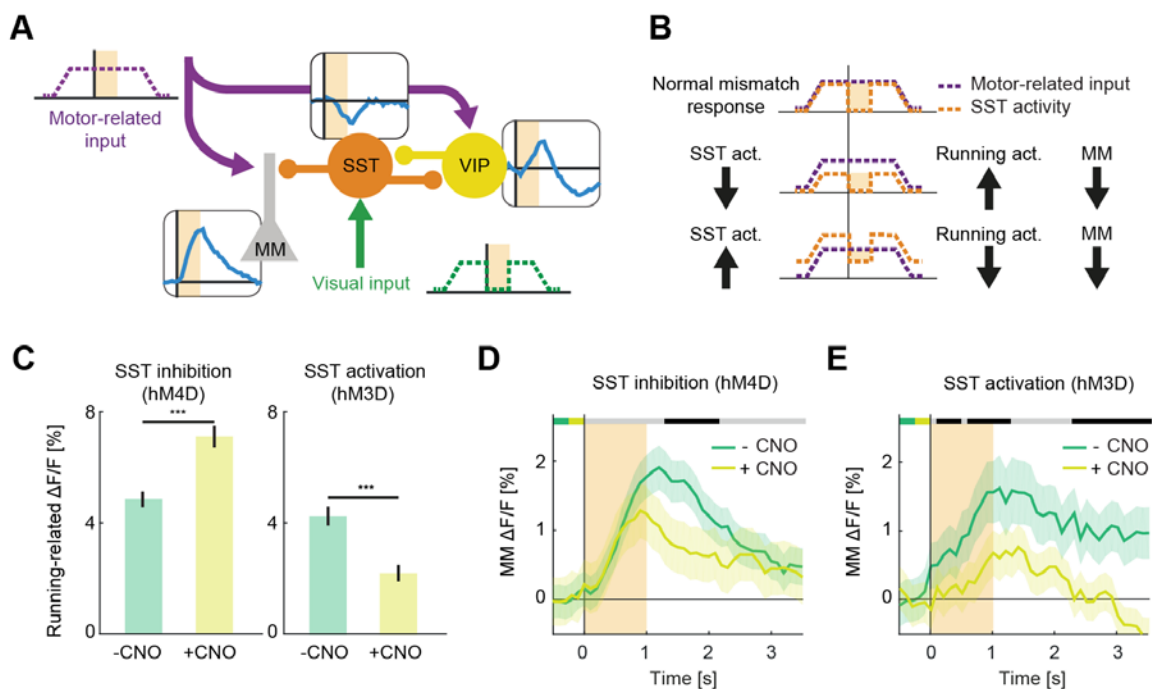


Figure 2.5. A drop in SST activity leads to a mismatch response in excitatory neurons.

(A) Schematic of a model circuit describing the computation of mismatch responses in layer 2/3 of V1. Excitatory neurons and VIP interneurons receive excitatory motor-related input (purple arrow; dashed purple line depicts idealized running profile around a mismatch, indicated by orange shading). SST interneurons receive visual flow input (green arrow; dashed green line depicts idealized visual flow around a mismatch, indicated by orange shading). Blue lines next to neurons depict average mismatch responses of excitatory neurons (**Figure 2.1E**), SST (**Figure 2.4A**) and VIP (**Figure 2.4C**) interneurons from CT mice. During mismatch, visual flow is halted and the activity of SST interneurons decreases, thereby disinhibiting the apical dendrites of mismatch neurons and allowing the excitatory motor-related input to activate the neuron. VIP interneurons amplify this effect by further suppressing SST interneuron activity. **(B)** Predicted effects of pharmacogenetic manipulation of SST interneurons on excitatory neurons. Idealized activity profiles of excitatory motor-related activity (purple line) and SST interneuron activity for a short period of running during a closed-loop session including a mismatch (onset marked by vertical line). In normal conditions (top), SST interneuron activity balances the motor-related input and the mismatch response of excitatory neurons is maximal (mismatch-triggered difference between excitatory and inhibitory input, orange shading). Inhibition of SST interneurons (middle) should result in a smaller mismatch-induced difference in inhibition and therefore a smaller mismatch response as well as increased running-related activity. Excitation of SST interneurons (bottom) should also result in smaller mismatch responses due to an over-inhibition of excitatory neurons, but decreased running-related activity. **(C)** Mean running related activity before and 30 min after injection of DREADD activator CNO (5 mg/kg i.p.) in mice expressing an inhibitory (left; 829 neurons, *** $p < 0.001$, Wilcoxon signed-rank test) or an excitatory (right; 411 neurons, *** $p < 0.001$, Wilcoxon signed-rank test) DREADD in SST interneurons. **(D)** Average population mismatch responses of excitatory neurons before (green trace) and 30 min after (yellow trace) the injection of CNO in mice expressing an inhibitory DREADD in SST interneurons (4 mice, 829 neurons). Orange bar indicates duration of mismatch, shading indicates s.e.m. Statistical comparisons as in **Figure 2.1E**. **(E)** Same as in **(D)**, but for mice expressing an excitatory DREADD in SST interneurons (2 mice, 411 neurons).

Our data indicate that layer 2/3 excitatory mismatch neurons and a subset of VIP interneurons receive excitatory, motor-related input, while a subset of SST interneurons is more strongly driven by visual flow. Consistent with the finding that SST interneurons receive strong input from surrounding excitatory neurons (Adesnik et al., 2012; Fino and Yuste, 2011; Jiang et al., 2015), we found that excitatory neurons whose activity correlates positively with visual flow (CT: 24% or 539 of 2259 of neurons; NT: 24% or 513 of 2104 neurons) exhibit a decrease in activity on mismatch similar to SST interneurons (**Figure S2.1J**). Based on the connectivity motif of excitatory neurons, SST and VIP interneurons (Pfeffer et al., 2013; Pi et al., 2013), we propose a schematic model circuit to explain mismatch responses in layer 2/3 excitatory neurons (**Figure 2.5A**). SST interneurons target the apical dendrites of layer 2/3 excitatory neurons (Markram et al., 2004). A reduction of visual input onto SST interneurons during mismatch thus relieves the apical dendrite of inhibition, and would allow excitatory motor-related input to activate the neuron. Based on this model we predict that both SST interneuron activation and inhibition should lead to a decrease of the mismatch response in excitatory neurons, but should have opposing effects on running-related activity in excitatory neurons (**Figure 2.5B**). To test this, we pharmacogenetically manipulated the activity of SST interneurons using DREADDs (Armbruster et al., 2007). We injected either AAV-EF1 α -DIO-hM4D(Gi)-mCherry or AAV-EF1 α -DIO-hM3D(Gq)-mCherry into V1 of normally reared SST-Cre mice. In addition, we unconditionally transfected neurons with GCaMP6f to record mismatch and running related activity in putative excitatory neurons. Note that in these experiments we cannot exclude the possibility that some of these putative excitatory neurons are non-SST interneurons. We found that DREADD inhibition of SST interneurons led to an increase in running-related activity in excitatory neurons, while DREADD activation of SST interneurons led to a decrease in running-related activity (**Figure 2.5C**). In addition, both inhibition and activation of SST interneurons led to a decrease in the mismatch response of excitatory neurons (**Figures 2.5D and 2.5E**). These results are consistent with a model of mismatch computation in which mismatch responses in layer 2/3 neurons are the result of a relief of SST interneuron mediated inhibition. To test the effect of a transient manipulation of SST and VIP activity on mismatch responses we injected AAV-EF1 α -GCaMP6f and either AAV-EF1 α -DIO-ChrimsonR-tdTomato (Klapoetke et al., 2014) or AAV-CAG-FLEX-ArchT-tdTomato (Han et al., 2011) into V1 of normally reared SST-Cre mice and VIP-Cre mice. We then identified putative excitatory mismatch neurons based on their responses to mismatch events in closed-loop sessions (in the following simply referred to as mismatch neurons) and measured the responses of these neurons to brief (1 s) activation or inhibition of SST or VIP interneurons (see Experimental Procedures). We found that activation of SST interneurons resulted in an inhibition of mismatch neurons that was strong enough to fully suppress mismatch responses in mismatch

neurons when SST interneurons were activated concurrently with a mismatch event (**Figure 2.6A**). Consistent with this, inhibition of SST neurons resulted in an activation of mismatch neurons and concurrent inhibition of SST interneurons with a mismatch event resulted in increased mismatch responses (**Figure 2.6B**). Conversely, activation of VIP interneurons resulted in an activation of mismatch neurons and an increase of the mismatch response when VIP interneurons were activated concurrently with a mismatch event (**Figure 2.6C**). Finally, inhibition of VIP interneurons resulted in an inhibition of mismatch neurons that was strong enough to suppress mismatch responses (**Figure 2.6D**). Note that even though these effects were stronger for mismatch neurons than for putative excitatory neurons that did not respond to mismatch (**Figure S2.4**), it is very likely only a subset of SST and VIP interneurons that are part of the circuit involved in mismatch responses in excitatory neurons. In summary, these results are consistent with the classical cortical SST-VIP disinhibitory circuit (Pfeffer et al., 2013; Pi et al., 2013), and suggest that this circuit plays a central role in mismatch computation with mismatch neurons under inhibitory control of SST interneurons. Thus, the relief of SST-mediated visual inhibition combined with a top-down motor-related excitatory drive can account for visuomotor mismatch responses in layer 2/3 excitatory neurons.

To test if both CT and NT mice learn to perform visuomotor tasks after exposure to visuomotor coupling, we repeated the training protocol with a separate cohort of mice. Instead of going through the imaging paradigm after coupled or non-coupled training, these mice were trained either to navigate a 2-dimensional (2D) virtual environment or to detect mismatch (see Experimental Procedures). Both CT and NT mice learned to perform the 2D virtual locomotion task over the course of 6 training sessions of 1 hour each (**Figures S2.5A, S2.5B**). Also, both CT and NT mice learned to report the occurrence of mismatch over the course of 3-5 training sessions of 1 hour each (**Figure S2.5C**). These findings suggest that visuomotor coupling rapidly establishes normal visuomotor processing even after prolonged absence of coupling in NT mice.

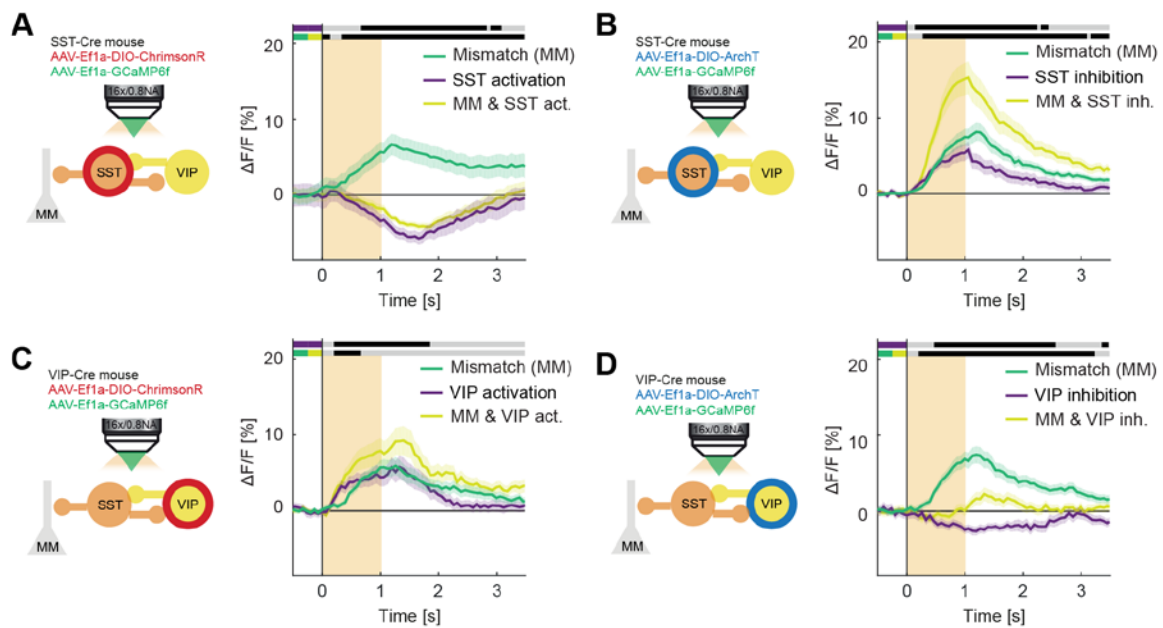


Figure 2.6. Mismatch neurons are inhibited by SST activation or VIP inhibition and activated by SST inhibition or VIP activation.

(A) Left: Schematic of the experimental design. ChrImsonR was selectively expressed in SST interneurons and GCaMP6f in all neurons. We then locally activated SST interneurons through the imaging objective while imaging GCaMP6f activity in all neurons. Right: Response of putative excitatory mismatch-responsive neurons (165 neurons, 5 mice) to mismatch (green line), optogenetic activation of SST interneurons during running (purple line), and concurrent mismatch and optogenetic activation of SST interneurons (yellow line). Orange area indicates duration of mismatch and duration of optogenetic stimulation respectively, shading indicates s.e.m. Statistical comparisons as in **Figure 2.1E**, but for 67 ms bins. Upper line marks comparison of manipulation-only against baseline, lower line marks comparison of mismatch only against concurrent mismatch and optogenetic stimulation. **(B)** Left: As in **(A)**, but expressing ArchT in SST interneurons. Right: Responses of mismatch neurons (236 neurons, 4 mice) as in **(A)**, but for optogenetic inhibition of SST interneurons. **(C)** Left: As in **(A)**, but expressing ChrImsonR in VIP interneurons. Right: Responses of mismatch neurons (114 neurons, 4 mice) as in **(A)**, but for optogenetic activation of VIP interneurons. **(D)** Left: As in **(A)**, but expressing ArchT in VIP interneurons. Right: Responses of mismatch neurons (107 neurons, 3 mice) as in **(A)**, but for optogenetic inhibition of VIP interneurons.

Normal visuomotor experience restores normal visuomotor integration in V1

Given that both CT and NT mice learned to perform visuomotor tasks over the course of a few days, visuomotor coupling should rapidly restore normal visuomotor processing in V1. To quantify the change in neural processing in V1 with the exposure to visuomotor coupling, we measured mismatch responses in both CT and NT mice over the course of 8 days following restoration of visuomotor coupling (exposure to both open-loop and closed-loop conditions and normal visuomotor experience with the transfer to rearing in a normal light/dark cycle; **Figure 2.1A**). We found that mismatch responses of excitatory neurons in CT and NT mice equalized rapidly with normal visuomotor experience (**Figures 2.7A – 2.7C**). The population mismatch responses of SST and VIP interneurons remained stable throughout the course of the experiment for both CT and NT mice (**Figures 2.7D and 2.7E**). This is consistent with the idea that the mismatch response of VIP and SST interneurons developed independent of visuomotor coupling. Similar to excitatory neurons, mismatch responses

in PV and NPY interneurons equalized after restoration of normal visuomotor coupling (**Figures 2.7F and S2.6A**). Interestingly, we found not only an increase of mismatch responses in NT mice with exposure to closed-loop sessions and normal visuomotor experience, but also a decrease of mismatch responses in CT mice with exposure to open-loop sessions and normal visuomotor experience. Similarly, we found that for the distribution of visual flow and running speed correlations, the angle of the first principle component (**Figures 2.2B and 2.2C**) equalized and approached zero for both CT and NT mice (**Figures 2.7G and S2.6B**). To quantify the behavioral response to mismatch on a timescale similar to that of the equalization of neural dynamics, we measured pupil dilation in response to mismatch. Mice exhibited a small but measurable pupil dilation response with a delay of approximately 450 ms after the neural response to a mismatch (400 ms for CT, 500 ms for NT; see Experimental Procedures; **Figure 2.7H**). This pupil dilation response was larger in CT mice and may reflect a startle response. The pupil dilation response also equalized with restoration of normal visuomotor experience, with the same time course as neural activity (**Figure 2.7I**). Altogether, these results suggest that the artificial restriction of visuomotor coupling to only a subset of movements (forward locomotion and eye movements) leads to an overrepresentation of the visuomotor processing of these movements that needs to be unlearned for the restoration of normal visuomotor behavior. This is consistent with the finding that a lack of visuomotor coupling for a specific range of movements leads to behavioral impairments that are specific to those movements (Hein and Held, 1967).

Discussion

Here we have shown that the development of responses to a mismatch between predicted and actual visual feedback in mouse V1 critically depends on coupled visuomotor experience. These mismatch responses are thought to be the consequence of predictive coding strategies that involve a comparison of actual and predicted sensory feedback to compute a prediction error or feedback mismatch. In this framework, predictions of sensory feedback are based on an internal model of the environment. Deviations from predictions in the form of mismatch signals are then used to update the internal model (Bastos et al., 2012; Rao and Ballard, 1999). As a consequence, it is likely that predictions are systematically shaped by experience, and can adapt to changes in the coupling between motor output and sensory feedback.

The mismatch responses we describe here could be the result of a weak excitatory visual response to the playback halt stimulus that is amplified by a running-related input. As mismatch is generated simply by halting visual flow, this would mean that the visual feature driving mismatch responses is either the negative acceleration of visual flow, or simply a stationary grating viewed while running.

However, any model for mismatch responses based on an excitatory visual drive fails to explain why mismatch responses scale linearly with the difference between running speed and visual flow speed in open-loop sessions (Zmarz and Keller, 2016). Additionally, a model for mismatch responses based on an excitatory visual input cannot explain why mismatch responses tend to decrease activity on playback onset (Figure S2.3A – S2.3C).

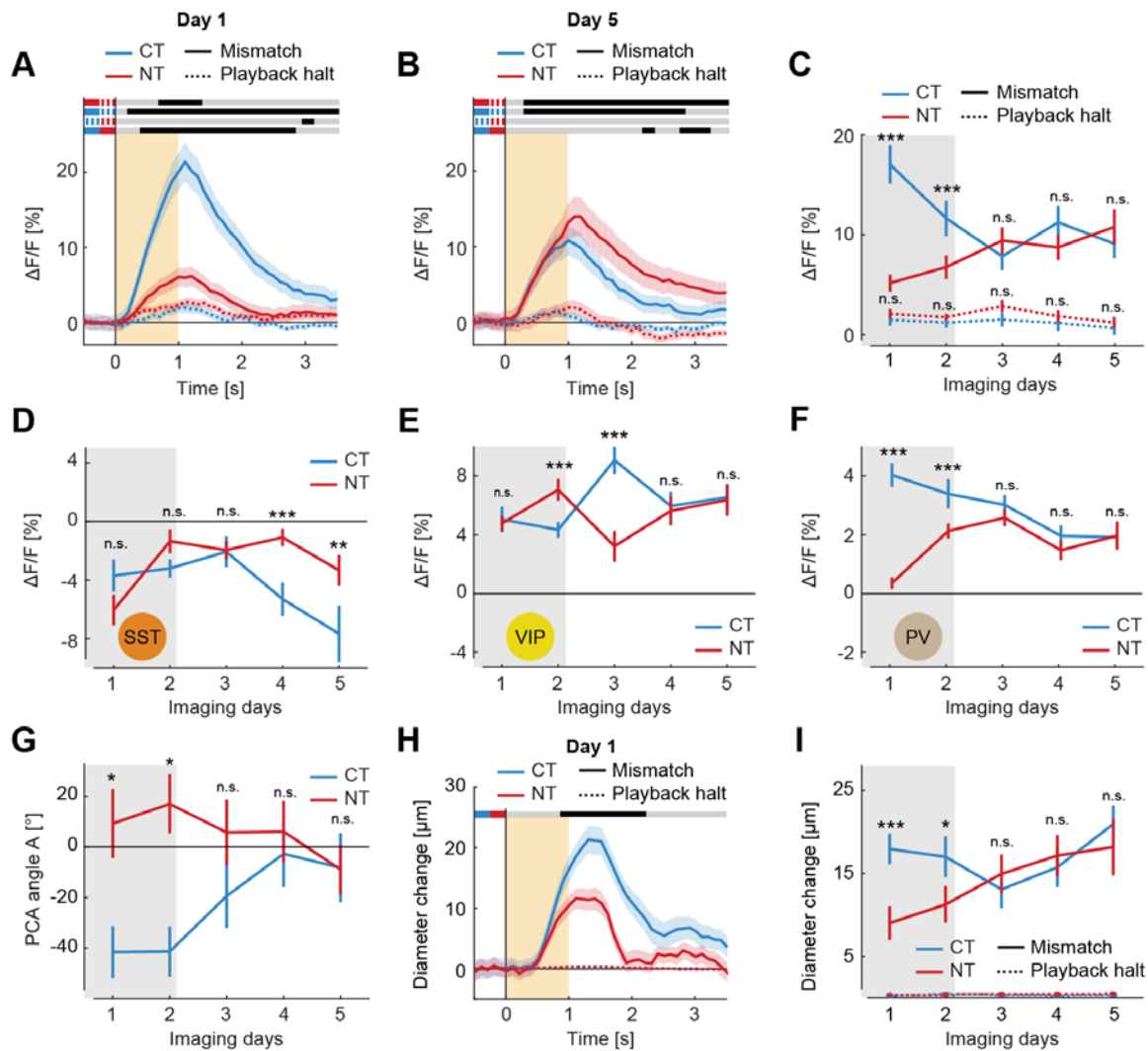


Figure 2.7. Normal visuomotor experience restores normal visuomotor integration.

(A) Average responses to mismatch (solid lines) and playback halt (dashed lines) of neurons with positive correlation of activity with running speed (running correlation greater than 0.05) and negative correlation of activity with visual flow (visual correlation smaller than -0.05) on the first imaging day (CT: 12% \pm 2% of neurons per mouse, 9 mice; NT: 10% \pm 3%, 9 mice). Orange area indicates duration of mismatch, shading indicates s.e.m. Statistical comparison as in Figure 2.1E. (B) Same as (A), but for last imaging day (CT: 10% \pm 2% of neurons per mouse, 8 mice; NT: 9% \pm 1%, 7 mice). (C) Average responses to mismatch and playback halt (see Experimental Procedures) of excitatory neurons selected as in (A) as a function of imaging days for CT and NT mice. Mice were dark reared until the second imaging session (indicated by gray area). Error bars indicate s.e.m. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, n.s., not significant, $p \geq 0.05$, Mann-Whitney-U test. (D) Average population responses to mismatch of SST interneurons, as a function of imaging days for CT and NT mice (CT, 5 mice, 118 neurons; NT: 5 mice, 157 neurons). Statistical test as in (C). (E) As in (D) but for VIP interneurons (CT: 3 mice, 189 neurons; NT: 3 mice, 137 neurons). (F) As in (D) but for PV interneurons (CT: 5 mice, 498 neurons; NT: 6 mice, 344 neurons). (G) Mean angle of first principal component (as in Figures 2.2B and 2.2C; see Experimental Procedures) relative to the y-axis for CT and NT mice as a function of imaging days. Gray area indicates dark rearing, error bars indicate s.e.m. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, n.s., not significant, $p \geq 0.05$, Mann-Whitney-U test. (H), Average pupil dilation in response to mismatch and playback halt for CT (25 mice) and NT mice (25 mice; see Experimental Procedures) on the first imaging day. Orange area indicates duration of mismatch, shading indicates s.e.m. Statistical comparisons as in

Figure 2.1E, but for $p < 0.05$. **(I)**, Average pupil dilation in response to mismatch and playback halt a function of imaging days for CT and NT mice. Gray area indicates dark rearing, error bars indicate s.e.m. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, Mann-Whitney-U test.

Our results are consistent with a model in which sensorimotor mismatch signals are computed locally in layer 2/3 by a comparator circuit that is shaped by experience. In this circuit, inhibition by visual flow is balanced against an excitatory motor-related input in mismatch neurons. SST interneurons mediate the inhibition by visual flow, while mutual inhibition between VIP and SST interneurons (Pfeffer et al., 2013) acts to amplify the responses of SST interneurons. Even though the average response of SST interneurons to mismatch is a decrease in activity, this effect is carried by only a subset of SST interneurons. Moreover, although such a simplified model is sufficient to explain mismatch responses, the interactions between the different interneuron subtypes are in all likelihood much richer than schematically summarized here. PV interneurons, for example, could act to normalize the mismatch response in excitatory neurons (Hofer et al., 2011; Kerlin et al., 2010).

Given that SST interneurons provide visual inhibition to excitatory mismatch neurons and that mismatch responses in SST interneurons do not depend on visuomotor experience (**Figure 2.4A**), it is likely that visuomotor experience predominantly modifies the synaptic inputs onto the excitatory neuron. In this way, a balance of excitation and inhibition is established, possibly via mechanisms similar to those resulting in the establishment of the balance between feed-forward excitation and inhibition mediated by PV interneurons (Xue et al., 2014). Note that the type of mismatch neuron we describe here balances an excitatory top-down input against an inhibition by visual flow and is active when there is less visual input than predicted. Conversely, if a neuron balances an inhibitory top-down input against an excitatory feed-forward input, it would signal a visual input that is stronger than predicted. Such a neuron would have classic visual responses in a passively observing animal. Computationally these two circuits are symmetric and would merely signal different types of mismatch (**Figure S2.7**).

We speculate that the framework of predictive coding can be used to describe cortical processing of sensory feedback for every movement that results in a predictable change of sensory input. We propose that the comparison of sensory input with a top-down prediction may be a general principle of cortical function, where predictions from higher areas are continuously compared to signals from lower areas, and mismatches between the two are used to refine these predictions (Clark, 2013; Friston, 2010). It is intriguing to speculate that impairments in this comparison may underlie cortical dysfunctions where the balance between predictions and sensory input is systematically perturbed (Frith et al., 2000; Sinha et al., 2014).

Author Contributions

A.A. and B.W. designed and conducted experiments. A.A. designed model and simulations. A.A. and B.W. analyzed data. A.A., B.W. and G.B.K wrote the manuscript.

Acknowledgements

We thank Antonia Drinnenberg, Thomas Mrcic-Flogel, and Rainer Friedrich for comments on earlier versions of the manuscript. We thank Daniela Gerosa-Erni for production of the AAV vectors, and the members of the Keller lab for discussion, support, and comments on the manuscript. This work was supported by the Swiss National Science Foundation and the Novartis Research Foundation.

