Mismatch receptive fields in mouse primary visual cortex

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Phliosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Pawel Zmarz

von Polen

Basel, 2019

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Dr Silvia Arber

Dr Georg Keller

Dr Tara Keck

Basel, 21.03.2017

_____

Prof. Dr. M. Spiess, Dekan

Table of Contents

Acknowledgements

I would like to thank Georg for giving me the opportunity to work with him. I have learnt, and continue to learn much from him, and I could not have asked for a better preparation for the next step of academia. I would like to thank all of my colleagues in the lab for continuous support and an excellent atmosphere that was highly conducive to solid research. Particular thanks to Aris Fiser and Bo Wang for the appropriate amount of absurdity necessary to maintain a grip on reality.

I would like to thank my thesis advisory committee, Dr Silvia Arbor and Dr Tara Keck for their support and feedback throughout the project. I would particularly like to thank Dr Tara Keck for her extraordinary involvement in the project, unwavering push for scientific clarity, and fabulous food recommendations. I would like to thank my colleagues from FMI, Dr Chie Sato, Elisabeth Meyer, Dr Kuo-Hua Huang, Dr Tamas Szkira, Dr Tasos Moresis, and Dr Tafheem Masoudi for making my FMI experience the best it could have been.

I would lastly like to thank my family, Babcia Basia, Jolanta Zmarz, Robert Gralek, Jacek Zmarz, and Tomek Miturski, for their continuous support and excitement for anything I choose to do.

Abstract

In primary visual cortex, a subset of neurons responds when a particular stimulus is encountered in a certain location in visual space. This activity can be modeled using a visual receptive field. In addition to visually driven activity, there are neurons in visual cortex that integrate visual and motor-related input to signal a mismatch between actual and predicted visual flow. Here we show that these mismatch neurons have receptive-fields and signal a local mismatch between actual and predicted visual flow in restricted regions of visual space. These mismatch receptive-fields are aligned to the retinotopic map of visual cortex and are similar in size to visual receptive-fields. Thus, neurons with mismatch receptive fields signal local deviations of actual visual flow from visual flow predicted based on self-motion and could therefore underlie the detection of objects moving relative to the visual flow caused by self-motion.

Introduction

A major goal of neuroscience has been to understand how our brains generate perception – our interpretation of our environment informed by our senses. One very fruitful strand of research aimed at answering this question centered on the idea that the brain builds a representation of the external world from individual neurons extracting specific features of their sensory input. This notion that individual neurons could function as 'feature detectors' (Barlow, 1953) was an extension from the concept of a neuronal 'receptive field' (Hartline, 1937; Sherrington, 1906) – the idea that there is a specific set of stimulus features that optimally excites a given neuron. In their seminal work, David Hubel and Torsten Wiesel used this idea to account for the response properties of simple and complex cells in primary visual cortex (V1) of the cat (Hubel and Wiesel, 1962; **Figure 1A**). An important feature of this hierarchical framework was that neuronal responses became increasingly invariant to stimulus features so that the cortex could represent ever more complex stimuli through a feedforward hierarchy of receptive fields (Riesenhuber and Poggio, 1999; **Figure 1B**). Consistent with these ideas, cortical areas that harbor more complex (and more invariant) representations such as faces (Tsao et al., 2006), concepts of individuals (Quiroga et al., 2005) or space (Hafting et al., 2005; O'Keefe and Dostrovsky, 1971) have since been identified.