Supplementary Information

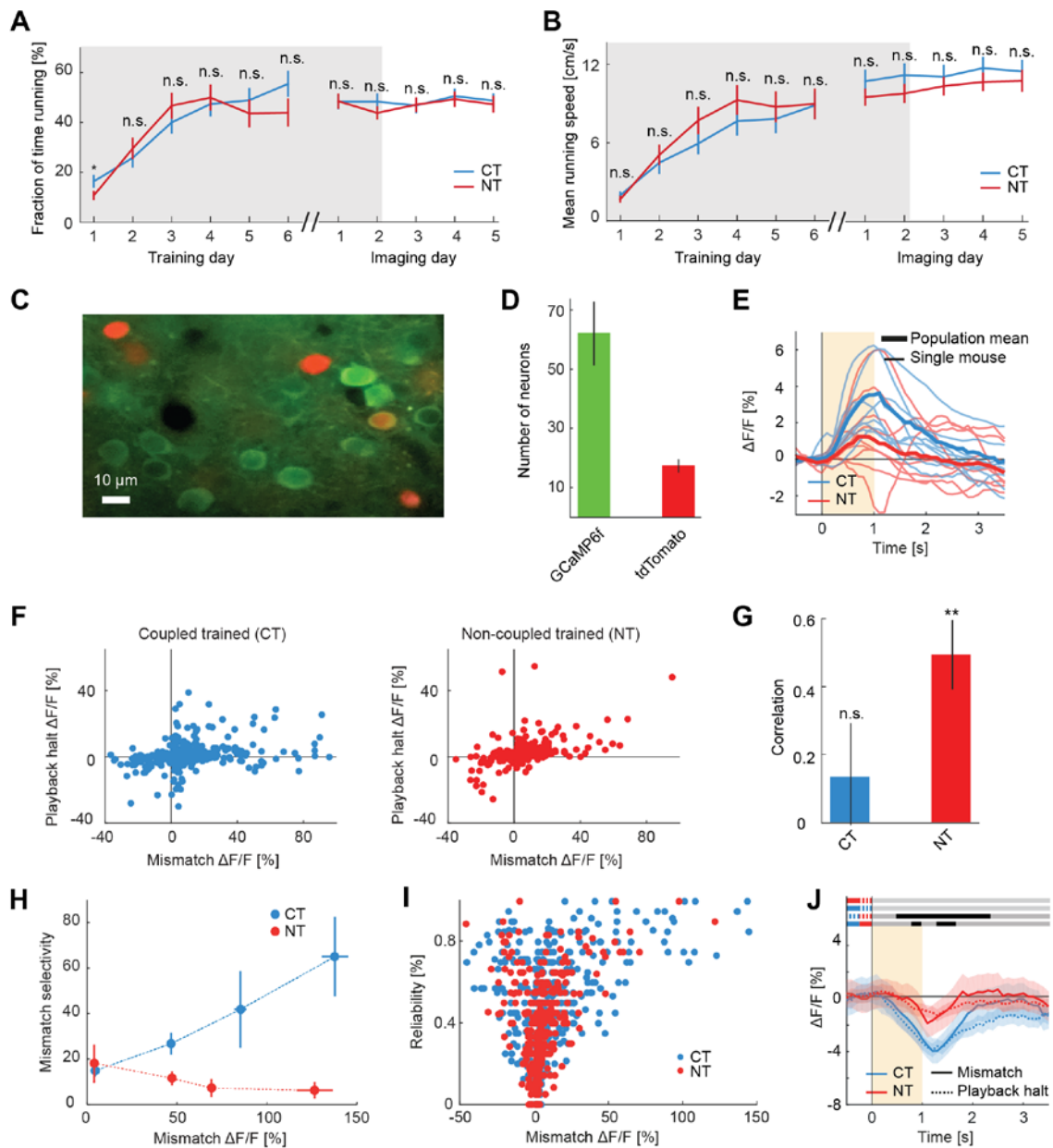


Figure S2.1. Analysis of running behavior, tdTomato expression, and additional analysis of mismatch and playback halt responses. Related to Figure 2.1.

(A) Fraction of time spent running increased during training period (left) for CT mice and NT mice and remained stable during imaging sessions (CT: 31 mice; NT: 30 mice). Error bars indicate s.e.m., gray shading indicates dark rearing. n.s., not significant, $p \geq 0.05$, Student's *t* test.

(B) Same as in **(A)**, but for average running speed. Average running speed increased during training (left) and remained stable during imaging sessions (right). Error bars indicate s.e.m., gray shading indicates dark rearing.

(C) Two-photon image of layer 2/3 neurons in mouse V1. In vGAT-Cre x Ai14 mice, inhibitory interneurons express tdTomato and are shown in red. Expression of GCaMP6f (green) under the EF1 α promoter resulted in $96.8\% \pm 0.7\%$ (3321 of 3438 in total, mean \pm s.e.m.) of GCaMP6f-positive neurons that were not tdTomato positive. The surprisingly small overlap between GCaMP6f and tdTomato expression may in part result from the selection bias of GCaMP6f-positive neurons towards active neurons.

(D) Number of GCaMP6f-expressing neurons and tdTomato-expressing neurons per field of view (375 μm x 300 μm ; note that the example shown in **(C)** is not a full field of view; 7 mice). The ratio of GCaMP6f-expressing neurons to tdTomato-expressing neurons is approximately 3.5 across mice. If labeling were complete, and assuming that roughly 20% of neurons in cortex are interneurons (Markram et al., 2004), one would predict a ratio of 4.

(E) As in **Figure 2.1E**, but for individual mice. Average population response to mismatch in all CT (thin blue lines) and NT (thin red lines) mice. The mean response over all neurons of CT (NT) mice is the thick blue (red) line.

(F) Scatter plot of average mismatch and playback halt responses for CT mice (left; 865 neurons, 25 outside axis limits and not shown) and NT mice (right; 425 neurons, 3 outside axis limits and not shown) in excitatory neurons with significant response to mismatch (see Experimental Procedures).

(G) Correlation between mismatch responses and playback halt responses for neurons with significant response to mismatch was significantly different from 0 in NT mice (9 mice, Student's t test, $p = 0.001$), but not in CT mice (9 mice, Student's t test, n.s., $p = 0.43$, see Experimental Procedures). Error bars indicate s.e.m.

(H) Mismatch selectivity measured as the absolute ratio of the mismatch response to the playback halt response as a function of the mismatch response, in CT (blue) and NT mice (red). With increasing mismatch response, mismatch neurons in CT, but not NT, mice become increasingly selective for mismatch versus playback halt. Data are represented as mean plus s.e.m.

(I) Scatter plot for response reliability as a function of mismatch response for CT (blue; 865 neurons) and NT mice (red; 425 neurons) in excitatory neurons with significant responses to mismatch. Response reliability was calculated for each neuron as the fraction of mismatch events with a significant response (see Experimental Procedures). Reliability was variable, but tended to increase with increasing mismatch responses.

(J) Average responses to mismatch (solid lines) and playback halt (dashed lines) of neurons with positive correlation of activity with visual flow (correlation greater than 0.05) on the first imaging day (CT: 24% or 539 of 2259 of neurons, 9 mice; NT: 24% or 513 of 2104 neurons, 9 mice). Orange area indicates duration of mismatch, shading indicates s.e.m. Statistical comparison as in **Figure 2.1E**.

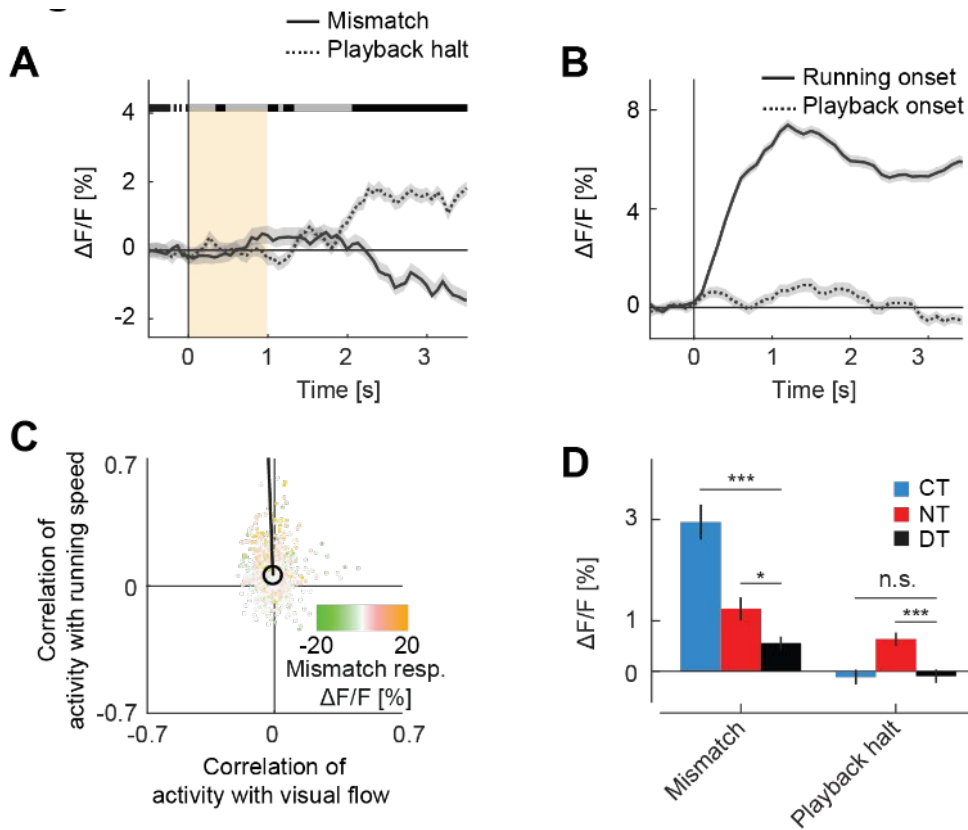


Figure S2.2. Population responses and correlation patterns in dark trained (DT) mice. Related to Figures 1 and 2.

(A) Average population responses to mismatch (solid line) and playback halt (dashed line) for excitatory neurons in DT mice (6 mice, 1076 neurons). Shading indicates s.e.m. Statistical comparisons as in **Figure 2.1E**, but for 67 ms time bins.

(B) Same as in **(A)**, but for running onset in closed-loop session and playback onset in open-loop session.

(C) As in **Figure 2.2B**, but for DT mice. Correlation coefficients between neural activity ($\Delta F/F$) of layer 2/3 excitatory neurons with running speed or with visual flow during open-loop sessions in DT mice. Each dot represents a single neuron. The color of each point indicates the amplitude of the mismatch response. The black circle marks mean correlation values. The solid black line indicates the angle of the mean first principle component of the distribution (see Experimental Procedures).

(D) Comparison of mismatch responses and playback halt responses of CT, NT and DT mice. *: $p < 0.05$, ***: $p < 0.001$, Student's t test.

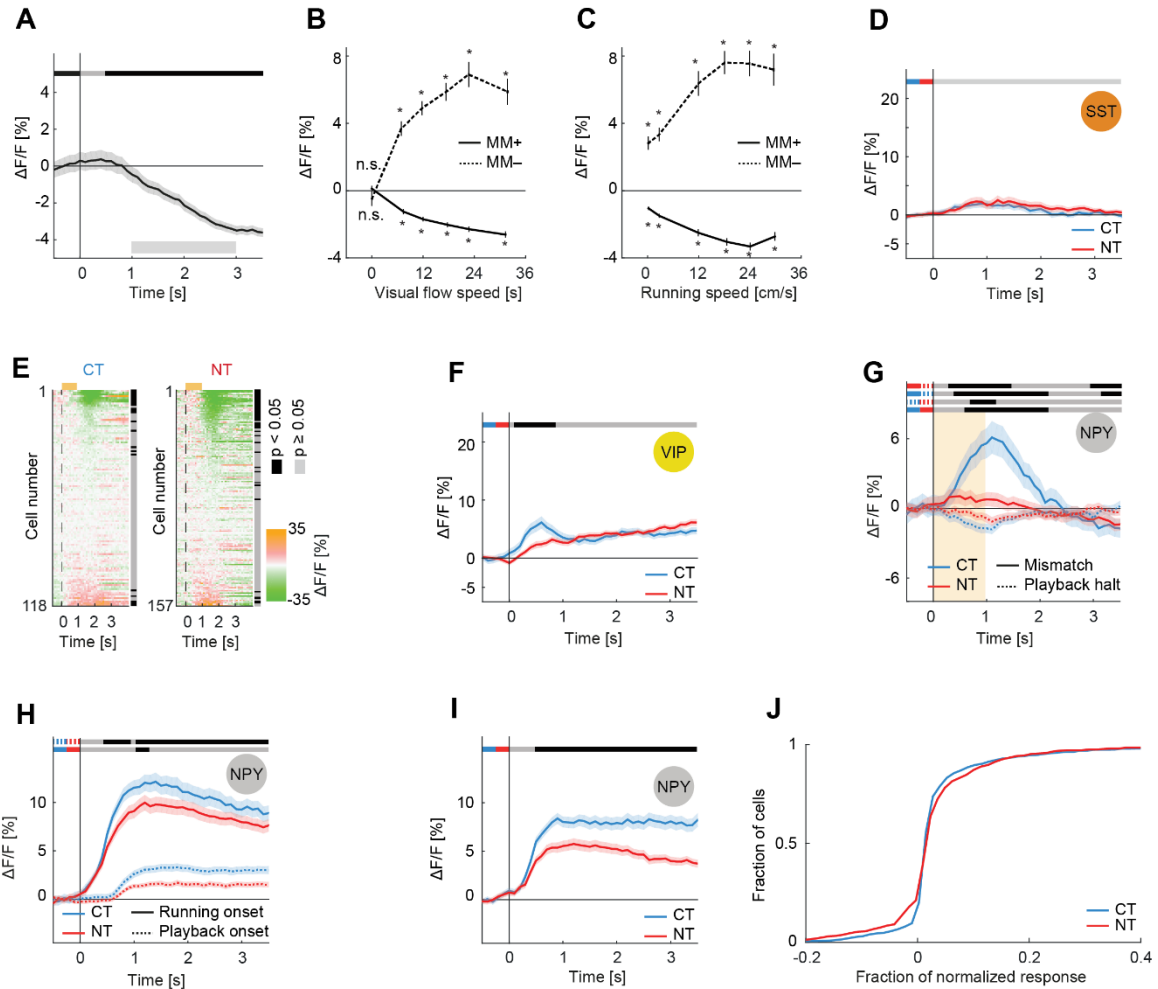


Figure S2.3. Additional analysis of responses of excitatory neurons and interneurons. Related to Figures 2 and 4.

(A) Average activity decreases in response to playback onset in open-loop sessions for excitatory neurons with a significant positive response to mismatch (see Experimental Procedures; 1110 neurons, data from CT and NT mice were pooled for this panel and in **(B)** and **(C)**). Note that for this analysis all playback onset events were included, independent of the running behavior. Shading indicates s.e.m. Statistical comparisons to baseline as in **Figure 2.1E**. Gray bar indicates window for calculating responses in **(B)** and **(C)**.

(B) The suppressive effect of visual flow on the activity of neurons with a significant positive response to mismatch increases with the speed of the visual flow (MM+, solid line, 1110 neurons). Similarly, the excitatory effect of visual flow on neurons with a significant negative response to mismatch also increased with visual flow speed (MM-, dashed line, 208 neurons). A visual flow speed of 36 cm/s corresponds to a temporal frequency of approximately 6 Hz in our virtual environment. Error bars indicate s.e.m. over events. n.s.: $p > 0.05$. *: $p < 0.001$, Student's t test.

(C) Average response to playback onset for the same neurons as in **(B)** for different running speeds. The inhibitory effect of visual flow is present in stationary mice and increases with increasing running speed. Note that the increase of the inhibitory effect with running speed is likely the consequence of an increase of baseline activity with locomotion. *: $p < 0.001$, Student's t test.

(D) Running-onset responses in complete darkness for SST interneurons (CT: 5 mice, 118 neurons; 5 mice, 157 neurons) on first imaging day. Statistical comparisons as in **Figure 2.1E**.

(E) Average mismatch response ($\Delta F/F$) of SST interneurons in CT mice (left, 5 mice, 118 neurons) and NT mice (right, 5 mice, 157 neurons), sorted by amplitude of mismatch response. Black and grey shading to the right indicates statistical significance of responses (grey: $p \geq 0.05$, black: $p < 0.05$, Mann-Whitney-U test; see Experimental Procedures). Orange bar marks the duration of mismatch.

(F) Running-onset responses in complete darkness for VIP interneurons (CT: 3 mice, 189 neurons; NT: 3 mice, 137 neurons) on first imaging day. Statistical comparisons as in **Figure 2.1E**.

(G) Average population responses to mismatch (solid line) and playback halt (dashed line) for NPY interneurons from CT (9 mice, 456 neurons) and NT (7 mice, 445 neurons) mice on first imaging day. Orange area indicates duration of mismatch, shading indicates s.e.m. Statistical comparisons as in **Figure 2.1E**.

(H) Same as in (E), but for running-onset responses in closed-loop sessions (solid lines) and playback onset in open-loop sessions (dashed lines) on first imaging day. Statistical comparison as in **Figure 2.1E**.

(I) Same as in (F), but for running-onset responses in complete darkness on first imaging day. Statistical comparison as in **Figure 2.1E**.

(J) Cumulative density plot of normalized mismatch responses of excitatory mismatch responsive neurons for CT (890 neurons) and NT (428 neurons) mice on first imaging day. Note the response distribution of the NT neurons is wider than that for CT neurons ($p < 0.01$, Kolmogorov–Smirnov test). This indicates that the distribution of mismatch responses over neurons is sharpened in CT mice, potentially by the selective activation of PV interneurons (**Figure 2.4E**).

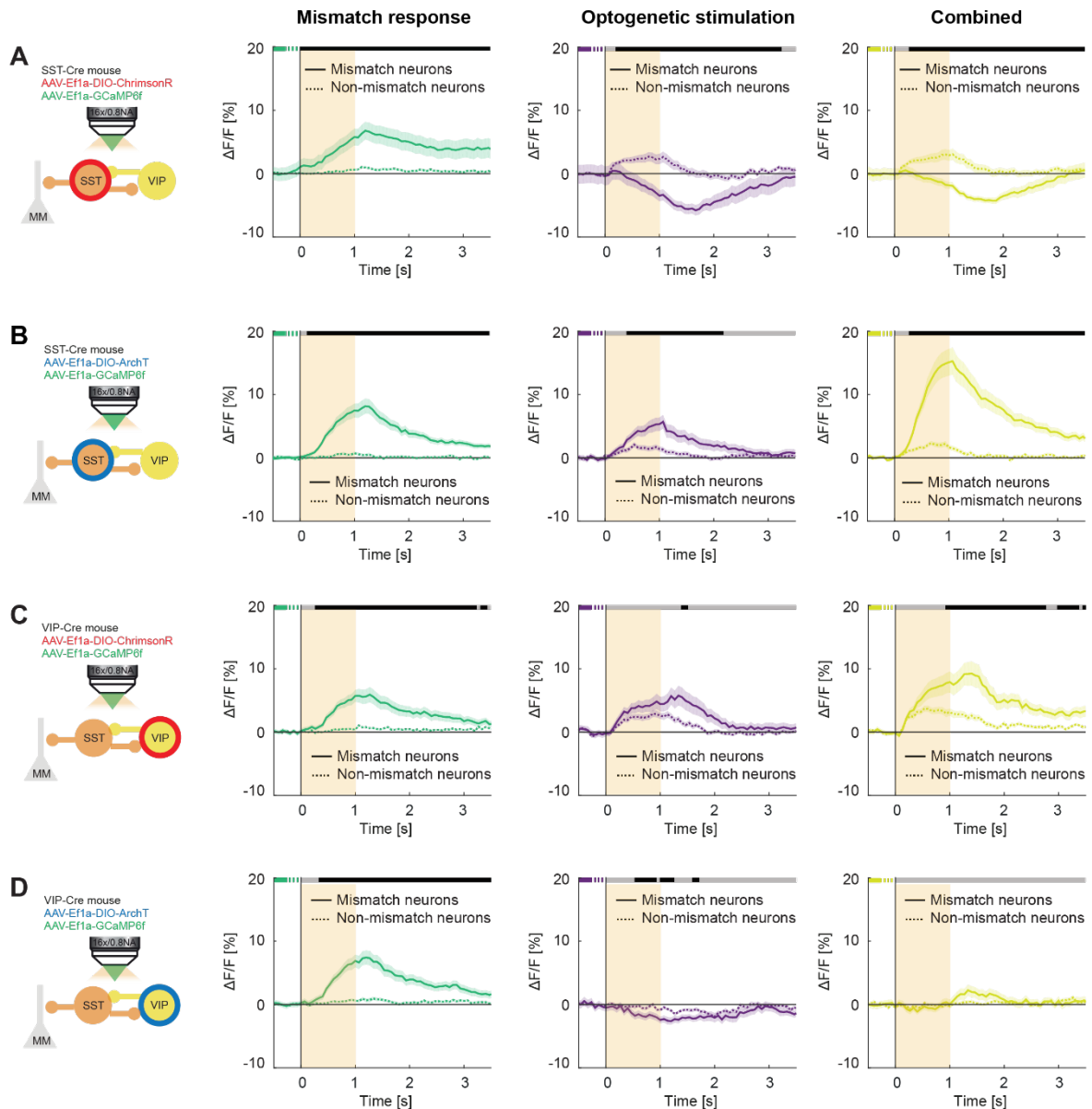


Figure S2.4. Effects of optogenetic stimulation on mismatch and non-mismatch neurons. Related to Figure 6.

(A) Left: Average responses to mismatch of 20% of neurons with the largest, positive mismatch response (solid green line, 165 neurons, 4 mice, the same neurons shown in **Figure 2.6**) and of 20% of neurons with no mismatch response (dashed green line, 165 neurons). Middle: Average response of the same mismatch responsive (solid purple line) and non-mismatch responsive neurons (dashed purple line) to optogenetic activation of SST interneurons. Right: Average response of mismatch responsive (solid yellow line) and non-mismatch responsive neurons (dashed yellow line) to concurrent optogenetic activation of SST interneurons and mismatch. Orange area indicates duration of mismatch and duration of optogenetic stimulation, shading indicates s.e.m. Statistical comparison as in **Figure 2.1E**, but for 67 ms time bins.

(B) As in **(A)**, but for optogenetic inhibition of SST interneurons (236 neurons in each group, 4 mice).

(C) As in **(A)**, but for optogenetic activation of VIP interneurons (114 neurons in each group, 4 mice).

(D) As in **(A)**, but for optogenetic inhibition of VIP interneurons (107 neurons in each group, 3 mice).

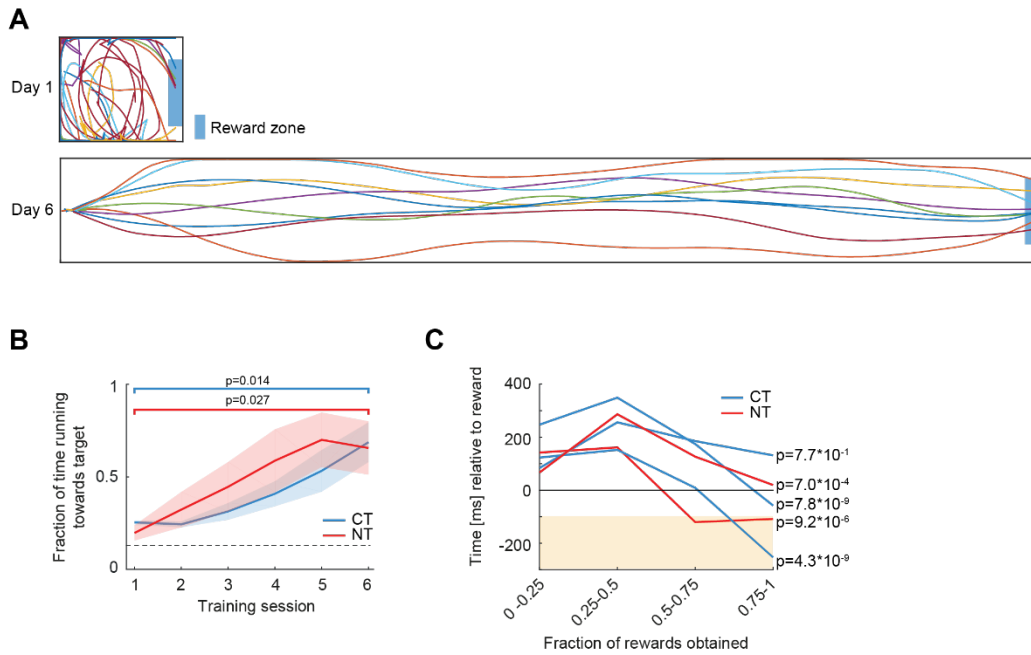


Figure S2.5. Both CT and NT mice learn to perform visuomotor tasks. Related to Figure 7.

(A) In a two-dimensional (2D) virtual locomotion task, water restricted mice need to learn to control a freely rotating styrofoam ball in order to traverse a linear corridor and reach a reward zone (blue shaded area). The length of the corridor was automatically increased as performance increased to keep reward rate constant. Upper panel: Sample trajectories of a single mouse in the 2D virtual environment on day 1. Trajectories were random, however the tunnel was short (approximately 0.5 m) and mice obtained rewards by chance. Once a reward was obtained, the mouse was teleported back to the beginning of the tunnel after a brief timeout (2 s) to start the next trial. Lower panel: Sample trajectories of the same mouse as in upper panel, but on day 6 when the mouse had learned the task (tunnel length approximately 6 m).

(B) Both CT (6 mice) and NT mice (4 mice) learned the 2D virtual locomotion task over the course of 6 training sessions (1 h/day). Task performance was quantified as the fraction of time spent running towards the reward zone. There was a significant increase from training session 1 to training session 6 for CT and NT mice (Student's t test). Shading indicates s.e.m.

(C) In a mismatch detection task, mismatch is followed by a water reward (100 ms delay after end of mismatch). Behavior is quantified as the latency to the first lick relative to the water reward (see Experimental Procedures). Mice were water restricted and habituated to licking from the water spout prior to testing. In the first training sessions, mice only licked after reward delivery. Over the course of 3 to 5 training sessions (1 h per day), both NT ($n = 2$) and CT ($n = 3$) mice started to lick during mismatch, before the reward was delivered. Orange area indicates mismatch. Shown is the mean time to the first lick as a function of the fraction of rewards obtained throughout training. To assess learning, we compared the distribution of the lick times of the first training session to the last training session for each mouse (sided Mann-Whitney-U test, p-values indicated adjacent to learning curve).

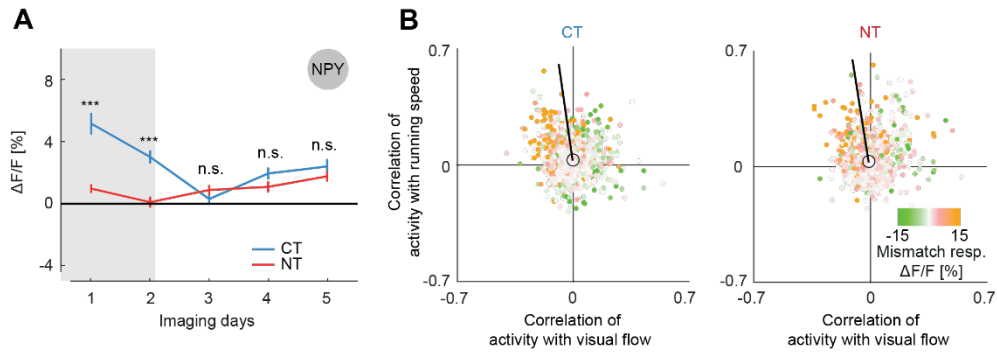


Figure S2.6. Normal visuomotor experience equalizes mismatch responses in NPY interneurons and correlation patterns in excitatory neurons. Related to Figure 7.

(A) Mismatch responses of NPY interneurons averaged over a 1 s window (see Experimental Procedures) as a function of imaging days for CT and (9 mice, 456 neurons) and NT (7 mice, 445 neurons). Mice were dark reared until the second imaging session (indicated by gray area). Error bars indicate s.e.m. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, Mann-Whitney-U test.

(B) Correlation coefficients between neural activity ($\Delta F/F$) and running speed or visual flow in CT (left) and NT (right) mice during open-loop sessions on imaging day 5. Each dot represents a single neuron (CT: 8 mice, 2213 neurons; NT: 7 mice, 1686 neurons). Dot color indicates the neuron's mismatch response. Black circles indicate mean correlations. Solid black lines indicate mean angle of first principle component of the distributions for each mouse (see Experimental Procedures).

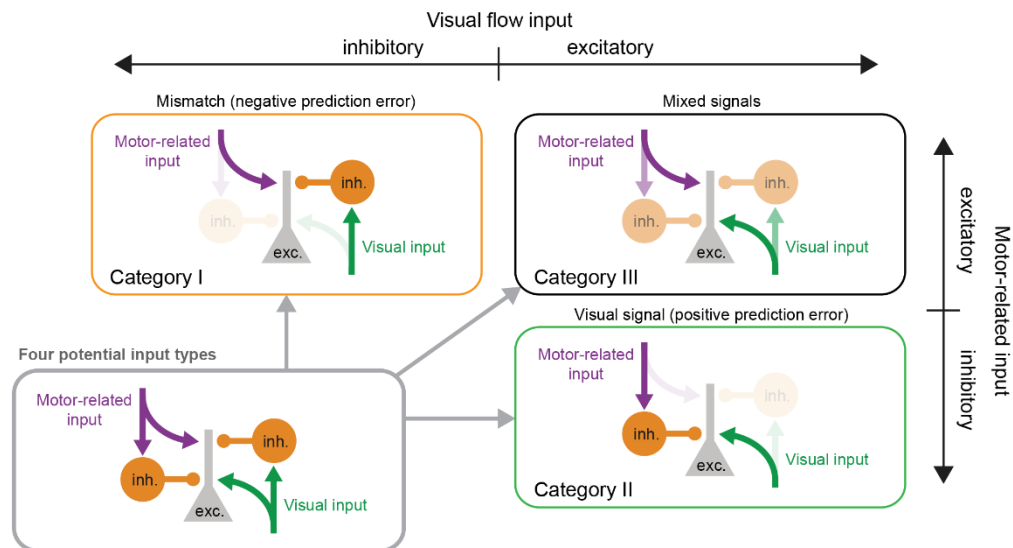


Figure S2.7. Layer 2/3 comparator circuits. Related to Figure 5.

In the cortical hierarchy, inputs to a layer 2/3 neuron can be divided into four different functional types: feed-forward excitatory, feed-forward inhibitory, top-down excitatory, top-down inhibitory. In this simplified model we will ignore neuromodulatory inputs for the moment. Here we use the terms feed-forward and top-down to denote the functional type and not necessarily the anatomical origin of the signal. Visual inputs to a layer 2/3 neuron in mouse V1 could either originate directly in dorsal lateral geniculate nucleus, or layer 4 neurons, or could be relayed by other layer 2/3 neurons. We propose that through visuomotor learning a balance between excitatory and inhibitory inputs is established in three different categories. Either top-down excitatory input is balanced against feed-forward inhibitory input to form a mismatch signal (category I), or feed-forward excitatory input is balanced against top-down inhibitory input (category II), or feed-forward and top-down excitation are combined (category III). In this model mismatch corresponds to a negative prediction error that signals occurrences of less visual input than expected. Visual inputs of category II correspond to a positive prediction error signaling more visual input than expected. Lastly, there is a population of neurons that appears to combine excitatory motor-related and visual input (category III). This particular population of neurons is prominent in NT mice (**Figure 2.2B**). One could speculate that either this is the default or naïve state of neurons in V1 before sensorimotor experience, or it is correlated with running speed and visual flow merely because it is comparing a top-down prediction of a movement whose frequency increases during locomotion. It is conceivable that a top-down prediction could for example be generated for visual feedback from eye movements. As the frequency of eye movements increases during running (Keller et al., 2012), the frequency of eye-movement related prediction errors would increase during visual flow and during running. Hence such an eye-movement prediction error neuron would appear running and visual flow correlated. Similar arguments can be made for most movements that are not predictable simply by running speed, but whose frequency systematically increases during running. This is not an easily solvable problem as we typically do not have experimental access to predictions of sensory feedback based on internal models. Here we have used locomotion as a proxy for a prediction of visual flow and thereby artificially reduced predictions to one dimension (forward locomotion). A more complete description of the behavior and the dynamics of the top-down inputs to V1 would have to be measured simultaneously to make a more informed estimate of the specific form of top-down predictions.

Note that in the absence of a top-down prediction (e.g. in experiments on anaesthetized or immobile animals passively viewing moving stimuli) neurons in both categories II and III will appear to be purely visually driven.

Methods

Contact for Reagent and Resource Sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Georg Keller (georg.keller@fmi.ch).

Experimental Model and Subject Details

The data for **Figures 2.1-2.4** and **2.7** were collected from a total of 61 mice. Experiments were started with surgery on P30. We used males and females of six different mouse lines: C57BL/6J mice (n = 6), vGAT-Cre x Ai14 tdTomato mice (n = 12), PV-Cre (Pvalb^{tm1(cre)Arbr}) (Hippenmeyer et al., 2005) (n = 11), SST-Cre (Sst^{tm2.1(cre)Zjh}) (Taniguchi et al., 2011) (n = 10), VIP-Cre (Vip^{tm1(cre)Zjh}) (Taniguchi et al., 2011) (n = 6) and NPY-Cre (Npy^{RH26Gsat/Mmucd}) (Gerfen et al., 2013) (n = 16). The data for **Figure S2.2** were collected from 6 C57BL/6J mice and experiments started at P30. The data for **Figure 2.5** were collected from a total of 6 SST-Cre mice, 7-16 weeks of age at the start of the experiment. The neural activity data for **Figure 2.6** were collected from 9 SST-Cre mice and 7 VIP-Cre mice 7-16 weeks of age at the start of the experiment. The behavioral data for Figure S5 were collected from 15 C57BL/6J mice, experiments were started on P30. For all experiments, mice were group housed throughout and both female and male mice were used. Unless stated otherwise, mice were housed in a 12h light/12h dark cycle. All animal procedures were approved by and carried out in accordance with guidelines of the Veterinary Department of the Canton Basel-Stadt, Switzerland.

Method Details

Surgery. Mice were briefly anesthetized with isoflurane and then received a subcutaneous injection of a Fentanyl (0.05 mg/kg; Actavis), Midazolam (5.0 mg/kg; Dormicum, Roche) and Medetomidine (0.5 mg/kg; Domitor, Orion) mixture. A 4 mm craniotomy was made over the right V1, centered on 2.5 mm lateral and 1 mm anterior of lambda. We labelled neurons with a calcium indicator by injecting an AAV2/1 vector (see section “Viral constructs” below for details) into right monocular V1, centered on 2.5 mm lateral and 0.5 mm anterior of lambda (3-4 injections per mouse, approx. 100-150 nl per injection). A 4 mm circular cover glass was glued in place using gel superglue (Ultra Gel, Pattex). The remaining exposed surface of the skull was scored with a needle to increase adhesion with glue and dental cement, and covered with Histoacryl (B. Braun). A titanium head bar was fixed to the skull using dental cement (Paladur, Heraeus Kulzer) (Leinweber et al., 2014). Mice were returned to their home cage after anesthesia was antagonized by an intraperitoneal injection of a Flumazenil (0.5 mg/kg; Anexate, Roche) and Atipamezole (2.5 mg/kg; Antisedan, Orion Pharma) mixture.

DREADD and optogenetic experiments. We used 7 to 16 weeks old male and female SST-Cre mice (for DREADD and optogenetic experiments) and VIP-Cre mice (for optogenetic experiments) that were reared in normal conditions. Craniotomy, virus injection, and headbar fixation was performed as described above. Mice were habituated to the setup, and imaging experiments started 2 weeks post surgery.

Viral constructs. We used AAV2/1-EF1 α -GCaMP5 (titer: 3.4×10^{11} GC/ml) for wild-type, EF1 α -CGaMP6f (titer: 5.6×10^{11} - 4.4×10^{12} GC/ml) for vGAT-Cre x Ai14, pharmacogenetic, and optogenetic experiments, and EF1 α -DIO-GCaMP6f (titer: 3.0×10^{11} - 7.8×10^{11} GC/ml) for inhibitory interneuron marker lines. To manipulate neural activity pharmacogenetically, we injected AAV2/1-EF1 α -DIO-hM4D(Gi)-mCherry (titer: 7.0×10^{11} GC/ml) or AAV2/1-EF1 α -DIO-hM3D(Gq)-mCherry (titer: 3.4×10^{11} GC/ml). For optogenetic manipulations of SST and VIP interneuron activity, we injected AAV2/1-EF1 α -DIO-ChrimsonR-tdTomato (titer: 2.2×10^{11} GC/ml) or AAV1-CAG-FLEX-ArchT-tdTomato (titer: 3.1×10^{12} GC/ml).

We initially attempted to label Cre-positive interneurons by means of a floxed RFP virus (AAV2/1-EF1 α -DIO-tdTomato) and bulk label all neurons with an unconditional GCaMP6f (AAV2/1-EF1 α -GCaMP6f) to concurrently record the activity of the selected interneuron type and all other neurons. However, for reasons unclear to us, this led to a very low co-labeling yield and signals in interneurons were often contaminated by the much stronger signals in surrounding excitatory neurons. We speculate that the reason for this may be that the promoter used (EF1 α) to drive the GCaMP6f expression is stronger in excitatory neurons than interneurons.

Virtual reality environment setup. The setup is based on the design of Dombeck and colleagues (Dombeck et al., 2007). Briefly, mice were head-fixed and free to run on an air-supported spherical treadmill. Rotation of the ball was restricted around the vertical axis with a pin. The virtual reality environment was projected onto a toroidal screen covering approximately 240 degrees horizontally and 100 degrees vertically of the mouse's visual field using a projector (Samsung SP-F10M) synchronized to the resonant scanner of the two-photon microscope. The virtual environment consisted of an infinite corridor with walls patterned with vertical sinusoidal gratings with a spatial frequency of approximately 0.04 cycles per degree (Leinweber et al., 2014).

Two-photon imaging. Functional two-photon calcium imaging was performed using 2 custom-built two-photon microscopes (Leinweber et al., 2014). Illumination source was a tunable femtosecond laser (Insight, Spectra Physics; Coherent Chameleon) tuned to 990 nm. Emission light was band-pass filtered using a 525/50 filter for GCaMP and a 607/70 filter for tdTomato/mCherry (Semrock) and detected using a GaAsP photomultiplier (H7422, Hamamatsu). Photomultiplier signals were

amplified (DHPCA-100, Femto), digitized (NI5772, National Instruments) at 800 MHz, and band-pass filtered around 80 MHz using a digital Fourier-transform filter implemented in custom-written software on an FPGA (NI5772, National Instruments). The scanning system of the microscopes was based either on a 12 kHz or an 8 kHz resonant scanner (Cambridge Technology). Images were acquired at a resolution of 750 x 400 pixels (60 Hz / 40 Hz frame rate, respectively), and a piezo-electric linear actuator (P-726, Physik Instrumente) was used to move the objective (Nikon 16x, 0.8 NA) in steps of 15 μm between frames to acquire images at 4 different depths. This resulted in an effective frame rate of 15 Hz or 10 Hz, respectively. The field of view was 375 μm x 300 μm .

Simultaneous two-photon imaging and optogenetic stimulations. ChrimsonR or ArchT stimulation and functional imaging of GCaMP6f-expressing neurons was done by using a modified Thorlabs B-Scope with a 12 kHz resonance scanner (Cambridge Technology) for line scanning. Illumination source for the optogenetic stimulation was a fast LED (UHP-T-595, Prizmatix) with a wavelength of 595 nm and which allowed fast TTL triggered operation. For spectral filtering we used a dichroic mirror (ZT775sp-2p, Chroma) to combine the two-photon laser and stimulation light. A second long-pass dichroic mirror (F38-555SG, Semrock) was used to split the GFP emission from both illumination light sources. Light leak from the 595 nm stimulation LED was reduced by synchronizing the LED light output to the turnaround times of the resonant scanner (during which imaging data were not acquired). Lastly, amplified PMT signals were digitally bandpass filtered at 80 MHz to reduce the effect of ringing in the amplifier. This allowed for near stimulation-artifact free synchronous imaging and optogenetic stimulation.

Experimental design. Mice were kept in the dark for an additional 2 days following surgery, after which they were introduced to the virtual reality environment. Mice were briefly anesthetized with isoflurane in the dark and then head-fixed on the setup. CT and NT mice were trained in pairs. The visual flow projected onto both screens was coupled to the locomotion of the CT mouse (**Figure 2.1B**). For dark training, mice were head-fixed and trained on the setup in complete darkness. All mice were free to run on the ball throughout training. In total, all CT, NT, and DT mice underwent 6 training sessions of 2 hours every other day (**Figure 2.1A**).

The first imaging experiment was performed 2 days after the last training session. The design of the imaging experiments was as previously described (Keller et al., 2012). Typically, an imaging experiment consisted of 1 closed-loop session and 2 open-loop sessions. In closed-loop sessions, the visual flow was coupled to the locomotion of the mouse, and was randomly perturbed with brief (1 s) halts (mismatch; one perturbation every 15 seconds on average). In open-loop sessions, the visual flow generated in the closed-loop session (including perturbations, here referred to as playback halt)

was replayed to the mouse independent of its locomotion. For some mice, open-loop sessions were followed by a dark session, where the virtual reality and all other light sources in the room were turned off. Each closed-loop, open-loop or dark session lasted 500 s. To minimize the effect of altered visuomotor experience (non-coupled experience in open-loop sessions for CT mice, and vice versa), we controlled the visual stimuli between imaging sessions so as to be the same as they were experienced in the training sessions, such that CT mice experienced closed-loop conditions (no perturbations) and NT and DT mice experienced open-loop conditions. Mice were kept in darkness between training and imaging sessions until after the second imaging session at which point they were transferred to in a normal 12 h/12 h light/dark cycle (**Figure 2.1A**). Note that DT mice were only imaged on time points 1 and 2.

At the end of each experiment intrinsic optical imaging was performed as described previously (Zmarz and Keller, 2016) to verify that the retinotopic location of recording sites corresponded to a part of the visual field covered by the toroidal screen.

Mismatch detection paradigm and 2D virtual locomotion task. Mice were dark-reared from birth and trained as either CT or NT, as described above. After the last training session, mice were water restricted. Throughout these experiments mice received water in the training paradigms. We monitored body weight of the mice and water was supplemented if body weight dropped below 80% of initial weight. Mice were first habituated to the setup and the lick spout in 2 sessions of 1 h each. Experiments started 2 days after the last training session. For the mismatch detection paradigm, we put the mice into the same virtual reality environment as described above in a closed-loop configuration including visual perturbations (mismatch) as described above. A droplet (approx. 10 μ l) of sucrose solution (15% in water) was delivered 100 ms after a mismatch via a metal spout placed in front of the mouse's snout. As mice learned the task, we observed anticipatory licking, which manifested as mice starting to lick during the 1 s mismatch, prior to reward delivery. A single experiment consisted of a 1 h closed-loop session during which the mice received approximately 100 rewards for a total of approximately 1 ml of sucrose solution. To assess learning for each mouse, the distribution of lick response times from the first training day was compared to the distribution of lick response times to the last training day using a Mann-Whitney-U test.

For the 2D virtual locomotion task, mice learned to navigate a virtual tunnel to a reward area. Mice had to learn to control heading in the virtual tunnel by rotation of the ball (rotation of the ball around the vertical axis was not restricted). Upon reaching the reward area, mice received a droplet (approx. 10 μ l) of sucrose solution (15% in water) via a lick spout. After a brief timeout (2 s), the position of the mouse in the virtual reality was reset to the starting location. The virtual tunnel was

kept very short initially and the tunnel length was increased progressively as mice learned the task, such that the average number of rewards received per minute was held approximately constant (at 1.3 rewards per min). The behavior was quantified as the amount of time the mice spent running in the direction of the reward area ($\pm 36^\circ$ from reward-area direction) normalized by the total time spent running. To quantify learning, the fraction of time spent running towards the target during training session 1 was compared to the fraction of time spent running towards the target on the last session for each mouse using a Student's t test.

Quantification and Statistical Analysis

Extraction of neuronal activity. Calcium imaging data were processed as previously described (Keller et al., 2012) and all data analysis was done in MATLAB (MathWorks). Briefly, raw images were full-frame registered to correct for brain motion. Neurons were manually selected based on mean and maximum fluorescence images. Raw fluorescence traces were corrected for slow drift in fluorescence using an 8th-percentile filtering with a 15 s window (Dombeck et al., 2007). $\Delta F/F$ traces were calculated as mean fluorescence in a selected region of every imaging frame and subsequently subtracted and normalized by the median fluorescence.

Data analysis. The details of the statistical analysis are noted in the figure legends. We did not test the distribution of the data for normality. To quantify average response traces, we first calculated the average event-triggered fluorescence trace for each neuron. The responses of all neurons were then averaged and the baseline (mean $\Delta F/F$ in a 0.5 s window pre event onset) was subtracted. To quantify the significance of the difference of two average calcium responses as a function of time, we performed a separate Student's t test for every bin of the calcium trace (10 Hz or 15 Hz) and marked bins as significantly different for $p < 0.01$. For visual clarity, we removed isolated significant bins, such that a significant bin was only marked if at least one of the two neighboring bins was also significant.

To calculate the average response of each neuron to mismatch or playback halt, we first calculated the difference between the average event-triggered response and the average response to 1000 randomly triggered events to generate a random-corrected trace. Average responses to mismatch and playback halt were then calculated as the mean fluorescence of the random-corrected average in a response window minus the mean fluorescence in a baseline window for each neuron (the response window for mismatch, playback halt, running onset and playback onset was +500 ms to +1500 ms, and the baseline subtraction window was -1000 ms to 0 ms). To determine the significance of a neuron's response, we calculated individual neuron responses to each mismatch event as described above and compared this distribution to the distribution generated by 1000

randomly triggered events. Significance was determined with a two sided Mann-Whitney-U test ($p < 0.05$). For mismatch and random events to be included in the analysis, mice had to be running above threshold (10^{-2} cm/s) before and after event onset (from -600 ms to + 1100 ms). In addition, for playback halt events to be included, mice had to be stationary during the playback halt (no running from -600 ms to +1100 ms). For running onset, mice had to be stationary for at least 600 ms prior to the running onset and continue running for 1100 ms above threshold following the onset. Similarly, for playback onset (quantified only during open-loop sessions) there had to be no visual flow for 600 ms prior to visual flow onset, followed by continuous visual flow above threshold for at least 1100 ms after onset, mice had to be stationary during this time.

To determine correlation between mismatch responses and playback halt responses (**Figures S2.1F and S2.1G**), we calculated Pearson's linear correlation coefficients for each mouse between the vector containing the mismatch responses of all neurons and the vector containing the playback halt responses for each neuron.

To calculate the response to playback onset as a function of visual flow speed and running speed (**Figures S2.3B and S2.3C**), we calculated the response to each playback onset for each neuron as the mean fluorescence in a response window minus the mean fluorescence in a baseline window (the response window was +1000 ms to +3000 ms, and the baseline window was -1000 ms to 0 ms). All playback onset events were used, irrespective of the running behavior. The same response window was also used to determine the visual flow speed and the running speed. For **Figure S2.3B**, The data were then split into different bins according to the visual flow speed. For each bin, both visual flow speed and playback onset response were then averaged. To estimate the baseline response (response to stationary grating), we calculated the response as described above, but for periods of no visual flow (stationary grating from a -2000 ms to +3000 ms). For **Figure S2.3C**, the data were split and averaged into different bins according to running speed.

We calculated Pearson's linear correlation coefficients to determine the correlation between individual neural activity and visual flow or running speed during the open-loop sessions. To minimize the influence of running-induced z-motion on the correlation coefficients, we calculated a threshold for each neuron ($3.72 * \text{standard deviation of the lower half of the fluorescence distribution}$) (Keller et al., 2012) and set all activity below this threshold to 1 (note that for $\Delta F/F$, baseline is at 1). To calculate the average correlations over days (**Figure 2.3**), we first calculated the average correlations per day and then averaged these across all imaging time points.

To calculate the principal component of the correlation distributions, we used the standard implementation available in MATLAB. We calculated the principal component for each imaging

region separately. To calculate the average angle, we averaged the vector sum of the normalized principle components of all imaging regions.

We calculated the average traces for the optogenetic experiments (**Figures 2.6, S2.4**), as described above. To further reduce the stimulation artefact after filtering, we used the following approach. The remaining stimulation artifact was approximated as a box function and subtracted from the average stimulation response of each neuron. The amplitude of the box function was estimated as the average of the of the absolute difference between the calcium signal on frame n-1 and n, and m and m+1, where the stimulation light was switched on between frame n-1 and n and switched off between m and m+1. On average this signal was 0.8% dF/F and much smaller than the typical neural response (**Figure 2.6 and S2.4**).

Average running speed during training and imaging sessions was calculated as the mean speed while the mouse was running above threshold (10^{-2} cm/s). Fraction of time running during training and imaging sessions was calculated as the fraction of time running speed was above threshold (10^{-2} cm/s) over total session duration. Note that during imaging sessions, fraction of time spent running and average running speed were calculated on the combined closed-loop and open-loop sessions.

Modified leaky integrate-and-fire neuron model. The model consists of a modified conductance based leaky integrate-and-fire neuron (cLIF) where inhibitory and excitatory conductances are linear combinations of running speed and visual flow speed. Parameters were adapted from (London et al., 2008; Salinas and Sejnowski, 2001). The membrane voltage was updated according to:

$$V_i = V_{i-1} - \frac{dt}{\tau_M} * [(V_{i-1} - E_L) + g_{ex} * (V_{i-1} - E_{ex}) + g_{in} * (V_{i-1} - E_{in})]$$

where $dt = 1$ ms, $\tau_M = 15$ ms, $E_L = -59$ mV, $E_{ex} = 0$ mV and $E_{in} = -78$ mV. When the voltage crossed the threshold (-40 mV), it was reset to $V_R = -48$ mV. The refractory period was 6 ms. The excitatory and inhibitory conductances were updated according to:

$$g_{ex,i} = g_{ex,i-1} * \left(1 - \frac{dt}{\tau_{ex}}\right) + X_{ex,i} * \Delta g_{ex}$$

$$g_{in,i} = g_{in,i-1} * \left(1 - \frac{dt}{\tau_{in}}\right) + X_{in,i} * \Delta g_{in}$$

where $\tau_{ex,in} = 5$ ms, $\Delta g_{ex} = 0.05$ and $\Delta g_{in} = 0.08$. Finally, to update the conductances, input X was calculated based on running speed R and visual flow V with scaling factors α, β :

$$X_{ex} = [\alpha R]^+ + [\beta V]^+, \text{ where } [x]^+ = 0, \text{ if } x \leq 0 \text{ and } [x]^+ = x, \text{ if } x > 0$$

$$X_{in} = |[\alpha R]^-| + |[\beta V]^-|, \text{ where } [x]^- = 0, \text{ if } x \geq 0 \text{ and } [x]^- = x, \text{ if } x < 0$$

Raw visual flow V and running speed R signals were normalized with the 95th percentile and upsampled to match the sampling frequency of the model (1 kHz).

To generate a simulated fluorescence signal F , we convolved the binary spike train S with a calcium kernel K :

$$F(i) = K(i) * S, \text{ where } K(i) = a * e^{-\tau*i}$$

The parameters were adjusted to the calcium indicator used and are based on estimates of published data (Chen et al., 2013) (for GCaMP5: $a = 0.05$, $\tau = 1.5$ s; for GCaMP6f: $a = 0.08$, $\tau = 1$ s) and the resulting trace F was downsampled to match the imaging frequency.

To predict neural activity of excitatory neurons during closed-loop sessions (**Figure 2.2E**), we trained the model on all available open-loop session data. We chose a correlation-based approach to find optimal values for α and β for every neuron. To maximize the correlation between simulated activity and neural activity, we employed a grid search approach. Neural activity was simulated over a wide range of combinations of the parameters α and β . We then calculated the correlation coefficient (Pearson's linear correlation) between simulations and the activity of a neuron, resulting in correlation maps (**Figures 2.2D and 2.2F**) characteristic for each neuron. The values of α and β resulting in maximal correlation were chosen to simulate activity during closed-loop sessions. To calculate average parameters α and β for the top 50% of excitatory mismatch neurons per mouse (**Figure 2.2F**), the vector $v_j = (\beta_j, \alpha_j)$ was transformed into polar coordinates and the mean was then calculated as:

$$\bar{m} = \frac{1}{N} \sum_{j=1}^N z_j, \text{ where } z_j = r_j * e^{i*\vartheta_j}$$

To estimate the fraction of explained variance (FEV) (**Figure 2.2F**), we used a cross validation approach. α and β were optimized as above, but using only 80% of the available open-loop session data. Optimal α and β were then used to predict the remaining 20% of the open-loop session. This was repeated 1000 times using randomly selected subsamples for training and testing. The FEV was estimated as the average squared correlation coefficient between prediction and actual neural activity. To estimate the FEV based on running speed only, β was held constant ($\beta = 0$), and similarly for simulations based on visual flow only, α was held constant ($\alpha = 0$). All simulations were performed in MATLAB using custom-written code.

Pupil dilation. Images of the left eye, contralateral to the craniotomy, were recorded with a CMOS camera at 30 Hz (DMKBUC03, Imaging Source). Pupil position was computed offline by smoothing

and thresholding the images and fitting a circle to the pupil. Data containing eye blinks were excluded from analysis. To extract mismatch induced pupil diameter changes, we computed the difference between the average dilation triggered on mismatch and the average dilation triggered on 1000 randomly chosen onsets. Average responses to mismatch was calculated as the difference between the amplitude averaged over a window pre (-100 ms to 0 ms) and post (+500 ms to +1500 ms) mismatch on the random-subtracted traces. To quantify significant difference as a function of time, we used the same bin-by-bin comparison described for calcium responses above, but with black bars indicating $p < 0.05$.