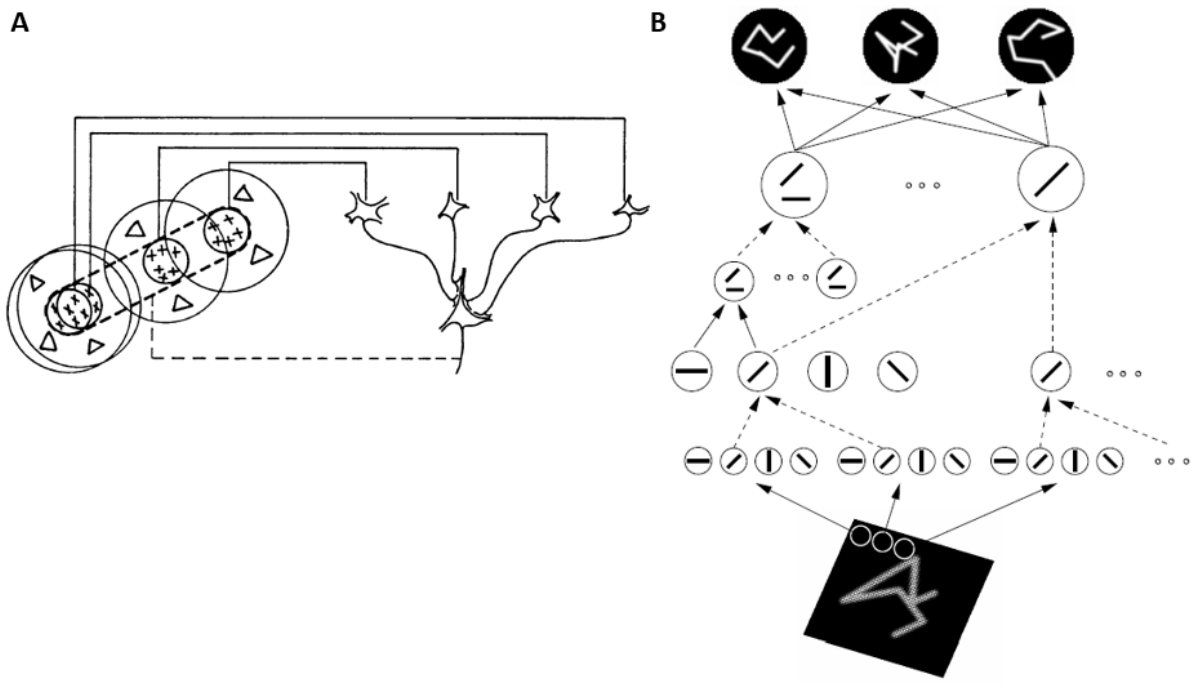

Figure 1. Classical, feedforward hierarchy of representations in the visual system based on receptive fields. (A) Hubel and Wiesel's model for simple-cell receptive fields. Thalamic cells (three example cells on the left) that project to V1 have circular 'on-off' receptive fields. A simple cell (on the right) responds to a black bar stimulus if it appears in a specific location in the visual field. The simple cell could respond to a bar stimulus by linear summation of the 'on' regions of the thalamic neurons' receptive fields (dotted lines connecting the on regions of the three cells on the left) that sample adjacent regions of visual space. Adapted from (Hubel and Wiesel, 1962). (B) Extension of the Hubel and Wiesel model from (A) for more complex representations. Adapted from (Riesenhuber and Poggio, 1999).

Although the concept of feedforward hierarchical processing has been an important model for understanding cortical function, it has long been appreciated that sensory processing must be influenced by motor activity (Pollen, 1999). This notion is intuitive since a significant fraction of sensory input is self-generated and there is a noticeable difference in the way we perceive self-generated from externally-generated stimuli: hearing your own