To determine neural response times, we calculated the time point of significant deviation between mismatch response traces of neurons with significant response to mismatch and randomly triggered traces (see above). For each neuron, we compared the fluorescence distributions of mismatch responses to random responses for each frame after the event onset (from 0 ms to +1500 ms). The response time was then taken as the first frame where the two distributions were significantly different (Mann-Whitney-U test, $p < 0.05$). The response time was only scored if the response distributions at 0 ms were not different and the responses diverged within the time window. This was the case for all excitatory neurons with significant response to mismatch. Pupil response times were calculated similarly.

CHAPTER 2: A SENSORIMOTOR CIRCUIT IN MOUSE CORTEX FOR VISUAL FLOW PREDICTIONS

This chapter is based on a paper that has been published in *Neuron* (Leinweber et al., 2017). The text and figures of this chapter correspond largely to the submitted manuscript, with minor adaptations to formatting and numbering to conform to the style of this thesis.

Marcus Leinweber¹, Daniel R. Ward¹, Jan M. Sobczak^{1,2}, Alexander Attinger^{1,3} & Georg B. Keller^{1,3,4}

¹*Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland*

²*Current address: Brain Research Institute, University of Zurich, Zurich, Switzerland*

³*Faculty of Natural Sciences, University of Basel, Basel, Switzerland*

⁴*Lead contact: georg.keller@fmi.ch*

Summary

Cortex is organized as a hierarchical processing structure. Feedback from higher levels of the hierarchy, known as top-down signals, have been shown to be involved in attentional and contextual modulation of sensory responses. Here we argue that top-down input to primary visual cortex (V1) from A24b and adjacent secondary motor cortex (M2) signals a prediction of visual flow based on motor output. A24b/M2 sends a dense and topographically organized projection to V1 that targets most neurons in layer 2/3. By imaging the activity of A24b/M2 axons in V1 of mice learning to navigate a 2D virtual environment we found that their activity was strongly correlated with locomotion and resulting visual flow feedback in an experience-dependent manner. When mice were trained to navigate a left-right inverted virtual environment, correlations of neural activity with behavior reversed to match visual flow. These findings are consistent with a predictive coding interpretation of visual processing.

Introduction

Visual processing in cortex is often described as a feedforward hierarchy of increasingly complex representations that functions to extract objects from visual input (Felleman and Van Essen, 1991; Marr, 1982; Riesenhuber and Poggio, 1999). In this framework top-down, or feedback, signals are thought to modulate visual responses based on visual representations in higher levels of the hierarchy. This modulation can be driven by attention (Roelfsema et al., 1998; Zhang et al., 2014),

context (Fiser et al., 2016; Zipser et al., 1996), or expectations (Gilbert and Li, 2013; Gilbert and Sigman, 2007). A mechanistic interpretation of these top-down signals is predictive coding (Rao and Ballard, 1999; Spratling, 2010). This theory postulates that a prediction of the component features based on visual representations in higher levels of the hierarchy is sent to lower levels, where it is compared to feedforward signals to compute mismatch between the two. Predictions are computed based on an internal model of the environment that is updated by mismatches fed forward from lower levels of the hierarchy. Predictive coding is a central idea in a whole family of theories of brain function (Barlow, 1994; Clark, 2013; Friston, 2010; Hawkins and Blakeslee, 2004; Wolpert et al., 1995).

Evidence for predictive coding in primary visual cortex (V1) has come mainly from its ability to explain non-classical visual response properties of V1 neurons (Grosf et al., 1993; Rao and Ballard, 1999; Spratling, 2010). One of the central problems with testing the hypothesis of predictive coding is that predictions are difficult to constrain experimentally. Typically, this is attempted using learned associations between behavior and sensory feedback. For example, sensory feedback couples in a predictable way to motor output. Hence, the experimental assumption is that signals generated during movement that are fed back to sensory areas should constitute an experience-dependent prediction of sensory feedback. This is referred to as an efference copy in the theory of internal models (Blakemore et al., 2000; von Holst and Mittelstaedt, 1950; Jordan and Rumelhart, 1992). Using paradigms of sensorimotor coupling, it has indeed been demonstrated that layer 2/3 of sensory cortices signals a mismatch between predicted and actual sensory feedback (Eliades and Wang, 2008; Keller et al., 2012). The source of the motor-related prediction necessary to generate these mismatch signals is still unknown.

Locomotion is sufficient to drive activity in mouse V1 (Keller et al., 2012; Saleem et al., 2013) and has been shown to modulate visually evoked activity (Niell and Stryker, 2010). Two hypotheses for the function of these motor-related signals are a gain modulation of visual responses (Ayaz et al., 2013; Fu et al., 2014; Niell and Stryker, 2010), and a prediction of visual flow based on motor output (Keller et al., 2012). Gain modulation of visual responses could be mediated by a low-dimensional signal such as a neuromodulatory input. A prediction of visual flow, on the other hand, would require a dense long-range excitatory input capable of conveying a high-resolution signal. Compatible with the idea of gain modulation, both cholinergic (Fu et al., 2014) and noradrenergic (Polack et al., 2013) inputs to V1 have been described. These neuromodulatory inputs drive locomotion-related gain changes in V1. However, the idea of neuromodulatory gain modulation during movement, acting through a disinhibitory circuit (Fu et al., 2014), cannot account for locomotion-driven activity in the

complete absence of visual input (Keller et al., 2012; Pakan et al., 2016; Saleem et al., 2013) or the context dependence of gain modulation (Pakan et al., 2016).

Here, we hypothesize that top-down projections to V1 carry a prediction of visual input based on motor output. The motor command for a left turn, for example, would lead to a prediction of full-field visual flow to the right. Generating a prediction of visual flow would thus require a transformation of the motor command to a signal in visual coordinates. In the framework of internal models, this transformation from motor coordinates (which muscles are activated) to sensory coordinates (how does the visual stimulus change) is referred to as a forward model (Wolpert et al., 1995). Anatomically, this could be implemented in different ways. A motor-related brain area could send an efference copy of the motor command to visual cortex in motor coordinates and the connectivity between axons from the motor-related area and visual neurons could act as the forward model transforming the signal to visual coordinates. Alternatively, the motor command could be transformed to visual coordinates in upstream areas and sent to visual cortex in visual coordinates. These two possibilities can be disambiguated by recording from axons of motor-related areas in visual cortex.

We sought to identify top-down inputs to V1 that could convey a prediction of visual flow based on motor output. One candidate area is the anterior cingulate cortex (ACC) (Miller and Vogt, 1984; Vogt and Miller, 1983; Zhang et al., 2014, 2016), which is bidirectionally connected with the adjacent secondary motor cortex (Vogt and Miller, 1983). Stimulation of ACC in rats causes head and eye movements (Sinnamon and Galer, 1984) and recent work demonstrates that the input from ACC to V1 has a role in attentional modulation of visual responses (Zhang et al., 2014) and experience-dependent spatial predictions (Fiser et al., 2016).

If the ACC input to V1 conveys a prediction of visual flow based on the mouse's movements, it should be characterized by a few defining features. First, the projection should have a bandwidth comparable to that of the feedforward visual input, measured either as the number of afferent neurons or as the fraction of V1 neurons targeted. Second, it should convey motor-related activity to V1, such that an inactivation of the input results in a decrease in motor-related activity in V1 neurons. Consequently, visuomotor mismatch signals in V1, which result from a comparison of an excitatory motor-related input and an inhibitory visual input (Attinger et al., 2017; Zmarz and Keller, 2016), should also decrease upon inactivation. Third, activity of this projection should depend on visuomotor experience and adapt to changes in visuomotor coupling as the mouse learns to control visual feedback in a novel environment. Lastly, artificial stimulation of the projection should result in illusory visual flow.

In the work we present here, we identify A24b, a subdivision of ACC, and an adjacent part of secondary motor cortex (M2), as one origin of a motor-related input to V1, and revisit the question of the function of this top-down projection. We suggest that one component of this top-down input is a prediction of visual flow based on the motor output of the mouse, consistent with the idea that internal models govern the processing of sensory input.

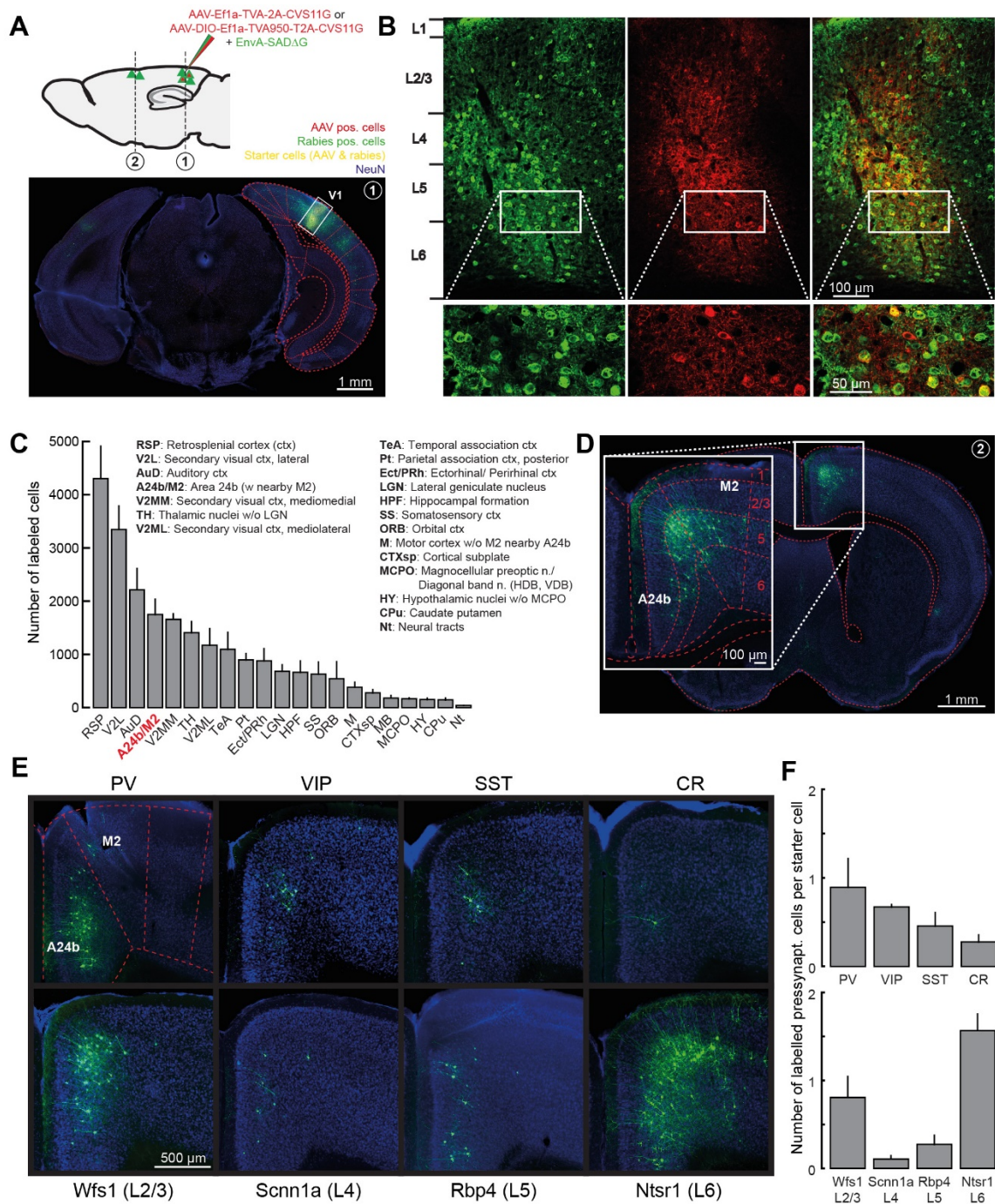


Figure 3.1. A24b/M2 is one of the main inputs to V1 and primarily targets layer 2/3 and 6 excitatory neurons as well as PV interneurons.

(A) Top: Injection scheme for monosynaptic input mapping experiments using rabies virus. Vertical dashed lines indicate position of coronal sections in (A) and (D). Bottom: Coronal section (λ -0.25 mm) through an injection site in V1. AAV

positive cells are labeled in red and rabies positive cells are labeled in green, NeuN staining in blue. **(B)** Top: Magnification of the injection site shown in (A). Left: Cells infected by the rabies virus express GFP and appear green. Middle: Cells infected by the AAV virus were identified by staining against the 2A linker peptide and appear red. Right: Overlay of the green and the red channels. Cells that appear yellow have been infected by both the AAV and the rabies virus and are putative starter cells. Bottom: Magnification of the area indicated by the white rectangle. **(C)** Number of cells labeled in different brain regions by rabies tracing from V1 (3 mice). Brain regions were identified using a mouse brain atlas (Franklin and Paxinos, 2012). Note, A24b/M2 is one of the largest inputs to V1, with twice as many cells labeled as in the dorsolateral geniculate nucleus (dLGN), the main source of thalamic input to V1. Here and in the following panels, error bars indicate SEM. **(D)** Example of a coronal section through A24b/M2 (bregma +0.50 mm) with cells labeled by rabies tracing from V1. **(E)** Example images of A24b/M2 showing cells labeled by rabies tracing restricted to different starter cell populations in V1. PV: parvalbumin positive interneurons, VIP: vasoactive intestinal peptide positive interneurons, SST: somatostatin positive interneurons, CR: calretinin positive interneurons, Wfs1: expression mainly in layer 2/3, Scnn1a: expression mainly in layer 4, Rbp4: expression mainly in layer 5, Ntsr1: expression mainly in layer 6. **(F)** Average number of labeled cells in A24b/M2 normalized by the number of starter cells in V1 for each Cre-line.

Results

A24b/M2 is one of the main inputs to V1 and mainly targets layer 2/3 and layer 6 excitatory neurons and PV interneurons.

To survey all afferents to V1 and identify potential motor-related input, we used transsynaptic rabies tracing (Wickersham et al., 2007). We injected an adeno-associated virus (AAV; AAV2/1-Ef1a-TVA950-T2A-CVS11G) into V1 to express the rabies virus G protein and the TVA receptor in V1 neurons. Monosynaptic rabies tracing was initiated using an EnvA-coated SADΔG rabies virus (**Figure 3.1A**). The AAV vector was injected into V1 two days prior to the injection of the rabies virus at the same location. As the Ef1a promoter used in the AAV virus tends to bias expression to layer 2/3 and layer 5 excitatory neurons (Attinger et al., 2017), tracing was likely also biased to presynaptic input to those neurons. Five days after the rabies virus injection, mice were sacrificed and their brains were sectioned and immunohistologically processed (see Experimental Procedures). We found an average of 2179 ± 877 (mean \pm std) labeled starter cells per mouse throughout all layers of V1 (**Figure 3.1B**), and 30437 ± 6112 presynaptic cells throughout the entire brain (3 mice). Presynaptic cells were mapped onto brain regions as defined by a mouse brain atlas (Franklin and Paxinos, 2012) (**Figure 3.1C**). We found the largest fraction of presynaptic cells in primary (7723 \pm 2124 cells, or 25%) and secondary visual areas (V2L, V2ML, and V2MM, 6177 \pm 1444 cells, or 20%), followed by retrosplenial (4304 \pm 1055 cells, or 14%) and auditory cortex (2212 \pm 694 cells, or 7%). The largest input from a motor-related area came from a region of cortex that was composed of A24b (1748 \pm 501 cells, or 6%), a sub-region of the anterior cingulate cortex (ACC) approximately corresponding to the dorsal part of ACC in a region just anterior of bregma (0 mm to +1 mm) (Vogt and Paxinos, 2014), and the medial part of M2 directly adjacent to A24b (**Figure 3.1D**). Using a boundary between A24b and M2 as defined in the Franklin and Paxinos mouse brain atlas (Franklin and Paxinos, 2012), 57% (1001 of 1748) of the neurons were in A24b, and 43% (747 of 1748) of the rabies positive neurons were in M2. As we find no evidence of a separation between the population of neurons in A24b and

the subset of neurons in the immediately adjacent medial part of M2, we refer to this region of cortex that projects to V1 as A24b/M2. Most of the presynaptically labeled cells in A24b/M2 were located in layer 5 (**Figures 3.1D** and **S3.1A,B**). All further analysis was focused on A24b/M2 as a potential source of motor-related input to V1.

To identify the target cell types of the projection from A24b/M2 to V1, we performed a series of cell-type specific rabies tracing experiments by restricting the expression of the TVA receptor and G protein to genetically defined populations of neurons. We injected AAV2/1-Ef1a-DIO-TVA950-T2A-CVS11G and EnvA-coated SADΔG into V1 of mice expressing Cre in different subpopulations of inhibitory neurons (parvalbumin (PV), vasoactive intestinal peptide (VIP), somatostatin (SST), or calretinin (CR)) or excitatory neurons predominantly in layer 2/3 (Wfs1-Cre), layer 4 (Scnn1a-Cre), layer 5 (Rbp4-Cre) or layer 6 (Ntsr1-Cre). We found that all tested neuron types in V1 receive input from A24b/M2 (**Figure 3.1E,F**). The largest fraction of presynaptically labeled cells in A24b/M2 was found when tracing from layer 2/3 and layer 6 excitatory neurons and PV interneurons.

Axonal input from A24b/M2 to V1 is dense and maintains topography.

To quantify the innervation pattern of A24b/M2 axons in V1 we labeled axons from A24b/M2 with a red fluorescent protein by injecting AAV2/1-Ef1a-tdTomato in A24b/M2 (**Figure 3.2A**). A24b/M2 axons densely innervate layer 1 and layer 6 of V1 (**Figure 3.2B-D**). To test if there is a topographic structure to the projection, we injected two AAVs expressing either eGFP or tdTomato 500 μm apart into A24b/M2, either separated along the anterior-posterior (**Figure 3.2E**) or the medio-lateral axis (**Figure S3.1C-E**). We found that posterior A24b/M2 projects to medial V1 and anterior A24b/M2 projects to lateral V1. Conversely, medial A24b/M2 projects to anterior V1 and lateral A24b/M2 projects to posterior V1. Thus, the projection from A24b/M2 to V1 maintains topography.

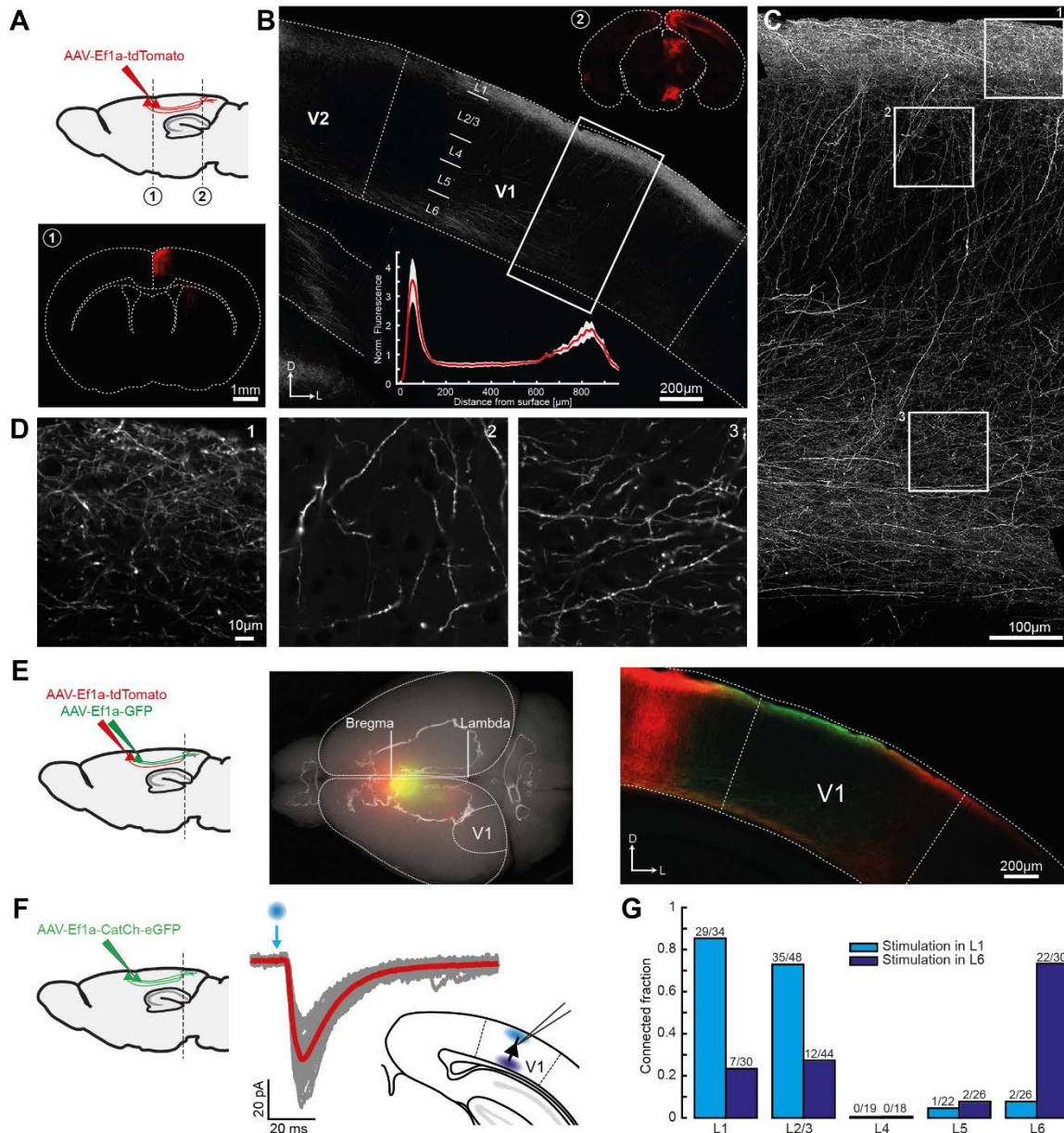


Figure 3.2. Axonal input from A24b/M2 to V1 is dense and maintains topography.

(A) Top: Injection scheme to label A24b/M2 axons in V1. Dashed lines indicate the locations of the coronal sections shown below and in **(B)**. Bottom: Wide-field fluorescence image of a coronal section at the location of the injection site. **(B)** Coronal section showing A24b/M2 axons in V1 and adjacent structures. A24b/M2 axons mainly innervate layer 1, 5 and 6 in V1. Box indicates magnified view shown in **(C)**. Top inset: Wide-field fluorescence image of a coronal section of the whole brain at the location indicated in **(A)** (lambda +0.40 mm). A24b/M2 axons shown in red innervate V1, retrosplenial cortex and subcortical structures. Bottom inset: Fluorescent labelling density as a function of depth from the surface of cortex in V1 (averaged over data from 7 mice, shading indicates SEM). **(C)** Maximum intensity projection of confocal images of the region marked in **(B)** showing dense innervation of V1 by A24b/M2 axons. Boxes correspond to the magnified views shown in **(D)**. **(D)** Confocal images of axons in layer 1 (1), layer 2/3 (2) and layer 6 (3). **(E)** Left: Injection scheme for dual color AAV topographic mapping. An AAV2/1-Ef1a-eGFP and an AAV2/1-Ef1a-tdTomato were injected separately, spaced by approximately 500 μm along the anterior-posterior axis in A24b/M2. Dashed line indicates position of the coronal section shown on the right. Middle: Wide-field fluorescence image of the brain three weeks post injection. Right: Pattern of labeling of A24b/M2 axons in V1. Anterior A24b/M2 (in red) projects to lateral V1 and posterior A24b/M2 (in green) projects to medial V1. **(F)** Left: Injection scheme for CRACM. An AAV2/1-Ef1a-CatCh-eGFP was injected into A24b/M2. Dashed line indicates position of coronal slices used for patch-clamp recordings. Middle: Example stimulation responses of a layer 2/3 neuron to stimulation of layer 1 axons (gray: single trials, red: average). Right: Recording scheme. Neurons were patched in different layers of V1 while axons were stimulated either in layer 1 or layer 6. **(G)** Fraction of neurons with a significant response to A24b/M2 axon stimulation for each cortical layer, stimulated either in layer 1 or layer 6. Numbers above bars indicate the number of neurons with a significant response out of the total number of neurons recorded.

To quantify the fraction of neurons in different layers of V1 that receive direct input from A24b/M2, we used channelrhodopsin-2-assisted circuit mapping (CRACM) (Petreanu et al., 2007). We expressed channelrhodopsin-2 in A24b/M2 neurons by injecting an AAV2/1-Ef1a-CatCh-eGFP into A24b/M2 and sacrificed mice 21 days later for slice recording experiments (see Experimental Procedures). We performed voltage-clamp recordings from neurons throughout all layers of V1 and recorded excitatory postsynaptic currents (EPSCs) in response to optical stimulation of A24b/M2 axons in either layer 1 or layer 6 (**Figure 3.2F**). To ensure that the evoked EPSCs were monosynaptic, the stimulation intensity was set far below the threshold necessary to trigger action potentials in V1 neurons in current-clamp recordings (see Experimental Procedures). Depending on the site of stimulation we found evoked responses in most neurons of layer 1 (29 of 34, or 85 %), layer 2/3 (35 of 48, or 73 %) and layer 6 (22 of 30, or 73 %; **Figure 3.2G**). Only a small fraction of layer 5 neurons (2/26, or 8%) and no layer 4 neurons (0/19, or less than 5%) exhibited A24b/M2 axon photo-stimulation evoked responses. On average the response latencies of the EPSCs were short (6.74 ± 3.27 ms, mean \pm std), consistent with a monosynaptic input (**Figure S3.1F-H**) (Petreanu et al., 2007). Thus, consistent with rabies tracing experiments, we found that A24b/M2 input mainly targets layers 1, 2/3 and 6, and directly innervates many of the neurons in these layers.

A24b/M2 axons convey motor-related signals to V1.

Activity in mouse V1 has been shown to reflect visual (Niell and Stryker, 2008), motor-related (Keller et al., 2012; Saleem et al., 2013), and visuomotor-mismatch signals (Keller et al., 2012; Saleem et al., 2013). To quantify a potential contribution of A24b/M2 input to these signals in V1, we transfected cells in A24b/M2 with an AAV2/1-Ef1a-GCaMP6f (Chen et al., 2013) and imaged activity of A24b/M2 axons in V1 (6007 axons in 10 mice) of head-fixed mice running on a spherical treadmill surrounded by a toroidal screen that provided visual flow feedback (**Figure 3.3A**) (Dombeck et al., 2007). Mice were exposed to either closed-loop conditions, in which visual flow feedback was coupled to locomotion and brief halts of visual flow were used to probe for mismatch responses, open-loop conditions, in which visual flow was presented independent of locomotion, or dark conditions (see Experimental Procedures). We found that the population of A24b/M2 axons exhibited strong motor-related activity that preceded running onsets by several hundred milliseconds (approximately 900 ms; see Experimental Procedures; **Figure 3.3B**). This anticipatory activity was stronger and started earlier than the anticipatory activity observed in layer 2/3 of visual cortex. To test whether the increased anticipatory activity was the result of a shift in the onset of activity in all axons or the result of the broadening of the distribution of onset times across the population of axons, we analyzed the distribution of the time differences between activity onset and running onset in the A24b/M2 axons and layer 2/3 V1 neurons. Although a substantial fraction of A24b/M2 axons (30%,

508 of 1,675) had activity onsets that preceded running onset, the average latency to activity onset followed the running onset (0.28 ± 0.13 s post running onset, 10 mice, mean \pm SEM), but preceded the average latency to response in layer 2/3 V1 neurons (0.43 ± 0.06 s post running onset, 8 mice, mean \pm SEM). However, the main effect contributing to the stronger anticipatory activity in A24b/M2 axons was a broader distribution of activity-onset times (FWHM of the distribution of A24b/M2 axon response times: 1.08 ± 0.18 s; FWHM of the distribution of layer 2/3 V1 neuron response times: 0.52 ± 0.16 s; **Figure S3.2A**). Consistent with the idea that this motor-related activity is not just the consequence of a brain state change, we found that the amplitude of the motor-related activity in the A24b/M2 axons increased with increasing running speed (**Figure S3.2B**). In addition to motor-related signals, we also found responses to visual stimulation in open-loop conditions in A24b/M2 axons that occurred with a delay of approximately 1000 ms after visual flow onset. We found no evidence of a correlation between the strength of an axon's response to visual flow onset and the strength of the axon's motor-related response (**Figure S3.2C**). Finally, we quantified the response of A24b/M2 axons to visuomotor mismatch, which results in strong mismatch responses in layer 2/3 neurons of V1 (Keller et al., 2012), and did not find any change in activity induced by mismatch (**Figure 3.3B**).

To test if motor-related activity in A24b/M2 axons depends on visual input prior to running onset, we analyzed activity in four different conditions: (1) running onset in the closed-loop condition, (2) running onset in the open-loop condition without visual flow preceding the running onset, (3) running onset in the open-loop condition with visual flow preceding running onset, and (4) running onset in darkness (**Figure 3.3C**, see **Figure S3.2D** for the same data before matching running speeds). We found that activity in A24b/M2 axons immediately following running onset depended strongly on visual flow preceding the running onset. Activity was highest when the mouse was seeing a static grating prior to the onset of running (conditions 1 and 2), and lower when the mouse either saw an already moving grating (condition 3) or nothing (darkness; condition 4) prior to running onset. The suppressive effect of visual flow was most pronounced when visual flow occurred in a 3 second window preceding running onset (**Figure S3.2E**). Consistent with a modulation of A24b/M2 responses by visual flow, A24b/M2 axons exhibited a marked decrease in activity when we switched to darkness while the mouse was running (**Figure S3.2F,G**). Hence, the motor-related activity in A24b/M2 axons is suppressed by either a lack of visual input or by visual flow preceding running onset.

To test if A24b/M2 axons specifically target neurons in layer 2/3 of V1 that exhibit strong motor-related activity we developed an *in vivo* version of CRACM (ivCRACM). We injected an AAV2/1-hSyn-ChrimsonR-tdTomato into A24b/M2 and an AAV2/1-Ef1a-GCaMP6f into V1 (**Figure 3.3D**). To identify

neurons driven by activity in A24b/M2 axons, we stimulated A24b/M2 axons locally in V1 and measured the calcium responses of layer 2/3 neurons (see Experimental Procedures). We identified 82 of 648 (or 13%) neurons that responded significantly to the local stimulation of A24b/M2 axons (**Figure 3.3E**, and **Figure S3.3A,B**). Note that with this method we likely only identify those neurons in V1 that receive the strongest A24b/M2 input and not all neurons that are connected. Consistent with the idea that motor-related activity in A24b/M2 axons drives motor-related activity in V1, stimulation responsive neurons also exhibited stronger motor-related activity than neurons that did not respond to A24b/M2 axon stimulation (**Figure 3.3F**). This was not simply the consequence of these neurons being more active than non-responsive neurons (**Figure S3.3C**). We also verified that the stimulation responses were absent in V1 neurons of control mice that did not express ChrimsonR in A24b/M2 axons under otherwise identical conditions (**Figure S3.3D**). This demonstrates that A24b/M2 input best activates neurons in V1 with the highest running-onset activity.

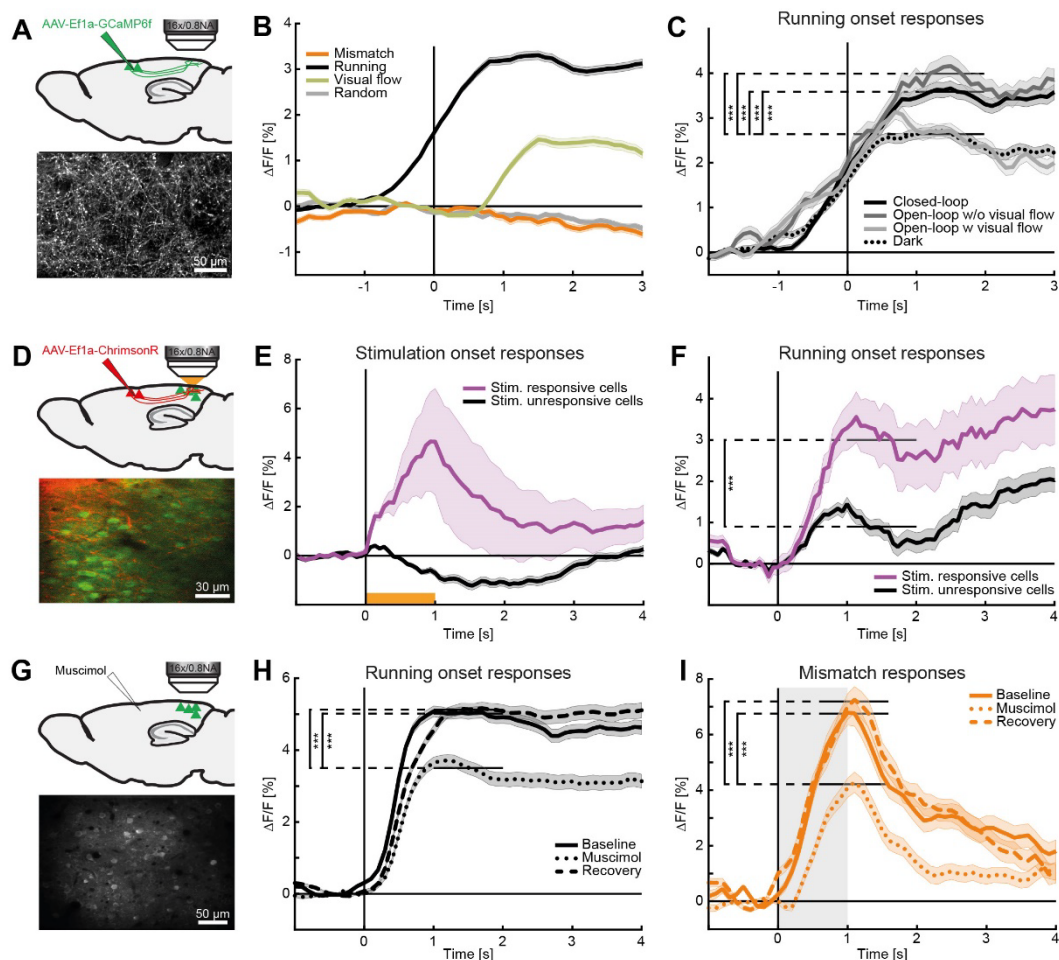


Figure 3.3. A24b/M2 axons convey motor-related signals to V1.

(**A**) Top: Schematic of axonal labeling and imaging. We injected an AAV expressing GCaMP6f in A24b/M2 and recorded the activity of A24b/M2 axons in V1. Bottom: Example two-photon image of A24b/M2 axons in V1. (**B**) Average population response of A24b/M2 axons (6007 axons, 100 imaging sites, 10 mice) in V1 to running onset (black line), visual flow onset (green line), mismatch onset (orange line) and triggered on random times (gray). Here and in the following panels, shading indicates SEM. (**C**) Average population response of A24b/M2 axons in V1 to running onset during a closed-loop session (black line), during an open-loop session without visual flow prior to running onset (dark gray line) and with visual flow

prior to running onset (light gray line), and during darkness (dotted black line). The visual stimulus preceding running onset strongly influences running onset activity in A24b/M2 axons (closed-loop versus open-loop with visual flow: $p = 1 \times 10^{-4}$, and versus dark: $p = 2 \times 10^{-6}$; open-loop without visual flow versus open-loop with visual flow: $p = 4 \times 10^{-8}$, and versus dark: $p = 2 \times 10^{-10}$; Student's t test, solid horizontal lines indicate testing window). **(D)** Top: Schematic of simultaneous axonal stimulation and imaging. We expressed ChrimsonR-tdTomato in A24b/M2 neurons and stimulated the A24b/M2 axons in V1 while recording GCaMP6f activity of layer 2/3 neurons in V1. Bottom: Example two-photon image of A24b/M2 axons expressing ChrimsonR (shown in red) and layer 2/3 neurons expressing GCaMP6f (shown in green) in V1. **(E)** We stimulated A24b/M2 axons in V1 of behaving mice while simultaneously recording the responses of V1 neurons. 82 of 648 neurons (or 13%) in V1 exhibited significant responses to the stimulation of A24b/M2 axons (purple line). Stimulation duration is indicated by the orange bar (1 s, 20 Hz, 50% duty cycle). **(F)** Running onset response of neurons selected in (E) that were strongly activated by A24b/M2 axon stimulation (purple line) and neurons that were not activated by A24b/M2 axon stimulation (black line; $p = 3 \times 10^{-5}$, Student's t test, solid horizontal lines indicate testing window). **(G)** Top: Schematic of A24b/M2 silencing during V1 imaging. We injected muscimol in A24b/M2 while recording GCaMP6f activity of layer 2/3 neurons in V1. Bottom: Example two-photon image of neurons in layer 2/3 of V1 expressing GCaMP6f. **(H)** Average population response of running onset responsive neurons (50 % most running onset responsive neurons selected on preceding time points; 1934 of 3868 neurons; see Experimental procedures) in V1 to running onset before (baseline; solid line), during (muscimol; dotted), and after (recovery; dashed) inactivation of A24b/M2 (baseline versus muscimol: $p = 5 \times 10^{-9}$; recovery versus muscimol: $p = 3 \times 10^{-10}$; Student's t test, solid horizontal lines indicate testing window). **(I)** Average population response of mismatch responsive neurons (50 % most mismatch responsive neurons selected on preceding time points; 1934 of 3868 neurons; see Experimental procedures) in V1 to mismatch before (baseline; solid line), during (muscimol; dotted), and after (recovery; dashed) inactivation of A24b/M2 (baseline versus muscimol: $p = 1 \times 10^{-10}$; recovery versus muscimol: $p = 1 \times 10^{-9}$; Student's t test, solid horizontal lines indicate testing window). Gray shading indicates the duration of feedback mismatch.

To test if A24b/M2 suppression affects motor-related signals in V1, we inactivated A24b/M2 in a separate set of mice, while recording neural activity in V1. We expressed GCaMP6f in V1 and recorded running onset and mismatch responses one day before, immediately after and one day after muscimol injection into A24b/M2 (**Figure 3.3G**). We found that both running onset responses (**Figure 3.3H**) and mismatch responses (**Figure 3.3I**) were significantly decreased in responsive neurons during inactivation of A24b/M2 (see Experimental Procedures). This inactivation experiment has two confounds. First, inactivating A24b/M2 led to changes in motor behavior. On average, mice decreased average running speed during A24b/M2 inactivation (**Figure S3.2H**). To correct for a possible bias of differences in running speed, average running speed was matched across conditions by subsampling individual running onset and mismatch trials (see Experimental Procedures; **Figure S3.2I, J**). Second, the inactivation volume likely only encompassed a fraction of the total A24b/M2 volume. Thus, the true effect size of a complete A24b/M2 inactivation is likely larger than the effects we report here. To rule out the possibility that the decrease in running onset responses and mismatch responses is the result of an acute off-target effect (Otchy et al., 2015), we also performed chronic lesions of A24b/M2 using ibotenic acid injections into A24b/M2. Consistent with the acute inactivation, both motor-related activity and mismatch responses were reduced following a chronic lesion of A24b/M2 (**Figure S3.2K, L**).

[A24b/M2 input to V1 correlates with expected visual feedback given a motor output.](#)

To test if activity in A24b/M2 axons correlates with specific motor behaviors, we trained mice to navigate to the end of a 2D virtual tunnel and simultaneously recorded the activity of ipsilateral A24b/M2 axons in either the left or right monocular V1 (**Figure 3.4A**; left V1: 153 sites, 34533 axons;

right V1: 55 sites, 7476 axons). On average we found that $22\% \pm 1.4\%$ (mean \pm SEM) of axons exhibited activity that was task-related (as defined by correlation with locomotion or rotation, see Experimental Procedures). These axons were then classified as either correlating with left turns or with right turns based on the correlation of their activity with the mouse's movement trajectory in the virtual environment (see Experimental Procedures). In left monocular V1, we found a bias for axons whose activity correlated with left turns (**Figure 3.4B**; left hemisphere: fraction 0.60 vs 0.40 \pm 0.02, $p = 4 \times 10^{-13}$, Mann-Whitney U test, 60% \pm 5% of axons). Conversely, in right monocular V1 we found a bias for axons whose activity correlated with right turns (**Figure 3.4B**; right hemisphere: 0.36 vs 0.64 \pm 0.04, $p = 2 \times 10^{-5}$, Mann-Whitney U test, 58% \pm 10% of axons). There are two possible interpretations of these observations. One is that A24b/M2 axons innervating a particular part of V1 are activated more strongly by the movements that increase visual flow in the corresponding area of visual space. The other is that this hemispheric bias is a consequence of a general bias of neural activity to the laterality of the movement, which has been described in motor cortex (Li et al., 2015) and the striatum (Cui et al., 2013). To distinguish between these alternatives, we tested if this bias in activity of A24b/M2 axons depends on visuomotor experience. We trained a different group of mice to perform navigation in left-right inverted virtual reality until they reached a certain performance criterion (up to 10 sessions, 1 session per day, see Experimental Procedures). Mice learned to perform this task despite left-right inversion, albeit slower than under normal conditions (**Figure S3.4A**). We recorded activity of A24b/M2 axons in right monocular V1 on the first and on the last training day. On the first day, we found a bias for axons whose activity correlated with ipsiversive turns, similar to the bias observed in mice trained in the non-inverted environment (**Figure 3.4C**). However, after mice learned to navigate the left-right inverted virtual environment, we found an inverse bias of the activity of A24b/M2 axons; most of the task-related axons in right V1 correlated best with a contraversive turn, which under left-right inversion maximizes visual flow in the left visual field. This shift in preference for contraversive turns was larger in mice with higher performance in the left-right inverted environment (**Figure 3.4D**). To test whether the shift in preference was specific to the navigation task in which visual feedback is coupled to the mouse's movement, or is simply a reversal in the preference of the motor-related or the visually-driven activity, we exposed a subset of the mice to both a no-visual flow condition and to an independent visual flow condition before and after training. In the no-visual flow condition, the walls of the corridor in the virtual environment were uniformly gray but the mouse was free to run on the spherical treadmill. When measured in the absence of visual flow coupled to the movement of the mouse, we found no reversal of the preference in either mice trained under normal conditions or under left-right inverted conditions (**Figure 3.4E**). Similarly, when measured in the independent

visual flow condition, which consisted of the mouse watching a playback of visual flow, the preference of A24b/M2 axons for either ipsi- or contraversive visual flow equalized but did not revert in mice trained under left-right inverted conditions and was unchanged in mice trained under normal conditions after training (**Figure S3.4B**). Hence, the reversal of preference is strongest in conditions in which visual flow is coupled to movement. In sum, activity in A24b/M2 axons in a particular retinotopic location in V1 is dynamically adapted to changes in visuomotor coupling with learning in such a way that A24b/M2 axons are activated most strongly by movements that maximize visual flow in the corresponding part of the visual space.

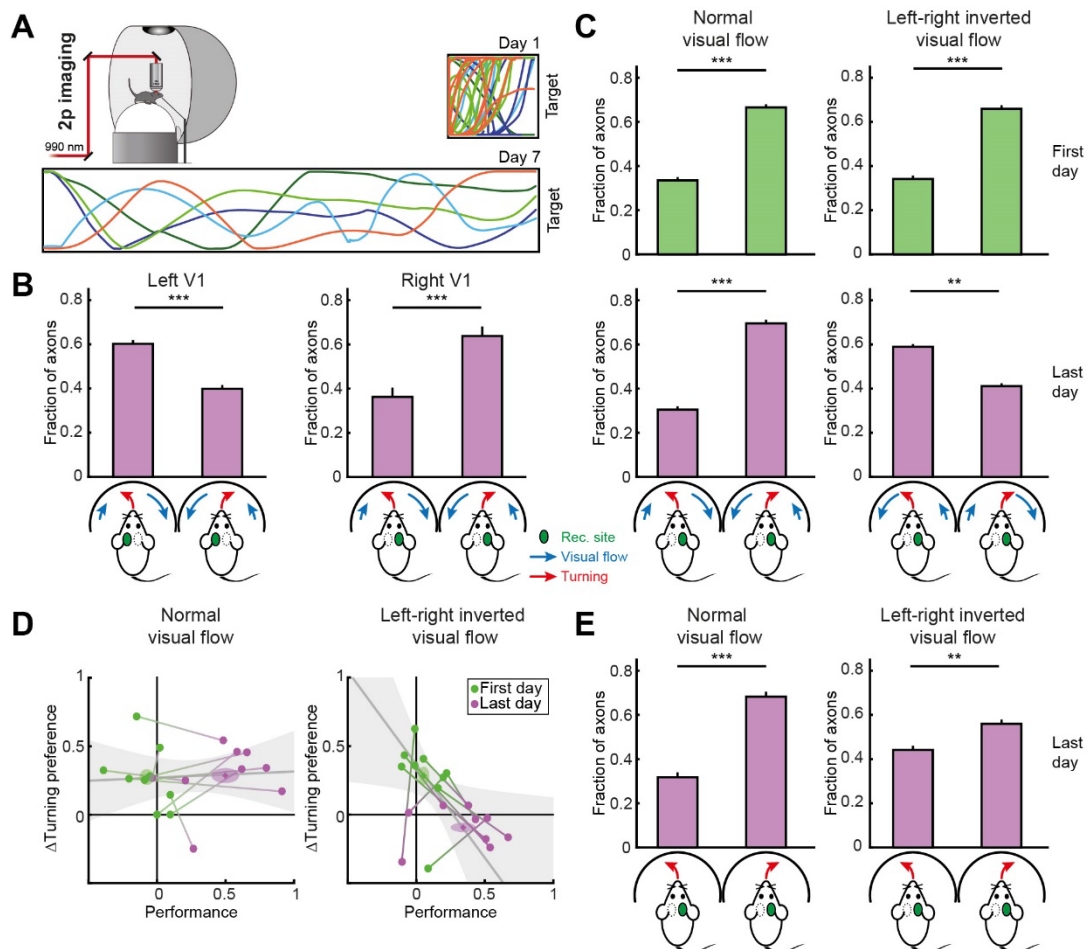


Figure 3.4. A24b/M2 input correlates with the expected visual feedback given a motor output.

(A) Top left: Schematic of the virtual reality setup used to test mice in a 2D navigation task while imaging GCaMP6f labeled A24b/M2 axons in V1. Top right: Five example trajectories of a mouse from the first training session on day 1. Bottom: Five example trajectories of a mouse from the last training session on day 7. Length of the tunnel is not shown to scale, the length/width ratio on day 1 was 5, and on day 7 it was 22. Length of the tunnel was increased as the mouse's performance increased. **(B)** Fractions of axons that correlate best with a turn to the left or a turn to the right in either left or right monocular V1 ($22\% \pm 1.4\%$ mean \pm SEM of axons per recording site correlated with behavior, only these axons were included in the analysis shown here and in (C); left V1: $p = 4 \times 10^{-13}$, 16 sites in 8 mice; right V1: $p = 3 \times 10^{-5}$, 40 sites in 9 mice; Mann-Whitney U test). Note that activity in A24b/M2 axons in either hemisphere correlated best with the turn that maximizes visual flow in the corresponding retinotopic location. Activity was recorded during the last training session. Here and in the following panel, error bars indicate SEM, and in the schematic below, the red arrow indicates the turning behavior of the mouse, the blue arrows indicate the direction and magnitude of the resulting visual flow, and the green region marks the hemispheric location of V1 recordings. **(C)** For mice trained with normal visual flow, activity in A24b/M2 axons in right V1 correlated best with right turns across learning (first day: $p = 9 \times 10^{-13}$, 47 sites in 9 mice; last day: $p = 2 \times 10^{-13}$, 47 sites in 9 mice; Mann-Whitney U test). For mice trained with left-right inverted visual flow, activity in A24b/M2 axons in left V1 correlated best with left turns across learning (first day: $p = 9 \times 10^{-13}$, 47 sites in 9 mice; last day: $p = 2 \times 10^{-13}$, 47 sites in 9 mice; Mann-Whitney U test).

¹⁰, 37 sites in 6 mice; Mann-Whitney U test). In mice trained in a left-right inverted virtual reality, activity in most A24b/M2 axons in right V1 initially correlated best with right turns, but after learning activity in most A24b/M2 axons correlated best with left turns (first day: $p = 1 \cdot 10^{-10}$, 39 sites in 9 mice; last day: $p = 0.002$, 23 sites in 5 mice; Mann-Whitney U test). Note that left turns maximize visual flow on the left of the mouse in the left-right inverted virtual reality. **(D)** Learning related changes in the A24b/M2 axon turning preference as a function of behavioral performance. For both groups of mice (normal trained and left-right inverted trained) the difference between the fraction of axons best correlated with ipsiversive turns and the fraction of axons best correlated with contraversive turns (defined as the turning preference) was positive (more axons best correlated with ipsiversive turns) on the first day of training (green dots). This turning preference became increasingly negative (more axons best correlated with contraversive turns) with improved behavioral performance (last day, magenta dots) for mice trained in left-right inverted virtual reality, but not for normally trained mice. Each pair of dots corresponds to one mouse. Shaded ovals mark mean and SEM across mice. The grey line is a linear fit to the data (Deming regression), and the shading marks the 95% confidence interval of the fit. **(E)** For both groups of mice (normal trained and left-right inverted trained), activity in A24b/M2 axons in right V1 correlated best with right turns after training in the absence of visual flow (normal trained animals: $p = 4 \cdot 10^{-4}$, 11 sites in 3 mice; left-right inverted animals: $p = 0.002$, 16 sites in 4 mice; Mann-Whitney U test).

Finally, to test if artificial activation of A24b/M2 axons in V1 could affect the behavior of the mouse we expressed channelrhodopsin-2 in A24b/M2 by injecting an AAV2/1-Ef1a-CatCh-eGFP into both left and right A24b/M2, and trained mice to navigate the virtual tunnel. On the seventh day of training we briefly stimulated axons (3 s at 25 Hz, see Experimental Procedures) in left, right, or both left and right monocular V1 as mice were navigating to the end of the tunnel. To control for unspecific effects of the stimulation on behavior we compared the stimulation of left or right V1 to a bilateral stimulation of both left and right V1, as A24b/M2 axon stimulation always led to a marked decrease in running speed (**Figure S3.4C,D**). In response to stimulation of A24b/M2 axons in left V1, mice turned to the right. Vice versa, in response to the stimulation of A24b/M2 axons in right V1, mice turned to the left (**Figure 3.5**). In summary, artificial activation of A24b/M2 axons in either left or right monocular V1 resulted in a behavioral response consistent with the bias in their activity we observed during voluntary turning behavior (**Figure 3.4**).

Discussion

Cortical function can be described in a representation framework that is based on the notion of the feature detector (Barlow, 1953; Marr, 1982) or in a predictive processing framework (Clark, 2013; Friston, 2005; Hawkins and Blakeslee, 2004). The predictive processing framework postulates that feedforward sensory signals are compared against a top-down prediction of sensory feedback to detect prediction errors. Evidence for the predictive processing framework has come amongst other things from the discovery of sensorimotor mismatch signals in sensory areas of cortex (Attinger et al., 2017; Eliades and Wang, 2008; Keller and Hahnloser, 2008; Keller et al., 2012; Zmarz and Keller, 2016) that could constitute prediction errors. Mismatch signals in mouse visual cortex are the result of a comparison of an excitatory motor-related prediction and an inhibitory visual input (Attinger et al., 2017). A missing piece of the puzzle has been evidence for top-down predictions of sensory input given a movement. We now show that A24b/M2 provides a strong and dense projection to V1, which conveys motor-related signals that depend on the mouse's visuomotor experience. We

suggest that these signals constitute a prediction of visual flow that is linked to the mouse's movement, consistent with an interpretation of the function of visual cortex in a predictive coding framework. This is detailed in the following:

1. A motor-based prediction of visual flow in V1 requires an anatomical basis in the form of either a direct or an indirect projection from a motor-related area of the brain. We found that one of the largest inputs to V1 originates in A24b/M2 (**Figure 3.1C**), a motor-related area of cortex (Sinnamom and Galer, 1984).
2. Visuomotor mismatch signals have been reported in layer 2/3 neurons of V1 (Keller et al., 2012; Saleem et al., 2013). These mismatch signals can be explained by a difference between an excitatory prediction of visual flow and an inhibitory visual signal (Attinger et al., 2017; Zmarz and Keller, 2016). We found that layer 2/3 excitatory neurons are one of the main targets of the projection from A24b/M2 (**Figure 3.1E,F**).
3. Movements lead to visual feedback that systematically changes with retinotopic location. Moving head or eye to the left leads to full field visual flow to the right. Moving forward, however leads to radially symmetric visual flow of opposite direction in the left and right visual field. Thus to simplify wiring one would expect a certain amount of conservation of topography in the process of a conversion from a myotopic coordinate system to a retinotopic coordinate system. We find that the projection from A24b/M2 to V1 coarsely maintains topography (**Figure 3.2E**).
4. To convey a prediction of visual input, the bandwidth of the projection would need to be roughly equivalent to the feedforward visual input from the dorsolateral geniculate nucleus (dLGN). The number of neurons in A24b/M2 providing input to V1 was roughly twice as large as that in dLGN (**Figure 3.1C**). In addition we found that the projection from A24b/M2 to V1 is dense in that it targets more than 75% of all neurons in layer 1, 2/3 and 6 (**Figure 3.2F,G**).
5. Motor-related predictions of visual input should occur mainly during movements that elicit visual feedback, and should not be activated by visuomotor mismatch. We found that A24b/M2 axons in V1 are strongly activated during locomotion and that activity precedes the onset of locomotion but remain silent during mismatch (**Figure 3.3B**).
6. The amplitude of the predictive signal should correlate with the magnitude of the predicted change in visual flow given the current visual stimulus. We found that running onset activity in A24b/M2 axons was significantly reduced if the running onset occurred either during darkness or during open-loop conditions with visual flow preceding running onset (**Figure 3.3C**). In both cases, the predicted change in visual flow resulting from a running onset is

lower, either due to a lack of visual information, or due to ongoing visual flow prior to the running onset.