voice versus hearing it back on tape, or touching your arm versus being tickled produces different percepts for the same sensory stimuli. These examples suggests that sensory and motor information is integrated to establish stable perception in the presence of self-generated stimuli, but how could this be implemented in the brain? One idea dating back to John Locke and later Herman von Helmholtz (Pollen, 1999) postulates that the motor system sends a copy of motor commands to sensory systems, which is integrated with feedforward input to generate perception (Sperry, 1950; von Holst, 1954). Evidence for such motor 'efference copies' initially came from recordings in electric fish whereby motor commands for activating the electric organ suppressed responses in electro-sensory neurons (Bell, 1981). Similar suppression effects have since been observed during vocalization in auditory areas of crickets (Poulet and Hedwig, 2006) and monkeys (Eliades and Wang, 2005, 2003), while microsaccades have been found to modulate activity throughout the visual system (Leopold and Logothetis, 1998; Martinez-Conde et al., 2000). Additionally, in rodents, locomotion has been found to differentially modulate neuronal responses in auditory (Schneider et al., 2014) and visual (Erisken et al., 2014; Niell and Stryker, 2010; Roth et al., 2016) areas. These data, coupled with the finding that only a small fraction of the input to a cortical neuron originates from thalamic (feed-forward) inputs (Ahmed et al., 1994; Binzegger, 2004; Garey and Powell, 1971), suggests that neuronal response properties need to be studied in the appropriate context of on-going cortical activity (such as in an awake animal, behaving in a naturalistic setting).

How can we reconcile the hierarchical organization of sensory systems with the diverse influence of motor-related activity on sensory processing? One possible way forward is to invert the idea of representations and suppose that the function of the cortex is to infer the causes of sensory stimuli (Friston, 2005; Rao, 1999). This approach to brain function, termed 'predictive coding' (predictive coding is itself an implementation of the general Bayesian brain hypothesis), postulates that higher order areas predict the activity of lower areas, while in every area a comparison is made between the higher order prediction and lower order input (Friston, 2005; Rao, 1999). Any deviation between the predicted and actual input is then fed back to higher order areas to update future

predictions (a simplified scheme is presented in **Figure 2**). In this view, the classical 'sensory' response is only present at the periphery, with subsequent areas only transmitting aspects of the input that are not predicted by higher areas (prediction error) and therefore cortical responses can be understood as transient error signals that attenuate as the predictive inputs are tuned to the current sensory input. This idea of active inference has been formalized by Karl Friston as the free-energy principle whereby the function of the brain is to minimize uncertainty by optimizing an internal model using sensory inputs (Friston, 2010).

Why would such a scheme be useful? A simple example comes from internal-model based theories of sensorimotor learning (Jordan and Rumelhart, 1992; Wolpert et al., 1995). As with the scheme above, the predicted sensory consequence of an action (such as hitting a ball with a racket) and the sensory feedback resulting from that action (such as feeling the force on your hand) can be compared and any resulting differences can be used to instruct the subsequent action. Learning is then optimized by iteratively minimizing the error between the prediction and the outcome. Another intriguing aspect of predictive coding schemes is that they can be implemented across various sensory systems, meaning that the different sensory areas could implement the same underlying circuit motif to perform predictive coding computations (such as in **Figure 2**). If true, such an insight would provide helpful constraints on experiments investigating cortical function. The concept of a repeated 'canonical circuit' in the neocortex that integrates thalamic and cortical input has in fact been previously proposed (Douglas et al., 1989)(**Figure 3**).

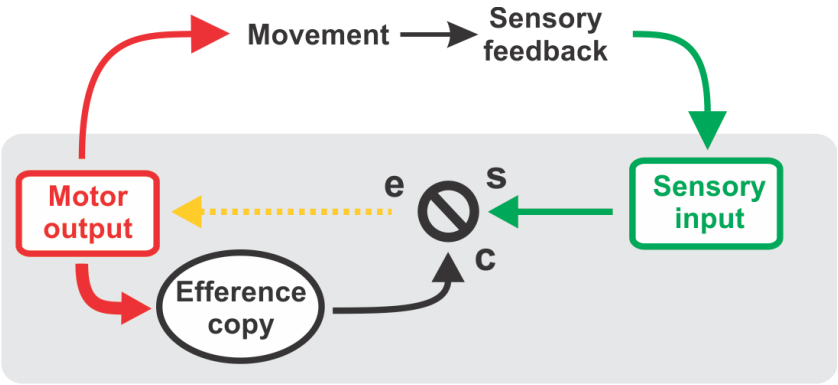

Figure 2. Model circuit for predictive coding computation. Motor commands generate movement that leads to sensory feedback, which is detected by primary sensory areas (for example, visual flow resulting from locomotion). At the same time an internal model of movement generates a prediction of the sensory consequences of the movement (sent via an 'efference copy') and is relayed from the motor area to the sensory area. A comparison of the predicted and the actual sensory input is computed and any residual error is detected and fed back to the motor system for updating the internal model and instructing subsequent action. Adapted from (Keller and Hahnloser, 2009).