7. An inhibition of an excitatory predictive signal should result in a reduction of both motor-related activity and mismatch signals in V1 neurons that respond strongly to running or mismatch. We found that inhibition of A24b/M2 reduces both motor-related and mismatch signals in running and mismatch responsive V1 neurons (**Figure 3.3H,I**).
8. If predictions are transmitted in visual coordinates, they should be stronger in a given retinotopic location in V1 for behaviors that maximize visual flow in the corresponding part of the visual field. We found that activity in A24b/M2 axons in monocular V1 correlated better with ipsiversive turns, which maximize visual flow in the contralateral visual field (**Figure 3.4B**).
9. A key premise of predictive coding theories is that predictions are continuously updated based on prediction errors. Thus, if there are changes in the way visual feedback is coupled to motor output, predictions of visual feedback should adapt to reflect the new form of visuomotor coupling. We found that in mice trained in left-right inverted visual environment, activity of A24b/M2 axons in right V1 correlated better with left turns, consistent with the fact that in this left-right inverted virtual environment, left turns maximize visual flow in the left visual field. (**Figure 3.4C**). This change was larger in mice that had a higher performance in left-right inverted navigation (**Figure 3.4D**), and was smaller for both correlation with running and visual flow separately (**Figure 3.4E**, and **Figure S3.4B**).
10. Artificial stimulation of a projection that conveys a prediction of visual flow in a particular retinotopic location in V1 should result in illusory visual flow in the corresponding part of the visual field. We found that stimulating A24b/M2 axons in either right or left monocular V1 resulted in a corrective behavior of the mouse consistent with an adaptation to an illusory visual flow (**Figure 3.5C**).

While many of the individual findings have alternate interpretations, the interpretation of the signals as a prediction of visual flow provides a unified explanation of all our findings. Other explanations likely raise more questions than they answer. For example, the finding that stimulation of A24b/M2 axons results in a turning behavior, could be explained by assuming that activity in A24b/M2 directly drives movement. However, this would not explain why neurons in A24b/M2 that directly drive movement would send collaterals to V1. Another concern with the interpretation of the signals as predictions of visual feedback may stem from the fact that activity in A24b/M2 axons in V1 is not absent in darkness. Here one should consider that V1 layer 2/3 neurons also exhibit motor-related activity in complete darkness (Keller et al., 2012; Saleem et al., 2013) that is persistent even under

complete retinal lesions (Keck et al., 2013). This motor-related activity cannot, strictly speaking, be explained in either the representation framework or the predictive processing framework. In complete absence of visual input, there is no visually driven activity and movement does not result in a predictable visual flow. However, in low light conditions, visual processing may systematically rely on predictions with decreasing strength of visual signals and complete darkness may be a singularity for which the system has not evolved to produce optimal predictions.

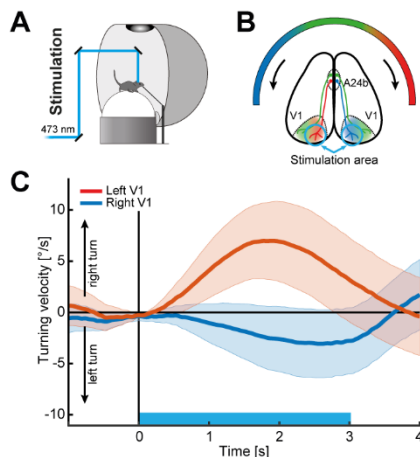


Figure 3.5: Artificial stimulation of A24b/M2 axons in monocular V1 causes turning behavior.

(A) Schematic of the virtual reality setup used for stimulation experiments. **(B)** Schematic top-down view of a mouse brain and the virtual reality screen illustrating the retinotopic mapping of the virtual reality onto V1, as well as the topographic mapping of axons from A24b/M2 to V1. Blue circles indicate areas stimulated with laser. Arrows indicate the direction of visual flow generated during forward locomotion. **(C)** Average turning velocity (7 mice) induced by stimulation of A24b/M2 axons in left (red) or right (blue) V1. Stimulation of the A24b/M2 axons in right (left) monocular V1 in mice navigating towards a target resulted in turning to the left (right). Shading indicates SEM. Blue bar indicates stimulation period.

The projection from A24b/M2 to V1 fulfills all the criteria to be a top-down signal that conveys a motor-related prediction of visual flow in visual coordinates. Nevertheless, it is likely that the same projection also has other functions. For example, it has been shown to be involved in surround suppression of visual responses (Zhang et al., 2014), and we have recently shown that it also conveys a prediction of visual input based on spatial location of the mouse (Fiser et al., 2016). Moreover, given that there are also delayed visual responses in A24b/M2 axons in V1 (**Figure 3B**), our data are consistent with a model in which A24b/M2 conveys a general prediction of visual input given recent experience. It is also likely that A24b/M2 is not the only source of predictive input to V1. Similar predictions could be conveyed by all top-down inputs to V1. Inputs from secondary visual areas, for example, can be described in a predictive coding framework with respect to higher level visual features (Rao and Ballard, 1999). We speculate that top-down input could in general be interpreted as a prediction of feedforward input. In this way, input to V1 from auditory cortex, for example, would provide a prediction of visual input given a learned association between a sound and a visual stimulus, while input from retrosplenial cortex, known to be activated by vestibular signals (Rancz et al., 2015), might provide a prediction of visual input based on a head or body rotation. Consistent

with this we find that A24b/M2 inactivation and lesion only result in a partial reduction of motor-related and mismatch signals in V1. Note that in the framework of predictive coding, also classically visual phenomena like attention and adaptation would find a mechanistic interpretation (Feldman and Friston, 2010; Keller et al., 2017).

It is unclear how these findings of a projection from A24b/M2 to V1 extend to primate anatomy. A24b is likely homologous to primate A24b (Vogt and Paxinos, 2014; Vogt et al., 1987), but a connection from A24b to V1 similar to the one described in rodents (Miller and Vogt, 1984; Vogt and Miller, 1983) has not been described in primates (Vogt et al., 1979). However, using classical neural tracing methods, the projection from ACC to V1 in rats was estimated to be much weaker (Miller and Vogt, 1984; Vogt and Miller, 1983) than what we have found using rabies tracing in the mouse. This means that either classical tracing underestimates long-range cortical projections, or that communication between ACC and visual cortex occurs primarily indirectly in rats and primates. In the human cortex, there is some evidence of functional connectivity between ACC and primary and secondary visual cortices (Zhou et al., 2016).

Predictive coding is a theoretical framework that describes the function of sensory processing in general, and visual cortex in particular. Evidence for predictive coding has mainly come from theoretical considerations (Rao and Ballard, 1999; Spratling, 2010) and the discovery of mismatch signals in primary sensory areas of cortex and avian pallium (Eliades and Wang, 2008; Fiser et al., 2016; Keller and Hahnloser, 2008; Keller et al., 2012). One of the key pieces of evidence lacking for a description of cortical function in a framework of predictive coding is predictive top-down input. We have demonstrated here that the input from A24b/M2 to V1 fulfills all the criteria necessary to be interpreted as a prediction of visual flow given a motor output. Our data are consistent with a theory of predictive coding where the function of cortex is to generate and maintain an internal model of the world (Craink, 1943) by continuously updating and comparing this model to actual sensory input.

Supplementary Information

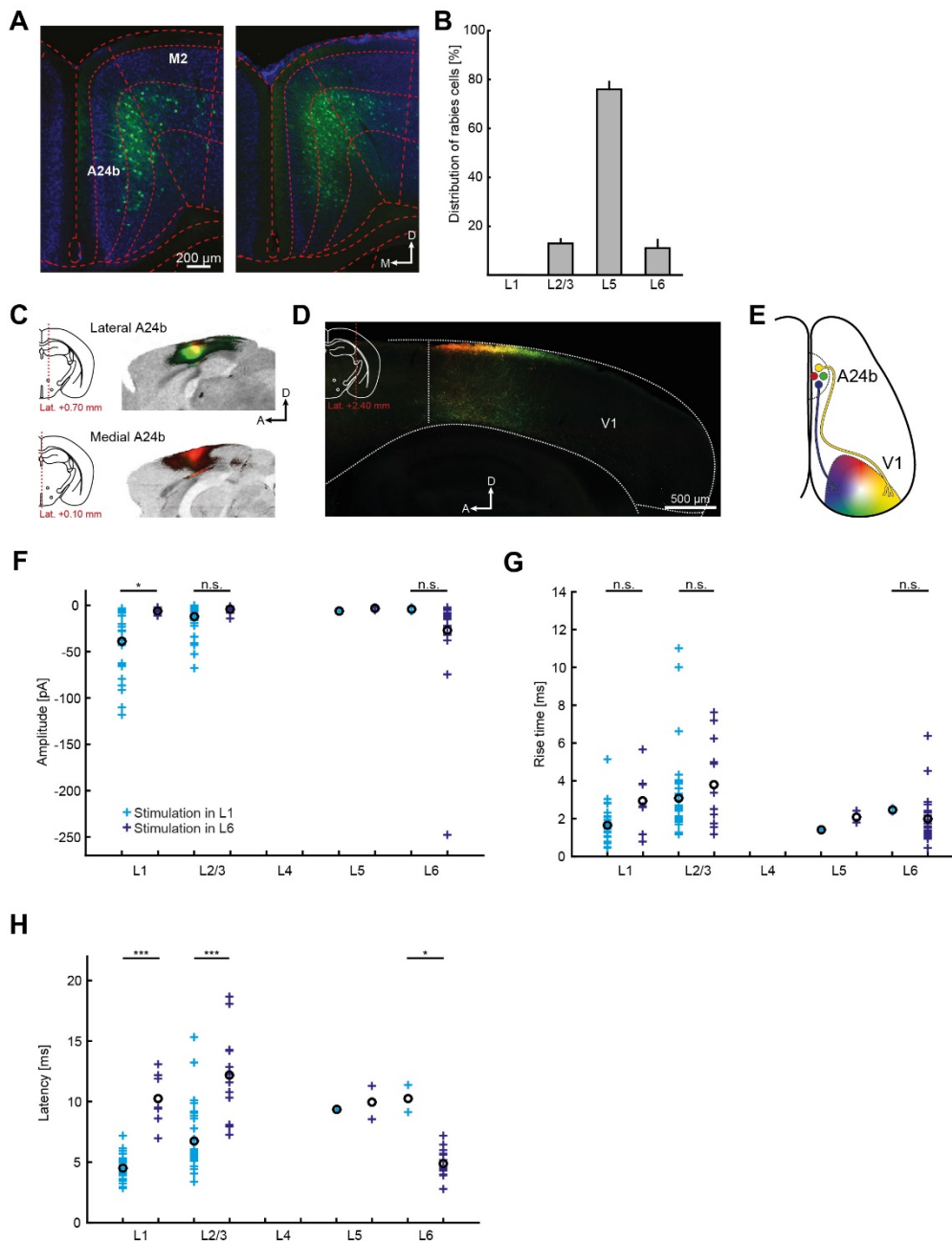


Figure S3.1. Supplementary anatomy and electrophysiology of the A24b/M2 to V1 projection.

(**A**) Coronal sections through A24b/M2 in two mice (C57BL/6J). Cells labeled by rabies tracing from V1 in A24b/M2 (as in **Figure 3.1D**) are shown in green. Area borders and cortical layers are indicated by dashed red lines. (**B**) Most presynaptically labeled cells were located in layer 5 of A24b/M2. Distribution of rabies positive cells in A24b/M2 sorted by layer (5 mice; error bars indicate SEM). (**C**) Injections of an AAV2/1-Ef1a-eGFP into lateral (0.7 mm lateral from midline) A24b/M2 and an AAV-Ef1a-tdTomato into medial (0.1 mm lateral from midline) A24b/M2. (**D**) Sagittal section (2.4mm lateral from midline) showing axonal projections in V1 of the two injections shown in (C). (**E**) Schematic of the mapping of the A24b/M2 projection to V1. Medial (lateral) A24b/M2 projects to anterior (posterior) V1, and anterior (posterior) A24b/M2 projects to lateral (medial) V1. Note, the exact path of the projection is illustrative only. Axons from A24b/M2 to V1 descend into the white matter below A24b/M2, traverse parallel to the midline in the posterior direction and then traverse lateral towards V1 where they ascend again into the gray matter. (**F**) Amplitudes of EPSCs induced by stimulation of A24b/M2 axons in neurons with a significant stimulation response (**Figure 3.2G**), by layer and stimulation site. Crosses

show data of single neurons; circles indicate mean. Note that analysis in this and the following panels only included neurons that exhibited significant stimulation responses. L1: $p = 0.012$; L2/3: $p = 0.209$; L6: $p = 0.130$, Mann-Whitney U test. **(G)** EPSC rise times were measured as the time from 20% to 80% of peak response. Crosses and circles as in (F). L1: $p = 0.072$; L2/3: $p = 0.400$; L6: $p = 0.192$, Mann-Whitney U test. **(H)** Latency to 20% rise from start of stimulation. Crosses and circles as in (F). L1: $p = 6.41 \cdot 10^{-5}$; L 2/3: $p = 2.6 \cdot 10^{-5}$; L6: $p = 0.025$, Mann-Whitney U test.

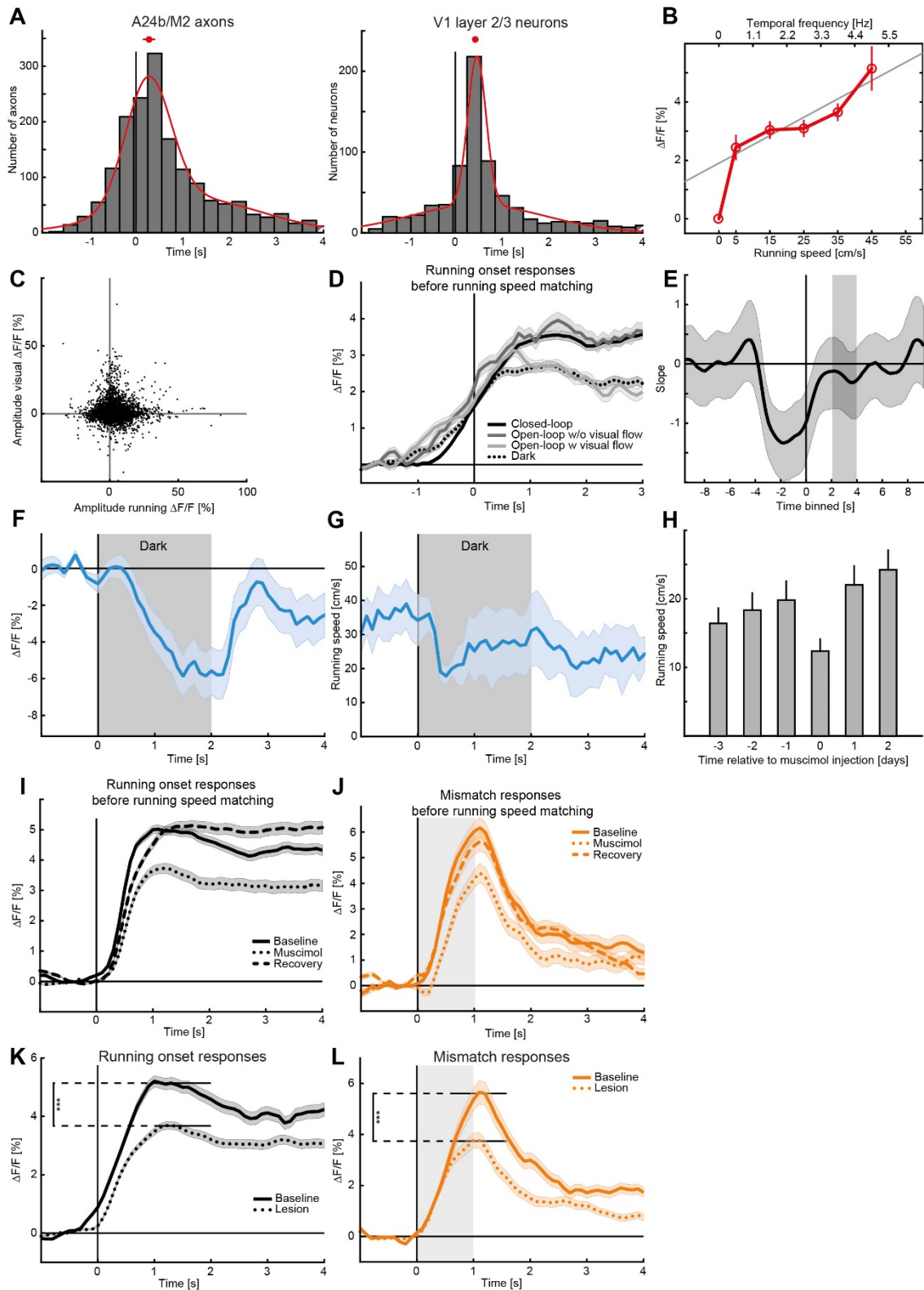


Figure S3.2. Suppressive effect of visual flow and running speed variability compensation.

(A) Distributions of the onset of activity relative to running onset in A24b/M2 axons in V1 (left) and layer 2/3 neurons in V1 (right). Red line is a sum-of-two-Gaussians fit to the distribution. Red dot and horizontal line indicate the mean peak of the fit and SEM (A24b/M2 axons: 0.28 ± 0.13 s, V1 neurons: 0.43 ± 0.06 s). **(B)** The average response in A24b/M2 axons to running onset in closed-loop conditions increases with increasing running speed. Gray line is a linear fit to the data, not including the zero bin. Axis on top of the panel indicates the visual flow speed (as temporal frequency of the grating) corresponding to the running speed. Error bars indicate SEM. **(C)** Scatter plot of the average running onset response in

darkness against the average visual flow onset response in open-loop for all A24b/M2 axons. We found no evidence of a correlation between the running onset response and the visual onset response of A24b/M2 axons in V1 (Pearson correlation, $R = -0.017 \pm 0.023$, mean \pm SEM, 10 mice). **(D)** Same as **Figure 3.3C**. Running onset responses are larger in closed-loop and open-loop conditions without visual flow than dark or open-loop with visual flow conditions also without running speed running speed correction in a window from 0.5 s to 2 s post running onset for the data shown in **Figure 3.3C**. Shading indicates SEM. **(E)** Reduction of running onset response (measured in the window from 2 s to 4 s after running onset, gray shading) as a function of visual flow at a given time (see STAR Methods). The strongest inhibition of running onset activity results from visual flow immediately preceding running onset (approx. -3 s to 0 s). Shading indicates the 95% confidence interval. **(F)** Average activity in A24b/M2 axons in V1 after an onset of darkness (at 0 s). Note, activity decreases and rebounds after lights are turned on again (at 2 s). Gray shading indicates period of darkness. **(G)** As in **(F)**, but for running speed. Note, there is a decrease in running speed concurrent with the drop in activity, however, this likely does not fully account for the darkness induced reduction in activity as the rebound of activity in A24b/M2 axons after lights on is not accompanied by a concurrent increase in running speed. **(H)** Average running speed as a function of the number of days relative to the muscimol inactivation of A24b/M2. Muscimol inactivation of A24b/M2 results in a reduction in average running speed on the day of the injection. **(I)** Same as **Figure 3.3H**. Average population responses to running onsets before, during and after muscimol injection in A24b/M2 for data without running speed matching. The decrease and recovery of activity following muscimol injections matches the results for running speed matched data (**Figure 3.3H**). **(J)** Same as **Figure 3.3I**. Average population responses to feedback mismatch before, during and after muscimol injection in A24b/M2 for data without running speed matching. The decrease and recovery of activity following muscimol injections matches the results for running speed matched data (**Figure 3.3I**). Gray shading indicates the duration of the feedback mismatch. **(K)** Ibotenic acid lesion of A24b/M2 reduces average population response of running onset responsive neurons (50 % most running onset responsive neurons selected on preceding time points; 921 of 1842 neurons; see STAR Methods) in V1. Shown are running onset responses before (baseline; solid line), and after (lesion; dotted) inactivation of A24b/M2 ($p = 5 \times 10^{-10}$; Student's t test). **(L)** Ibotenic acid lesion of A24b/M2 reduces average population response of mismatch responsive neurons (50 % most mismatch responsive neurons selected on preceding time points; 921 of 1842 neurons; see STAR Methods) in V1. Shown are mismatch responses before (baseline; solid line), and after (lesion; dotted) inactivation of A24b/M2 ($p = 2 \times 10^{-4}$, Student's t test). Gray shading indicates the duration of the feedback mismatch.

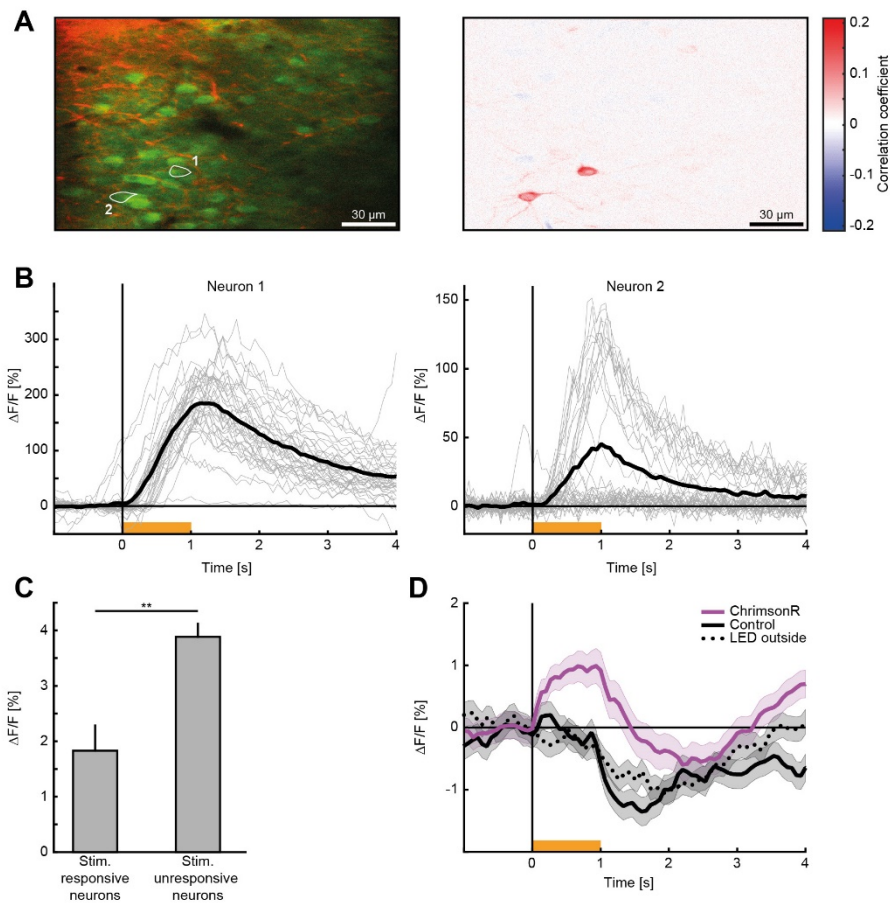


Figure S3.3. *In vivo* channelrhodopsin-2-assisted circuit mapping (ivCRACM).

(A) Left: Example two color two-photon image showing A24b/M2 axons labeled with ChrimsonR-tdTomato (red) and layer 2/3 neurons labeled with GCaMP6f (green) at the boundary of layer 1 and layer 2/3. Right: Pixel-wise correlation of GCaMP6f activity with stimulation of A24b/M2 axons. The two neurons marked on the left exhibited a strong positive correlation with the stimulation of A24b/M2 axons. **(B)** Example responses to the stimulation of A24b/M2 axons of the two neurons shown in (A). Orange bar indicates the duration of the stimulation. **(C)** Average activity of neurons that correlated with axonal stimulation (82 neurons; see STAR Methods) is lower than for neurons that did not correlate with axonal stimulation (566 neurons; $p = 0.0016$, Student's *t* test). Error bars indicate SEM. **(D)** Average population response of V1 neurons to stimulation of A24b/M2 axons (purple line; 648 neurons). In animals that did not express ChrimsonR in A24b/M2 axons, there was no stimulation response (black line; 1493 neurons). The stimulation induced decrease in activity is likely a visual response to the stimulation light, as it was also present when using a stimulation LED placed outside of the craniotomy and light path of the microscope (LED outside; dotted black line). Shading indicates SEM.

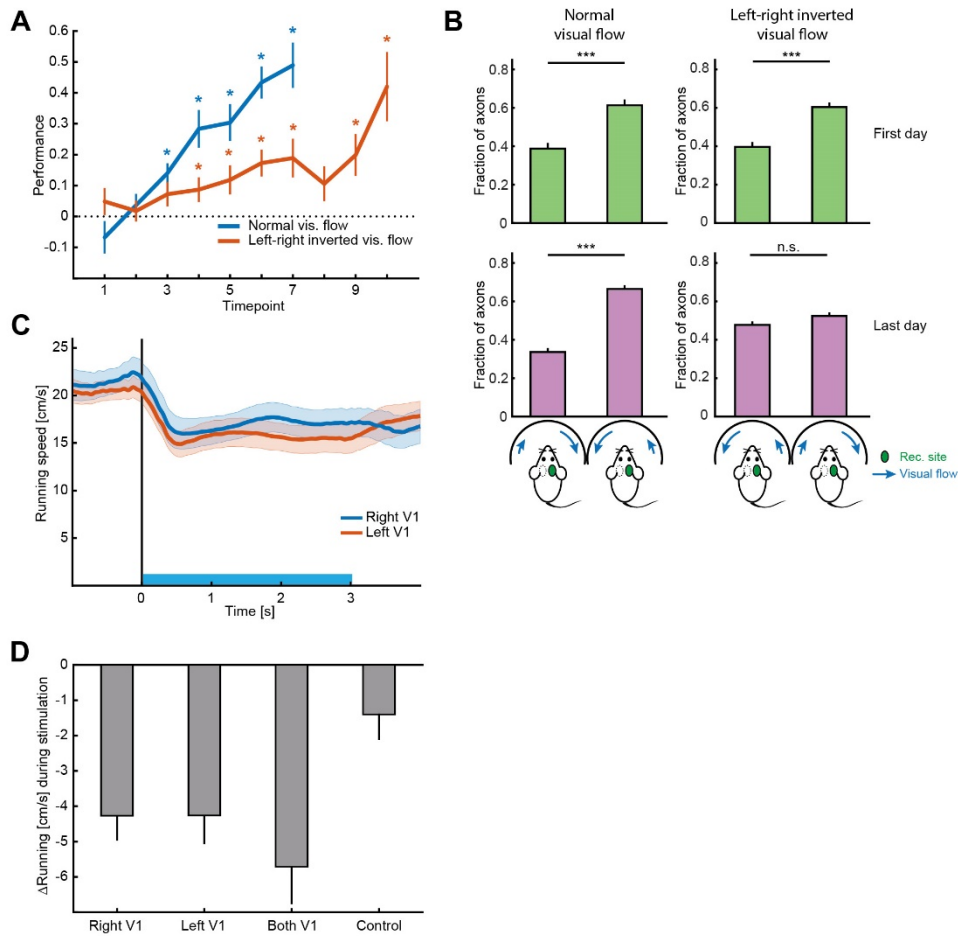


Figure S3.4. Learning curves for normal and left-right inverted trained mice and running speed decrease after stimulation of A24b/M2 axons in V1.

(A) Mice trained to navigate to the end of a virtual tunnel in a left-right inverted task (red line; 9 mice) matched the performance but learned slower than mice trained in a normal virtual environment (blue line; 9 mice). Performance is measured as the fraction of distance travelled towards the target normalized by the total distance travelled (A random walk, this results in a performance of 0, while movement in a straight line towards the target results in a performance of 1) * $p < 0.05$, one-sample Student's t test. Error bars indicate SEM. **(B)** Visual flow generated in the preceding closed-loop session was replayed to the mouse on the first and last day of training. For both groups of mice (normal trained and left-right inverted trained), activity in A24b/M2 axons in right V1 correlated best with the playback of the visual flow of right turns before training. After training the difference in preference was unchanged for normally trained mice but was no longer present for left-right inverted trained mice (normal trained mice, first day: $p = 5 \times 10^{-8}$, 24 sites in 5 mice, last day: $p = 7 \times 10^{-6}$, 16 sites in 3 mice; left-right inverted animals, first day: $p = 2 \times 10^{-8}$, 21 sites in 6 mice; last day: $p = 0.70$, 15 sites in 4 mice; Mann-Whitney U test). **(C)** Running speed is reduced upon stimulation of channelrhodopsin-expressing A24b/M2 axons in both right and left monocular V1. Blue bar indicates the duration of the stimulation. Shading indicates SEM. **(D)** Reduction of running speed is also apparent when stimulating A24b/M2 axons in both right and left monocular V1 or sham stimulation (cranial windows are covered with black tape). Thus, a part of this reduction in running speed is likely due to a combination of a startle response to the noise of the galvanometer changing position and a subtle change in lighting. The laser is always on and moved between the cranial windows and a blank position on the head bar located between the left and right cranial window. Error bars indicate SEM.

Methods

Contact for Reagent and Resource Sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Georg Keller (georg.keller@fmi.ch).

Experimental Model and Subject Details

Animals and surgery. Rabies tracing experiments (**Figure 3.1**) were performed on males and females of nine different mouse lines: C57BL/6J mice (n = 5), PV-Cre (Pvalb^{tm1(cre)Arbr}; n = 3) mice, VIP-Cre (Vip^{tm1(cre)Zjh}, n = 3), SST-Cre (CR^{tm1(cre)Zjh}; n = 4), CR-Cre (n = 3), Wfs1-Cre mice (Wfs1-Tg2-CreERT2; n = 3), Scnn1a-Cre (n = 4), 3 Rbp4-Cre (Rbp4^{KL100Gsat/Mmucd}; n = 3), Ntsr1-Cre (Ntsr1^{GN220Gsat/Mmucd}; n = 4). All mice used were between 71 and 370 days old. For all other experiments a total of 88 female C57BL/6J mice were used, 8-10 weeks old at the start of the experiment. All experiments were performed on healthy mice which were not used for prior procedures. Mice were group-housed in a vivarium (light/dark cycle: 12/12 hours). Female adult mice were mainly used to reduce aggression and dominance during group housing. All animal procedures were approved by and carried out in accordance with guidelines of the Veterinary Department of the Canton Basel-Stadt, Switzerland.

Method Details

Surgery. For all surgical procedures mice were anesthetized using a mix of fentanyl (0.05 mg/kg), medetomidine (0.5 mg/kg) and midazolam (5 mg/kg). Cranial windows were implanted as previously described (Keller et al., 2012; Leinweber et al., 2014).

Rabies tracing (Figure 3.1 and Figure S3.1). To initiate monosynaptic rabies tracing, either AAV2/1-Ef1a-TVA950-T2A-CVS11G or AAV2/1-Ef1a-DIO-TVA950-T2A-CVS11G (titer $2 \cdot 4 \cdot 10^{11}$ GC/ml) was injected into V1 of C57BL/6J mice or of mice of different Cre lines, respectively, through a small craniotomy. Following the injection, the craniotomy was sealed with cyanoacrylate and the skin sutured. After 2-4 days the seal was removed and an EnvA-coated, glycoprotein-G deleted GCaMP6s rabies virus (referred to as EnvA-SADΔG-GCaMP6s, titer $1.5 \cdot 10^9$ TU/ml, FACS titered (Wertz et al., 2015; Wickersham et al., 2010)) was injected at the same location. For Wfs1-Cre mice, Cre expression was induced by intraperitoneal injection of 100 μl tamoxifen (20mg/ml in saline) twice within a 24-48 h interval 4 days after the AAV injection. Mice were sacrificed for histological staining 5-6 days after injection of the rabies virus. Mice were transcardially perfused for 10 min with phosphate buffered saline (PBS), followed by 10 min of perfusion with a solution of 4% paraformaldehyde (PFA) in PBS. Brains were then isolated and post-fixed for 24h in 4% PFA in PBS. The PFA solution was then exchanged for 30% sucrose in PBS, in which brains were immersed until they sank. Afterwards, the brains were transferred to embedding medium (Tissue-Tek), frozen on

dry ice and stored at -80°C before they were sectioned into $80\ \mu\text{m}$ coronal sections using a cryostat. Brain sections were placed in well-plates separately and kept free-floating in PBS. Each section was triple immunostained for rabies derived GFP (Abcam 13970), peptide linker 2A expressed by the AAV (Millipore #ABS31) and a neuron-specific nuclear protein NeuN (Millipore MAB377). The conjugated fluorescent labels of secondary antibodies were Alexa Fluor 488 (Jackson Immuno #703-545-155), Alexa Fluor 568 (ThermoFisher #A10042) and Alexa Fluor 647 (ThermoFisher #A31571), respectively. Stained sections were subsequently mounted on microscope slides and imaged using the ZEISS Axio Scan.Z1 slide scanner with 10x magnification. These images were then used to manually count the starter cells and rabies labeled cells throughout the brain. All the brain areas were defined as in (Franklin and Paxinos, 2012).

AAV tracing (Figure 3.2A-E and Figure S3.1C-E). A small craniotomy was made over A24b (centered on $0.3\ \text{mm}$ lateral and 0.5 anterior of bregma) and AAV2/1-Ef1a-tdTomato-WPRE (titer 2.4×10^{11} GC/ml) and AAV2/1-Ef1a-eGFP-WPRE (titer 2.2×10^{11} GC/ml) were injected around bregma separated by $0.5\ \text{mm}$ either along the anterior-posterior axis or the medio-lateral axis. The craniotomy was sealed with cyanoacrylate and the skin sutured. Three weeks later mice were sacrificed and perfused. Brain sections ($80\ \mu\text{m}$) were cut on a cryostat and mounted on microscope slides and imaged using the ZEISS Axio Scan.Z1 slide scanner with 10x magnification (Figure 3.2B,C), or with a confocal microscope (Figure 3.2D).

CRACM (Figure 3.2F,G and Figure S3.1F-H). Adult (P90 – P150) C57BL/6J mice received AAV2/1-Ef1a-CatCh-eGFP (titer 2.3×10^{11} GC/ml) injections into the right A24b. After 3 - 5 weeks, mice were anesthetized with isoflurane and sacrificed. The brain was removed and coronal slices from right V1 prepared in ice-cold slicing solution (in mM: 87 NaCl, 2.5 KCl, 7 MgCl₂, 1.25 NaH₂PO₄, 25 NaHCO₃, 25 glucose, 75 sucrose, bubbled with 95% O₂/5% CO₂). $300\ \mu\text{m}$ thick slices were cut using a Campden Instruments 700 SMZ vibrating microtome. Slices were stored in ACSF (in mM: 125 NaCl, 2.5 KCl, 2 CaCl₂, 1 MgCl₂, 1.25 NaH₂PO₄, 26 NaHCO₃, and 25 glucose, pH 7.3, equilibrated with 5%CO₂/ 95%O₂) for 30 minutes at 35°C followed by up to 6 hours at room temperature until use. To quantify synaptic input from A24b to V1, whole-cell patch-clamp recordings were made from neurons in the visual cortex at room temperature ($22 - 24^{\circ}\text{C}$). The intracellular solution contained (in mM): 120 potassium gluconate, 5 mM NaCl, 2 MgCl₂, 0.1 CaCl₂, 10 HEPES, 1.1 EGTA, 4 magnesium ATP, 0.4 disodium GTP, 15 sodium phosphocreatine and 0.1% Biocytin (pH 7.3 with KOH; 290 mOsm). The borders of the cortical layers were identified by cortical depth, neuron morphology, and the current responses of excitatory neurons. Layer 1 was identified by its low cell density. In early experiments identification of layer 4 and layer 5 was facilitated with the use of Rbp4-Cre mice injected with an AAV2/1-Ef1a-DIO-eGFP in V1.

A24b axons were activated selectively in layer 1 or layer 6 under a 60x 0.9 NA Olympus water immersion objective with 2 ms pulses from a blue LED at a power of 0.42 mW. Stimulation intensity was far below the threshold to trigger action potentials in postsynaptic neurons; the threshold for eliciting postsynaptic spikes was consistently above 4.0 mW. Postsynaptic responses were recorded at a holding potential of -70 mV. To determine the source of inputs, the area of stimulation centered on the region above or below the neuron in layer 1 and layer 6. No responses were observed when the stimulation was centered directly adjacent to layer 1 outside of cortex. In untransfected control mice LED stimulation did not evoke responses at stimulation intensities of 4.0 mW. All recordings were acquired and analyzed with IgorPro (WaveMetrics) using NeuroMatic.

Axonal imaging (Figure 3.3A-C, Figure 3.4 and Figure S3.2A-G). A craniotomy was made over V1 and sealed with a 4 mm cover slip as previously described (Leinweber et al., 2014). Additionally, a small craniotomy was made over A24b (ipsilaterally, 0.3 mm lateral of bregma) and AAV2/1-Ef1a-GCaMP6f-WPRE (titer 1.0×10^{11} - 1.8×10^{12} GC/ml) was injected before the region was sealed with cyanoacrylate. A titanium head bar was fixed to the skull with dental cement. Prior to axon imaging, V1 was mapped with intrinsic signal imaging, as previously described (Wertz et al., 2015). Four weeks post injection, calcium signals in A24b axons were imaged using a modified Thorlabs B-Scope, as previously described (Leinweber et al., 2014). Illumination source was a pulsed infrared laser (Coherent Vision S) tuned to a wavelength of 990 nm. We used an 8 kHz resonance scanner (Cambridge Technology) for line scanning, which enabled frame rates of 40 Hz at 400 x 750 pixel resolution. In addition, we used a piezo actuator (P-726 PIFOC, Physik Instrumente) to move the objective (Nikon 16x, 0.8 NA) in steps of 15 μ m between frames to acquire images at 4 different depths, thus reducing the effective frame rate to 10 Hz.

The behavioral virtual reality setup was as previously described (Leinweber et al., 2014). Briefly, head-fixed mice were free to run on an air-supported polystyrene ball. The rotation of the ball was restricted to the forward and backward directions using a pin, and was coupled to linear displacement in the virtual environment projected onto a toroidal screen surrounding the mouse. The screen covered a visual field of approximately 240 degrees horizontally and 100 degrees vertically. The virtual environment presented on the screen was a virtual tunnel with walls consisting of continuous vertical sinusoidal gratings. Mice were first exposed to a closed-loop condition, during which motion of the grating was coupled to the mouse's locomotion on the ball, and then exposed to an open-loop condition during which the visual flow of a previous closed-loop session was replayed to the mouse. The open-loop condition was followed by a dark session where all light sources including the virtual reality were turned off. To measure responses of A24b/M2 axons to the onset of darkness, all light sources were turned off for period of 2 s in a closed-loop session, in a

subset of experiments. Additionally, visual feedback was perturbed during closed-loop conditions by stopping the gratings for a brief period (1 s) at random times, as described previously (Keller et al., 2012).

***In vivo* channelrhodopsin-2-assisted circuit mapping (ivCRACM) (Figure 3.3D-F and Figure S3.3).**

Viral injections and surgery were performed as described for the axon imaging with the exception that an AAV2/1-hSyn-ChrimsonR-tdTomato-WPRE (titer 3.0×10^{12} GC/ml) was injected into A24b and AAV2/1-Ef1a-GCaMP6f-WPRE (titer 1.0×10^{11} GC/ml) into V1. ChrimsonR stimulation and functional imaging of GCaMP6f expressing neurons was done by using a modified Thorlabs B-Scope.

Illumination source for two-photon imaging was a pulsed infrared laser (Spectra physics) tuned to a wavelength of 990 nm. We used a 12 kHz resonance scanner (Cambridge Technology) for line scanning, which enabled frame rates of 60 Hz at 400 x 750 pixel resolution. In addition, we used a piezo actuator (Physik Instrumente) to move the objective (Nikon 16x, 0.8 NA) in steps of 15 μ m between frames to acquire images at 4 different depths, thus reducing the effective frame rate to 15 Hz. Illumination source for ChrimsonR stimulation was a fast LED (UHP-T-595, Prizmatix) with a wavelength of 595 nm, which allowed for fast TTL triggered operation. Stimulation lasted for 1 s (20 Hz; 50% duty cycle) with an average power of 15 mW at random times during a closed-loop condition. For spectral filtering we used a dichroic mirror (ZT775sp-2p, Chroma) to combine the two-photon laser and stimulation light. A second long-pass dichroic mirror (F38-555SG, Semrock) was used to split the GFP emission from both illumination light sources. Light leak from the 595 nm stimulation LED was reduced by synchronizing the LED light output to the turnaround times of the resonant scanner (during which imaging data were not acquired). Lastly, amplified PMT signals were digitally bandpass filtered at 80 MHz to reduce the effect of ringing. Two-photon imaging of activity was otherwise performed as previously described (Leinweber et al., 2014).

Muscimol or ibotenic acid silencing of A24b (Figure 3.3G-I and Figure S3.2H-L). Viral injections and surgery were performed as described above for axon imaging with the exception that AAV2/1-Ef1a-GCaMP6f-WPRE (titer 2.9×10^{12} GC/ml) was injected into V1. At the time of cranial window implant a small craniotomy over A24b was made and sealed with cyanoacrylate to allow for injections of muscimol or ibotenic acid preceding imaging without further need for surgery. Muscimol (at a concentration of 5 g/l) or ibotenic acid (13.5 g/l) was injected 15 min, or 24h for ibotenic acid, prior to imaging at a volume of 50-100 nl at a depth of 500 μ m.

Virtual reality and navigation task (Figure 3.4, Figure 3.5 and Figure S3.4). Viral injections and surgery were performed as described above for axonal imaging. For training, mice were initially placed into a virtual environment with an infinite tunnel (identical to the one used for axonal

imaging) for two training sessions on two consecutive days to allow the mice to get accustomed to the setup and virtual environment. Rotation of the spherical treadmill was restricted to forward and backward rotation using a pin. During subsequent training, rotation of the ball was no longer restricted and mice were free to rotate in the virtual environment in addition to moving forward and backward. The virtual reality consisted of a tunnel with a reward zone located at one end. The walls of the tunnel were textured with white circles on a black background in the first half of the tunnel and sinusoidal vertical stripes in the second half. Reaching the reward zone triggered a 5 s timeout during which the mouse could lick from a water spout for reward. After the timeout, the virtual environment was reset to the beginning of the tunnel to start the next trial. As the mouse's performance in the task improved, the length of the tunnel was gradually increased to keep the rewards per hour at an approximately constant level of 100, throughout training. Mice were water restricted for the duration of the experimental series and were supplemented with water during early training sessions if they received less than 1 ml of total water reward. The weight of all mice was monitored daily to ensure that body weight would not drop below 80% of starting weight. A subset of mice were also exposed to an open-loop session on the first and last day of training. In the independent visual flow condition we replayed the visual flow generated by the same mouse in a preceding closed-loop session. In the no-visual flow condition, the reward location was still present in the virtual environment, but the walls of the environment were all uniformly gray.

For A24b axon stimulation experiments (**Figure 3.5**), viral injections and surgeries were performed as described above for axonal imaging with the exception that an AAV2/1-Ef1a-CatCh-eGFP (titer 2.3×10^{11} GC/ml) was injected into A24b of both hemispheres. Cranial windows were implanted over V1 of each hemisphere. Mice were trained and performed the same 2D navigation task as in **Figure 3.4**. A blue laser (473 nm) was directed either at monocular V1 or a blanking position using a pair of galvanometric mirrors. The stimulation (3 s, 25 Hz, 50% duty cycle) occurred once per traversal with a probability of 0.7 and at a random position in the tunnel (located in a region between 20% and 70% of the length of the tunnel). A CCD camera was used to calibrate and select stimulation locations using custom-written software. Laser power was adjusted to 16 mW/mm^2 at the surface of the brain.

Quantification and Statistical Analysis

Data analysis. All imaging data were corrected for lateral brain motion using custom-written software. Regions of interest were automatically selected using a combination of independent component analysis and image segmentation as previously described (Mukamel et al., 2009). Fluorescence traces were calculated as the mean pixel value in each region of interest per frame and

were median-normalized to calculate $\Delta F/F$. $\Delta F/F$ traces were filtered as previously described (Dombeck et al., 2007). The details of the statistical analysis are noted in the figure legends. We did not test the distribution of the data for normality.

Figure 3.2. The axon density profile was determined by calculating the mean fluorescence profile along the dorsal-ventral axis in V1 ($\lambda + 0.4$ mm). Individual traces were median normalized.

Figure 3.3. For stimulus-triggered fluorescence changes (onset responses, **Figure 3.3**, **Figure S3.2-3.3**), fluorescence traces were mean-subtracted in a window preceding the stimulus onset (running onset: -1900 ms to -1500 ms; mismatch, visual flow or random: -500 ms to -100 ms). To correct for differences in running speed between the different conditions, individual trials were subsampled to match the distribution of running speeds across the conditions. Subsampling was performed by removing trials with the highest and lowest running speed iteratively until the distributions had matching means. Onset response latencies were determined by finding the onset of a persistent, significant deviation from baseline ($p < 0.01$, paired Student's t test, for at least 1 s). Significance testing (Student's t test) was done in a window of 1 s centered on the peak of the distributions. To compare the distribution of running onset responses in A24/M2 axons and V1 neurons, we quantified the time to response as the time difference between the onset of the activity (more than 2 standard deviations above baseline for at least 1 s following this time) and the running onset. Each distribution of time differences was fitted with a two-term Gaussian model (**Figure S3.2A**). To quantify how visual flow influences running-related activity, we calculated a linear regression of the mean activity (in a window from 2 s to 4 s after running onset, averaged over all axons per imaging site) versus the average visual flow speed in a 1 s window that was varied between -10 s and 10 s. For each time step, we then calculated the average slope of the regression curves (averaged across imaging site) (**Figure S3.2E**). For the ivCRACM experiments, cells were considered to be driven by the stimulus light if the correlation between activity and stimulus was higher than 0.05. For the muscimol silencing experiments, the 50% of neurons most responsive to either running onset or mismatch were selected on two separate imaging time points preceding the baseline time point. Onset responses were then computed for these neurons for pre-muscimol (baseline, time point -1), muscimol inactivation of A24b (time point 0) and recovery (time point+1).

Figure 4. The activity $A(t)$ of each axon was correlated with different linear combinations of forward velocity $v(t)$ and rotation $r(t)$: $\alpha * v(t) + \beta * r(t)$, where $\alpha = \sin(\vartheta)$, $\beta = \cos(\vartheta)$. The interaction angle ϑ was varied from 0° to 180° in steps of 6° . For each axon an optimum interaction angle was determined, and axons were considered task-correlated if the correlation coefficient at optimum ϑ was higher than 0.005. Axons were classified as left (right) turning for an interaction angle between

6° and 66° (114° and 174°). Running forward ($\vartheta = 90^\circ$) and rotating without forward motion ($\vartheta = 0^\circ$, 180°) were excluded as they result in visual flow equal in magnitude in the left and right visual field. Learning was quantified as change in efficiency of moving towards the target. This was measured as the distance travelled towards the target normalized by the total distance travelled. Using this measure a random walk results in a performance of 0, movement in a straight line towards the target results in a score of 1.

Figure 3.5. To control for any unspecific effect of the stimulation on running speed, the stimulation effect of either left or right stimulation was normalized by subtracting the effect of bilateral stimulation.

EPILOGUE

We constantly make conscious predictions on multiple time scales, from predicting if it is safe to cross the street in front of an approaching car to predicting the questions in an upcoming exam. To achieve this, we employ internal models that capture regularities and relations of the external world and use them to make predictions about the future. Besides these more conscious forms of predictions, it is widely accepted that our brain also employs internal models for motor control (Körding and Wolpert, 2004; Wolpert et al., 1995) that are implemented in the cerebellum (e.g. (Ito, 2008; Wolpert et al., 1998) and motor cortex (e.g. (Stavisky et al., 2017)). To what extent internal models shape sensory processing is less clear. In the framework of predictive processing, internal models are distributed across all hierarchical levels of cortical processing and shape sensory processing already at the earliest cortical level in the respective modality (i.e. the primary sensory cortical areas) (Clark, 2013; Friston, 2005). My thesis work supports this view by showing that predictive processing can be used as a framework to explain visuomotor integration in the primary visual cortex (V1) of the mouse.

Visuomotor experience shapes predictive processing in visual cortex

In Chapter 1, I show that the coupling of sensory and motor experience during development shapes information processing in mouse V1. In particular, the neuronal response to a mismatch between locomotion and the resulting visual feedback depends on the mouse's experience of visual feedback coupled to locomotion during development (**Figure 2.1**). I show that neurons strongly responding to mismatch between locomotion and visual feedback (mismatch neurons) balance excitatory motor-related input against visual inhibition (**Figure 2.2**). By recording from genetically identified neuronal subpopulations, I uncovered that the inhibitory visual input onto mismatch neurons is mediated by somatostatin-positive interneurons (**Figure 2.4**). I then confirmed the central role of somatostatin-positive interneurons in this microcircuit by pharmacogenetic and optogenetic circuit manipulations (**Figure 2.5**, **Figure 2.6**). Finally, I demonstrate that normal visuomotor experience rapidly restores normal visual processing (**Figure 2.7**). These findings are consistent with the idea of predictive processing, where feedforward visual input is compared to a prediction that is based on an internal model, already in V1. Importantly, the internal model that generates the predictions of visual flow is based on a learned relationship between locomotion and visual flow (**Figure 4.1**). These results show that learned relationships between sensory input and motor output could shape our perception already at the earliest level.

Interestingly, our data also contains hints about predictions of visual input that are based on the structure of the visual input. We found that in mice that did not experience the coupling of visual feedback to locomotion, mismatch and playback-halt responses are very similar (**Figure 2.1, Figure S2.1**). This puzzling finding gives way for some intriguing speculations. During training, NT mice received visual flow input that was generated by other mice, and therefore not random. This means that the visual flow was inherently structured and predictable. The sudden halts that are characteristic for both mismatch and playback halt violate the statistics of the previously experienced visual flow. Mismatch and playback halt responses could therefore signal an error between the visual flow input and a prediction of visual flow that is based on the experience of visual flow input changing relatively smoothly. Such a prediction of visual input based on the statistics of the visual scene might originate in secondary visual areas. In coupled trained mice, we found only very few neurons that respond to playback halt (**Figure S2.1**). In these mice running speed was a perfect predictor of visual flow, predictions based on the statistics of visual flow might therefore be less pronounced.

The circuit model for the computation of mismatch proposes a motor-related input to V1 (Chapter 1, **Figure 4.1**). Based on the predictive coding framework, we assume that this motor-related input is a prediction of visual flow based on running speed that is relayed to V1 via feedback connections. Another possibility is that V1 receives excitatory input that signals running speed (and not a prediction of visual flow), and mismatch neurons compare running speed to visual speed. The experiments in Chapter 1 were not designed to disambiguate the two. Whether a motor-related input to V1 carries a prediction of visual flow and where it originates in the cortex was addressed in Chapter 2.