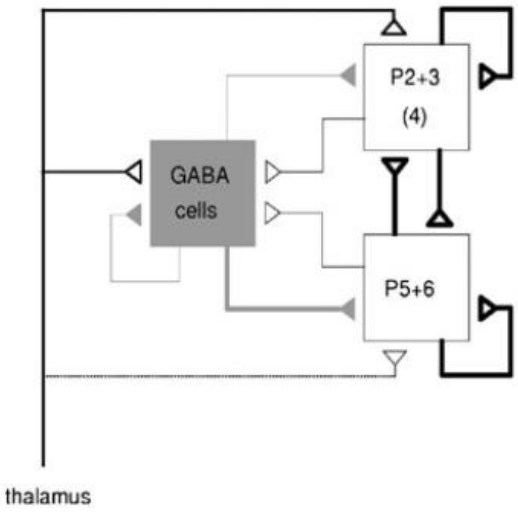

Figure 3. Canonical circuit model for the cortex. The model is composed of three population of neurons: one inhibitory (GABA cells) and two excitatory pools. Subpopulations of neurons in each pool receive excitatory thalamic input, which is integrated in the context of ongoing recurrent activity. The thickness of lines represents strength of input, and open/closed arrow heads correspond to excitatory/inhibitory synapses, receptively. Adapted from (Douglas and Martin, 2011). Note: along with these local connections, cortical neurons additionally receive dense long-range cortical and sub-cortical inputs (Miller and Vogt, 1984; Vogt and Miller, 1983).

A predictive coding framework for brain function has been suggested to account for a diverse range of phenomena from projection patterns and synaptic plasticity, to electrophysiological effects such as mismatch negativity (Friston, 2005). Neural network models of visual processing based on predictive coding principles have been shown to reproduce properties of classical (such as 'on/off' receptive fields and orientation tuning)(Olshausen and Field, 1996), as well as extra-classical (such as 'end-stopping') visual receptive field effects (Rao and Ballard, 1999), the latter of which are problematic to explain in a purely feedforward framework. By making use of the fact that local regions of visual scenes are correlated in space and time (and can therefore be predicted), predictive coding schemes have also been shown to reduce the redundancy of representations of the external environment, since only the unpredicted aspects of the sensory input are transmitted (Huang and Rao, 2011). Experimentally, neurons reporting deviations between expected and actual sensory stimuli have been found in the auditory pallium of songbirds (Keller and Hahnloser, 2009), auditory cortex of monkeys (Eliades and Wang, 2008) and visual cortex and higher order thalamic nuclei of mice (Keller et al., 2012; Roth et al., 2016). However, the question of whether core concepts of feedforward sensory processing, such as the receptive field, can in fact be integrated into the framework of predictive coding still lacks experimental confirmation. This study aims to address this in primary visual cortex of the mouse.