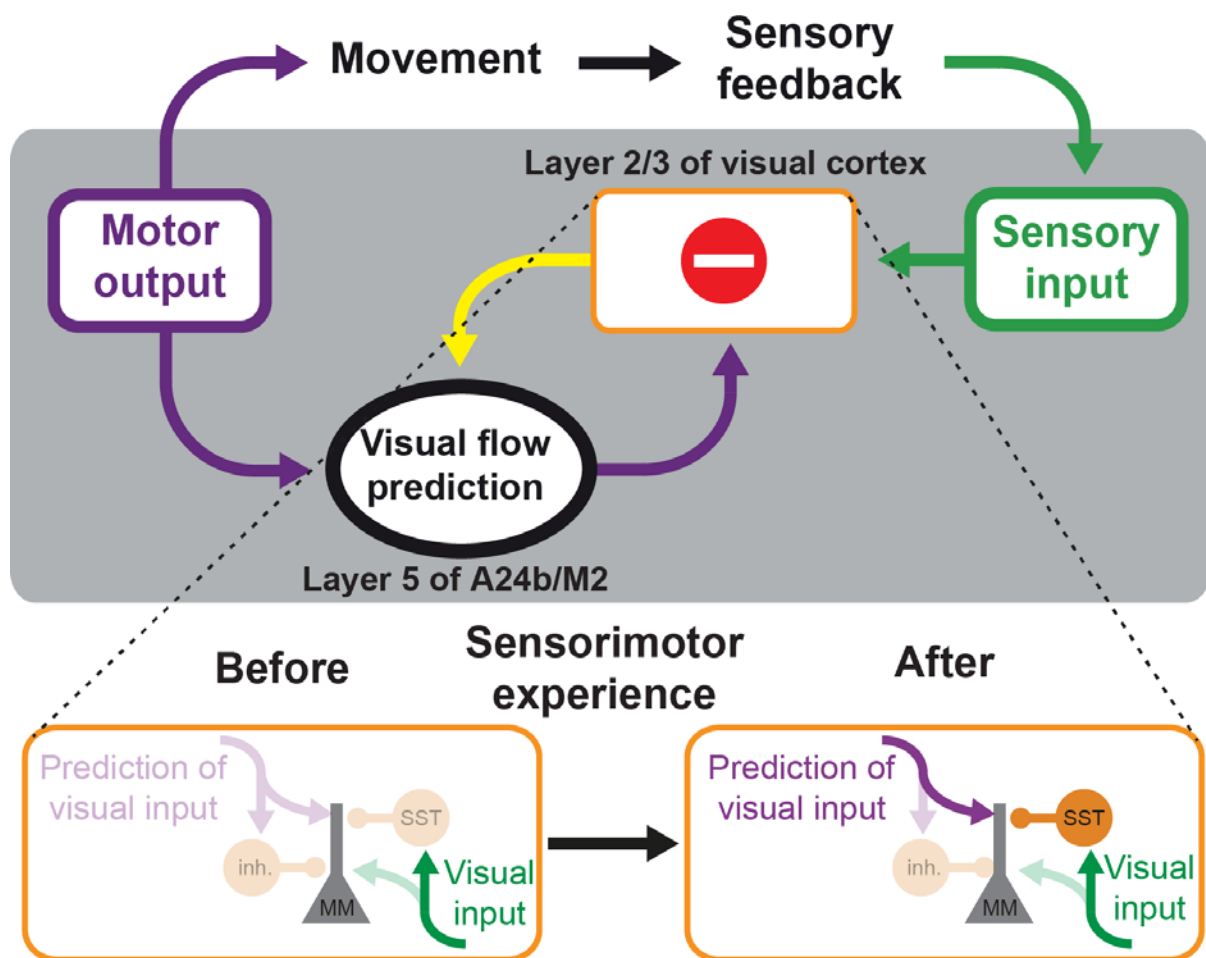


Figure 4.1 A sensorimotor circuit for visuomotor integration. Excitatory neurons in visual cortex integrate an excitatory motor related prediction of visual flow with actual visual flow input that is relayed by SST interneurons to compute mismatch between the two (Chapter 1). The motor related prediction of visual flow originates from the higher motor areas A24b and M2 (Chapter 2), possibly from layer 5 neurons. Mismatch responses depend on the experience of visual flow coupled to locomotion during development (Chapter 1), indicating that the internal model forming the visual-flow prediction is shaped by experience during development, but also in adult mice, as learning a task with novel visuomotor coupling changes the input from A24b/M2 to V1 (Chapter 2).

Motor-based prediction of visual flow from A24b to V1

By exhaustively mapping monosynaptic input to V1 with rabies virus, we find that mouse A24b, a higher motor area, sends a dense topographically organized input to V1 (**Figure 3.1**). By imaging the activity of axons from A24b in V1 and manipulating A24b activity, we find that they carry motor related signals that drive motor and mismatch signals in V1 (**Figure 3.3**). Importantly, this motor related input from A24b to V1 is shaped by experience, as training to navigate a left-right inverted world reverses A24b visuomotor coding (**Figure 3.4**). This is to me the strongest support for the idea that these connections send an actual prediction of visual flow. Thus, V1 receives a dense feedback input from a higher motor area that conveys a prediction of visual flow to V1. Such a predictive feedback input is an essential element in the description of cortical function in the theoretical framework of predictive coding.

The insights presented in both chapters complement each other and, together with additional recent work in the lab (Fiser et al., 2016; Zmarz and Keller, 2016), helped us to gain a mechanistic and comprehensive understanding of visuomotor integration in the visual cortex of the mouse. Together, our results suggest that motor-based predictions of visual flow, as well as predictions based on spatial location, are generated in A24b and sent to V1. There, excitatory neurons in L2/3 combine these predictions of visual input with the actual visual input that is relayed by somatostatin-positive interneurons (**Figure 4.1**). Importantly, these mismatch neurons have receptive fields resembling classical receptive fields (Zmarz and Keller 2016). The predictive processing framework therefore directly extends the more classic way of describing sensory processing in terms of receptive fields.

Outlook

While the puzzle pieces of a predictive processing framework of visuomotor integration are coming together, open questions about the computations and their implementation remain. In ongoing projects, I try to address a few of the questions that follow from the work presented in Chapter 1 and 2.

The proposed circuit model computes a negative prediction error, i.e. it is active when there is less sensory input than predicted. The computation of positive prediction error, that would signal more sensory input than expected, could be implemented with a circuit where a neuron combines excitatory sensory input with a prediction that is relayed by a local interneuron. To investigate this hypothesis, I am exposing mice to a virtual reality environment where we introduce different types of perturbations. This will allow me to investigate responses to different types of visuomotor perturbations, for example when there is suddenly more visual flow than expected.

Moreover, in addition to perturbations affecting the whole visual field, I am also using local perturbations where only small parts of the visual field are perturbed. In addition to visuomotor mismatch (i.e. violation of prediction based on locomotion), these perturbations also induce visual mismatch (i.e. violation of prediction based on adjacent regions in the visual world). This will allow me to investigate if and how different internal models might interact.

All of the data presented in this thesis was recorded from superficial layers of the visual cortex. The role of deep cortical layers in visuomotor integration of V1 is still unclear. By recording the activity of layer 5 neurons, I hope to get a better understanding of their role in visual processing and visuomotor integration. Circuit models for predictive processing localize state units based on projection patterns partially in superficial, but mainly in lower layers of cortex (Bastos et al., 2012; Shipp, 2016). Consistent with this hypothesis, preliminary data from experiments in layer 5 of V1

suggests that visuomotor mismatch events evoke only weak responses in neurons of layer 5, in contrast to the strong mismatch responses of layer 2/3 neurons. To explore the role of layer 5 neurons in visuomotor integration, I plan to characterize their activity more thoroughly and to perturb their activity by optogenetic activation and inhibition to see how activity manipulation of layer 5 neurons affects the activity of SST interneurons and excitatory mismatch neurons in layer 2/3.

Final remarks

The predictive processing framework offers a unifying framework to understand and explore brain function in general. It is an intriguing idea that the comparison between sensory input and a prediction based on internal models constitutes a generic computation of cortical processing. In the past decade, experimental evidence supporting this idea has accumulated. In the predictive processing framework, core symptoms of psychiatric diseases like schizophrenia can be understood as a disbalance between the weights of the sensory input and the predictions based on internal models. The effects of psychotic and antipsychotic drugs could be understood by affecting this balance, and delusions and hallucinations could be explained by dysfunctional, overactive predictions (Clark, 2013). Thus, a mechanistic and comprehensive understanding of predictive processing in the cortex could potentially open up novel ways for the understanding and treatment of psychiatric disorders.

References

- Adesnik, H., Bruns, W., Taniguchi, H., Huang, Z.J., and Scanziani, M. (2012). A neural circuit for spatial summation in visual cortex. *Nature* *490*, 226–231.
- Akerboom, J., Chen, T.-W., Wardill, T.J., Tian, L., Marvin, J.S., Mutlu, S., Calderon, N.C., Esposti, F., Borghuis, B.G., Sun, X.R., et al. (2012). Optimization of a GCaMP Calcium Indicator for Neural Activity Imaging. *J. Neurosci.* *32*, 13819–13840.
- Armbruster, B.N., Li, X., Pausch, M.H., Herlitze, S., and Roth, B.L. (2007). Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proc. Natl. Acad. Sci. U. S. A.* *104*, 5163–5168.
- Attinger, A., Wang, B., and Keller, G.B. (2017). Visuomotor Coupling Shapes the Functional Development of Mouse Visual Cortex. *Cell* *169*, 1291–1302.e14.
- Ayaz, A., Saleem, A.B., Schölvinck, M.L., and Carandini, M. (2013). Locomotion Controls Spatial Integration in Mouse Visual Cortex.
- Barlow, H.B. (1953). Summation and inhibition in the frog's retina. *J. Physiol.* *119*, 69–88.
- Barlow, H.B. (1994). What is the computational goal of the neocortex? In *Large-Scale Neuronal Theories of the Brain*, pp. 1–22.
- Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., and Friston, K.J. (2012). Canonical Microcircuits for Predictive Coding. *Neuron* *76*, 695–711.
- Blakemore, C., and Cooper, G.F. (1970). Development of the brain depends on the visual environment. *Nature* *228*, 477–478.
- Blakemore, S.J., Wolpert, D., and Frith, C. (2000). Why can't you tickle yourself? *Neuroreport* *11*, R11-6.
- Chen, T.-W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., et al. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* *499*, 295–300.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* *36*, 181–204.
- Craik, K.J. (1943). *The Nature of Explanation* (Cambridge University Press London).
- Cui, G., Jun, S.B., Jin, X., Pham, M.D., Vogel, S.S., Lovinger, D.M., and Costa, R.M. (2013). Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature* *494*, 238–242.
- Dayan, P., Hinton, G.E., Neal, R.M., and Zemel, R.S. (1995). The Helmholtz machine. *Neural Comput.* *7*, 889–904.
- Dhande, O.S., Stafford, B.K., Lim, J.-H.A., and Huberman, A.D. (2015). Contributions of Retinal Ganglion Cells to Subcortical Visual Processing and Behaviors. *Annu. Rev. Vis. Sci.* *1*, 291–328.
- Dombeck, D.A., Khabbaz, A.N., Collman, F., Adelman, T.L., and Tank, D.W. (2007). Imaging Large-Scale Neural Activity with Cellular Resolution in Awake, Mobile Mice. *Neuron* *56*, 43–57.
- Eliades, S.J., and Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* *453*, 1102–1106.
- Feldman, H., and Friston, K.J. (2010). Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* *4*, 215.
- Felleman, D.J., and Van Essen, D.C. (1991). Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cereb. Cortex* *1*, 1–47.
- Fino, E., and Yuste, R. (2011). Dense Inhibitory Connectivity in Neocortex. *Neuron* *69*, 1188–1203.

- Fiser, A., Mahringer, D., Oyibo, H.K., Petersen, A. V, Leinweber, M., and Keller, G.B. (2016). Experience-dependent spatial expectations in mouse visual cortex. *Nat. Neurosci.*
- Franklin, K.B.J., and Paxinos, G. (2012). Paxinos and Franklin's The mouse brain in stereotaxic coordinates.
- Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* *360*, 815–836.
- Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* *13*, 293–301.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* *11*.
- Friston, K., Kilner, J., and Harrison, L. (2006). A free energy principle for the brain. *J. Physiol.* *100*, 70–87.
- Frith, C.D., Blakemore, S., and Wolpert, D.M. (2000). Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. *Brain Res. Brain Res. Rev.* *31*, 357–363.
- Fu, Y., Tucciarone, J.M.M., Espinosa, J.S.S., Sheng, N., Darcy, D.P.P., Nicoll, R.A.A., Huang, Z.J.J., and Stryker, M.P.P. (2014). A cortical circuit for gain control by behavioral state. *Cell* *156*, 1139–1152.
- Garrido, M.I., Kilner, J.M., Stephan, K.E., and Friston, K.J. (2009). The mismatch negativity: a review of underlying mechanisms. *Clin. Neurophysiol.* *120*, 453–463.
- Gerfen, C.R., Paletzki, R., and Heintz, N. (2013). GENSAT BAC Cre-Recombinase Driver Lines to Study the Functional Organization of Cerebral Cortical and Basal Ganglia Circuits. *Neuron* *80*, 1368–1383.
- Gilbert, C.D., and Li, W. (2013). Top-down influences on visual processing. *Nat. Rev. Neurosci.* *14*, 350–363.
- Gilbert, C.D., and Sigman, M. (2007). Brain states: top-down influences in sensory processing. *Neuron* *54*, 677–696.
- Gong, S., Doughty, M., Harbaugh, C.R., Cummins, A., Hatten, M.E., Heintz, N., and Gerfen, C.R. (2007). Targeting Cre recombinase to specific neuron populations with bacterial artificial chromosome constructs. *J. Neurosci.* *27*, 9817–9823.
- Gregory, R.L. (1980). Perceptions as hypotheses. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* *290*, 181–197.
- Grosov, D.H., Shapley, R.M., and Hawken, M.J. (1993). Macaque V1 neurons can signal “illusory” contours. *Nature* *365*, 550–552.
- Han, X., Chow, B.Y., Zhou, H., Klapoetke, N.C., Chuong, A., Rajimehr, R., Yang, A., Baratta, M. V, Winkle, J., Desimone, R., et al. (2011). A high-light sensitivity optical neural silencer: development and application to optogenetic control of non-human primate cortex. *Front. Syst. Neurosci.* *5*, 18.
- Hawkins, J., and Blakeslee, S. (2004). *On intelligence* (Times Books).
- Hein, A., and Held, R. (1967). Dissociation of the visual placing response into elicited and guided components. *Science* *158*, 390–392.
- Held, R., and Hein, A. (1963). Movement-produced stimulation in the development of visually guided behavior. *J. Comp. Physiol. Psychol.* *56*, 872–876.
- Hippenmeyer, S., Vrieseling, E., Sigrist, M., Portmann, T., Laengle, C., Ladle, D.R., and Arber, S. (2005). A developmental switch in the response of DRG neurons to ETS transcription factor signaling. *PLoS Biol.* *3*, e159.
- Hirsch, H. V, and Spinelli, D.N. (1970). Visual experience modifies distribution of horizontally and vertically oriented receptive fields in cats. *Science* *168*, 869–871.
- Hofer, S.B., Ko, H., Pichler, B., Vogelstein, J., Ros, H., Zeng, H., Lein, E., Lesica, N.A., and Mrsic-Flogel,

- T.D. (2011). Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nat. Neurosci.* *14*, 1045–1052.
- von Holst, E., and Mittelstaedt, H. (1950). Das Reafferenzprinzip. *Naturwissenschaften* *37*, 464–476.
- Hubel, D., and Wiesel, T. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* *160*, 106.
- Hubel, D.H., and Wiesel, T.N. (1970). The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *J. Physiol.* *206*, 419–436.
- Ibrahim, L.A.A., Mesik, L., Ji, X.-Y., Fang, Q., Li, H., Li, Y.-T., Zingg, B., Zhang, L.I.I., and Tao, H.W.W. (2016). Cross-Modality Sharpening of Visual Cortical Processing through Layer-1-Mediated Inhibition and Disinhibition. *Neuron* *89*, 1031–1045.
- Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nat. Rev. Neurosci.* *9*, 304–313.
- Jiang, X., Shen, S., Cadwell, C.R., Berens, P., Sinz, F., Ecker, A.S., Patel, S., and Tolias, A.S. (2015). Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* *350*, aac9462.
- Jordan, M.I., and Rumelhart, D.E. (1992). Forward Models: Supervised Learning with a Distal Teacher. *Cogn. Sci.* *16*, 307–354.
- Kanai, R., Komura, Y., Shipp, S., and Friston, K. (2015). Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philos. Trans. R. Soc. London B Biol. Sci.* *370*.
- Keck, T., Keller, G.B., Jacobsen, R.I., Eysel, U.T., Bonhoeffer, T., and Hübener, M. (2013). Synaptic scaling and homeostatic plasticity in the mouse visual cortex in vivo. *Neuron* *80*, 327–334.
- Keller, G.B., and Hahnloser, R.H.R. (2008). Neural processing of auditory feedback during vocal practice in a songbird. *Nature* *457*, 187–190.
- Keller, A.J., Houlton, R., Kampa, B.M., Lesica, N.A., Mrsic-Flogel, T.D., Keller, G.B., and Helmchen, F. (2017). Stimulus relevance modulates contrast adaptation in visual cortex. *Elife* *6*.
- Keller, G.B., Bonhoeffer, T., and Hübener, M. (2012). Sensorimotor mismatch signals in primary visual cortex of the behaving mouse. *Neuron* *74*, 809–815.
- Kerlin, A.M., Andermann, M.L., Berezovskii, V.K., and Reid, R.C. (2010). Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. *Neuron* *67*, 858–871.
- Klapoetke, N.C., Murata, Y., Kim, S.S., Pulver, S.R., Birdsey-Benson, A., Cho, Y.K., Morimoto, T.K., Chuong, A.S., Carpenter, E.J., Tian, Z., et al. (2014). Independent optical excitation of distinct neural populations. *Nat. Methods* *11*, 338–346.
- Knill, D.C., and Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends Neurosci.* *27*, 712–719.
- Körding, K.P., and Wolpert, D.M. (2004). Bayesian integration in sensorimotor learning. *Nature* *427*, 244–247.
- Lee, S.-H., and Dan, Y. (2012). Neuromodulation of brain states. *Neuron* *76*, 209–222.
- Leinweber, M., Zmarz, P., Buchmann, P., Argast, P., Hübener, M., Bonhoeffer, T., and Keller, G.B. (2014). Two-photon calcium imaging in mice navigating a virtual reality environment. *J. Vis. Exp.* e50885.
- Leonardo, A., and Konishi, M. (1999). Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature* *399*, 466–469.
- Li, N., Chen, T.-W., Guo, Z. V., Gerfen, C.R., and Svoboda, K. (2015). A motor cortex circuit for motor planning and movement. *Nature* *519*, 51–56.

- Linden, J.F., Liu, R.C., Sahani, M., Schreiner, C.E., and Merzenich, M.M. (2003). Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex. *J. Neurophysiol.* *90*, 2660–2675.
- Livingstone, M.S., Freeman, D.C., and Hubel, D.H. (1996). Visual responses in V1 of freely viewing monkeys. *Cold Spring Harb. Symp. Quant. Biol.* *61*, 27–37.
- London, M., Larkum, M.E., and Häusser, M. (2008). Predicting the synaptic information efficacy in cortical layer 5 pyramidal neurons using a minimal integrate-and-fire model. *Biol. Cybern.* *99*, 393–401.
- Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., et al. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* *13*, 133–140.
- Makino, H., and Komiyama, T. (2015). Learning enhances the relative impact of top-down processing in the visual cortex. *Nat. Neurosci.* *18*, 1116–1122.
- Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., and Wu, C. (2004). Interneurons of the neocortical inhibitory system. *Nat. Rev. Neurosci.* *5*, 793–807.
- Marr, D. (1982). *Vision* (MIT Press).
- Miller, M.W., and Vogt, B.A. (1984). Direct connections of rat visual cortex with sensory, motor, and association cortices. *J. Comp. Neurol.* *226*, 184–202.
- Mukamel, E.A., Nimmerjahn, A., and Schnitzer, M.J. (2009). Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron* *63*, 747–760.
- Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol. Cybern.* *66*, 241–251.
- Niell, C.M., and Stryker, M.P. (2008). Highly selective receptive fields in mouse visual cortex. *J. Neurosci.* *28*, 7520–7536.
- Niell, C.M., and Stryker, M.P. (2010). Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* *65*, 472–479.
- Nordeen, K.W., and Nordeen, E.J. (1992). Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behav. Neural Biol.* *57*, 58–66.
- Otchy, T.M., Wolff, S.B.E., Rhee, J.Y., Pehlevan, C., Kawai, R., Kempf, A., Gobes, S.M.H., and Ölveczky, B.P. (2015). Acute off-target effects of neural circuit manipulations. *Nature* 1–16.
- Pakan, J.M., Lowe, S.C., Dylida, E., Keemink, S.W., Currie, S.P., Coutts, C.A., and Rochefort, N.L. (2016). Behavioral-state modulation of inhibition is context-dependent and cell type specific in mouse visual cortex. *Elife* *5*.
- Petreaunu, L., Huber, D., Sobczyk, A., and Svoboda, K. (2007). Channelrhodopsin-2-assisted circuit mapping of long-range callosal projections. *Nat. Neurosci.* *10*, 663–668.
- Pfeffer, C.K., Xue, M., He, M., Huang, Z.J., and Scanziani, M. (2013). Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nat. Neurosci.* *16*, 1068–1076.
- Pi, H.-J., Hangya, B., Kvitsiani, D., Sanders, J.I., Huang, Z.J., and Kepecs, A. (2013). Cortical interneurons that specialize in disinhibitory control. *Nature* *503*, 521–524.
- Pinto, L., Goard, M.J., Estandian, D., Xu, M., Kwan, A.C., Lee, S.-H., Harrison, T.C., Feng, G., and Dan, Y. (2013). Fast modulation of visual perception by basal forebrain cholinergic neurons. *Nat. Neurosci.* *16*, 1857–1863.
- Polack, P.-O., Friedman, J., and Golshani, P. (2013). Cellular mechanisms of brain state-dependent

gain modulation in visual cortex. *Nat. Neurosci.* *16*, 1331–1339.

Rancz, E.A., Moya, J., Drawitsch, F., Brichta, A.M., Canals, S., and Margrie, T.W. (2015). Widespread vestibular activation of the rodent cortex. *J. Neurosci.* *35*, 5926–5934.

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. *Nat. Neurosci.* *2*, 79–87.

Reimer, J., McGinley, M.J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D.A., Tolias, A.S., Iriki, A., Tanaka, M., Iwamura, Y., et al. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nat. Commun.* *7*, 13289.

Riesenhuber, M., and Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat. Neurosci.* *2*, 1019–1025.

Roelfsema, P.R., Lamme, V.A., and Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature* *395*, 376–381.

Saleem, A.B., Ayaz, A., Jeffery, K.J., Harris, K.D., and Carandini, M. (2013). Integration of visual motion and locomotion in mouse visual cortex. *Nat. Neurosci.* *16*, 1864–1869.

Salinas, E., and Sejnowski, T.J. (2001). Correlated neuronal activity and the flow of neural information. *Nat. Rev. Neurosci.* *2*, 539–550.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* *275*, 1593–1599.

Shipp, S. (2016). Neural Elements for Predictive Coding. *Front. Psychol.* *7*, 1792.

Sinha, P., Kjelgaard, M.M., Gandhi, T.K., Tsourides, K., Cardinaux, A.L., Pantazis, D., Diamond, S.P., and Held, R.M. (2014). Autism as a disorder of prediction. *Proc. Natl. Acad. Sci. U. S. A.* *111*, 15220–15225.

Sinnamon, H.M., and Galer, B.S. (1984). Head movements elicited by electrical stimulation of the anteromedial cortex of the rat. *Physiol. Behav.* *33*, 185–190.

Spratling, M.W. (2010). Predictive coding as a model of response properties in cortical area V1. *J. Neurosci.* *30*, 3531–3543.

Spratling, M.W. (2017). A review of predictive coding algorithms. *Brain Cogn.* *112*, 92–97.

Stavisky, S.D., Kao, J.C., Ryu, S.I., and Shenoy, K. V (2017). Trial-by-Trial Motor Cortical Correlates of a Rapidly Adapting Visuomotor Internal Model. *J. Neurosci.* *37*, 1721–1732.

Taniguchi, H., He, M., Wu, P., Kim, S., Paik, R., Sugino, K., Kvitsiani, D., Kvitsani, D., Fu, Y., Lu, J., et al. (2011). A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. *Neuron* *71*, 995–1013.

Vinck, M., Batista-Brito, R., Knoblich, U., and Cardin, J.A. (2015). Arousal and Locomotion Make Distinct Contributions to Cortical Activity Patterns and Visual Encoding. *Neuron* *86*, 740–754.

Vogt, B.A., and Miller, M.W. (1983). Cortical connections between rat cingulate cortex and visual, motor, and postsubicular cortices. *J. Comp. Neurol.* *216*, 192–210.

Vogt, B.A., and Paxinos, G. (2014). Cytoarchitecture of mouse and rat cingulate cortex with human homologies. *Brain Struct. Funct.* *219*, 185–192.

Vogt, B.A., Rosene, D.L., and Pandya, D.N. (1979). Thalamic and cortical afferents differentiate anterior from posterior cingulate cortex in the monkey. *Science* *204*, 205–207.

Vogt, B.A., Pandya, D.N., and Rosene, D.L. (1987). Cingulate cortex of the rhesus monkey: I. Cytoarchitecture and thalamic afferents. *J. Comp. Neurol.* *262*, 256–270.

Vong, L., Ye, C., Yang, Z., Choi, B., Chua, S., and Lowell, B.B. (2011). Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons. *Neuron* *71*, 142–154.

- Wang, Q., and Burkhalter, A. (2007). Area map of mouse visual cortex. *J. Comp. Neurol.* *502*, 339–357.
- Wang, Q., Sporns, O., and Burkhalter, A. (2012). Network analysis of corticocortical connections reveals ventral and dorsal processing streams in mouse visual cortex. *J. Neurosci.* *32*, 4386–4399.
- Wertz, A., Trenholm, S., Yonehara, K., Hillier, D., Raics, Z., Leinweber, M., Szalay, G., Ghanem, A., Keller, G., Rózsa, B., et al. (2015). Single-cell-initiated monosynaptic tracing reveals layer-specific cortical network modules. *Science* *349*, 70–74.
- Wickersham, I.R., Lyon, D.C., Barnard, R.J.O., Mori, T., Finke, S., Conzelmann, K.-K., Young, J.A.T., and Callaway, E.M. (2007). Monosynaptic restriction of transsynaptic tracing from single, genetically targeted neurons. *Neuron* *53*, 639–647.
- Wickersham, I.R., Sullivan, H.A., and Seung, H.S. (2010). Production of glycoprotein-deleted rabies viruses for monosynaptic tracing and high-level gene expression in neurons. *Nat. Protoc.* *5*, 595–606.
- Wilson, N.R., Runyan, C.A., Wang, F.L., and Sur, M. (2012). Division and subtraction by distinct cortical inhibitory networks in vivo. *Nature* *488*, 343–348.
- Wolpert, D., Ghahramani, Z., and Jordan, M. (1995). An internal model for sensorimotor integration. *Science* (80-). *269*, 1880.
- Wolpert, D.M., Miall, R.C., and Kawato, M. (1998). Internal models in the cerebellum. *Trends Cogn. Sci.* *2*, 338–347.
- Xue, M., Atallah, B. V., and Scanziani, M. (2014). Equalizing excitation–inhibition ratios across visual cortical neurons. *Nature*.
- Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J.P., Chang, W.-C., Jenvay, S., Miyamichi, K., Luo, L., and Dan, Y. (2014). Long-range and local circuits for top-down modulation of visual cortex processing. *Science* (80-). *345*, 660–665.
- Zhang, S., Xu, M., Chang, W.-C., Ma, C., Hoang Do, J.P., Jeong, D., Lei, T., Fan, J.L., and Dan, Y. (2016). Organization of long-range inputs and outputs of frontal cortex for top-down control. *Nat. Neurosci.*
- Zhou, Y., Shi, L., Cui, X., Wang, S., and Luo, X. (2016). Functional Connectivity of the Caudal Anterior Cingulate Cortex Is Decreased in Autism. *PLoS One* *11*, e0151879.
- Zipser, K., Lamme, V.A., and Schiller, P.H. (1996). Contextual modulation in primary visual cortex. *J. Neurosci.* *16*, 7376–7389.
- Zmarz, P., and Keller, G.B. (2016). Mismatch Receptive Fields in Mouse Visual Cortex. *Neuron* *92*, 766–772.