Anatomy of the early visual system

Visual input entering the eye is relayed out of the retina to the lateral geniculate nucleus through retinal ganglion cells (RGCs). Throughout the visual system, the topography of these retinal representation of the visual field is maintained, an organizational principle termed 'retinotopy'. Although there are multiple subcortical targets of RGCs, the lateral geniculate nucleus (LGN) of the thalamus is the major path for the visual information to reach cortex (Felleman and Van Essen, 1991). Historically LGN is thought to simply act as a relay station, however recent evidence has suggested that sensory processing also occurs at this stage (Cudeiro and Sillito, 2006; Erisken et al., 2014; Roth et al., 2016). Projections from LGN innervate primary visual cortex (V1) targeting cortical layers (L)4 and 6, with L4 receiving stronger input than L6. The visual information is then directed from L4 to L2 and L3 (in the mouse where the two layers are not readily histologically distinct L2 and L3 are referred to as L2/3), from which it is broadcast vertically to L5 (but avoiding L4) within V1, as well as to L4 of other 'downstream' cortical areas (V2, V4 etc.). V1 L5 neurons broadcast information to other cortical as well as sub-cortical targets. Although greatly simplified (Harris and Mrsic-Flogel, 2013; Harris and Shepherd, 2015), the information flow in the circuit described above constitutes the classical 'feedforward' model of information flow in the visual cortex. Other inputs to V1 neurons are classically termed 'feedback' and originate locally (horizontally and vertically) from within V1, as well as from other cortical and sub-cortical areas (Harris and Mrsic-Flogel, 2013; Harris and Shepherd, 2015). Within V1, there are dense horizontal and recurrent connections within L2, L3 and L5. L2 and L3 are also targeted by vertical projections from deeper layers within V1. Other cortical and sub-cortical areas largely target L2, 3 and 5 of V1. Of note is that LGN input to V1 is only a small fraction of the total input, and some intracortical projections (such as retrosplenial and anterior cingulate cortex) appear much denser than LGN projections (Miller and Vogt, 1984; Vogt and Miller, 1983).

Visual processing

Classical, hierarchical models of visual processing have largely focused on feedforward information flow in the visual system (Riesenhuber and Poggio, 1999). In cats and primates (possibly also in rodents (Marshel et al., 2011)), visual information flow is segregated into two parallel streams: dorsal stream concerned with motion perception and visually guided movements, and ventral stream concerned with object recognition (Goodale and Milner, 1992). LGN inputs are pooled and transformed in V1 to generate responses selective to simple features of the visual scene (e.g. orientation). These representations are then used to construct more complex representations along the visual hierarchy that become increasingly invariant (with respect to e.g. contrast), such that cortical areas located in later stages of the hierarchy process complex objects such as faces (Tsao et al., 2006) or complex representations of visual motion (Simoncelli et al., 1998). Feedback inputs are classically interpreted as 'modulatory' (Crick and Koch, 1998; Sherman and Guillery, 1998), sharpening feature selectivity or changing the gain of visual responses.

More recently, researchers have begun to dissect the role of feedback connections in the visual system in greater detail. Indeed, majority of input to V1 L2 and L3 does not originate from the thalamus, but rather from local (horizontal and vertical) V1 connections, and cortical and sub-cortical projections (Harris and Mrsic-Flogel, 2013; Harris and Shepherd, 2015). In order to constrain the possible wiring diagrams of cortical circuitry, a canonical circuit has been proposed based on V1 anatomy and functional connectivity (Douglas et al., 1989; Douglas and Martin, 2011), a motif that may be repeated throughout cortex (Heinzle et al., 2007). The classical notion of the visual hierarchy based on feedforward projections has also been opened for debate (Vezoli, 2004). In terms of function, recordings from behaving mice have shown that locomotion can modulate and drive activity throughout the early visual system of the mouse – consistent with the dense input that V1 receives from motor-related cortical areas such as anterior cingulate cortex and secondary motor cortex (Miller and Vogt, 1984; Vogt and Miller, 1983). In primary visual cortex, locomotion amplifies visual responses (Niell and Stryker, 2010) and drives activity even in the absence of visual input (Keller et al., 2012; Saleem et al., 2013) in a running speed-dependent manner

(Saleem et al., 2013). Similar locomotion-related effects have been reported in dorsolateral geniculate nucleus and lateral posterior nucleus of the visual thalamus (Erisken et al., 2014; Roth et al., 2016). These data suggest that classical notions of V1 function based on hierarchical representations of visual features have to be extended to include the prominent role of 'feedback' input.

Primary visual cortex as a model circuit for understanding cortical function

The visual system is a good candidate to study predictive coding mechanisms. From a conceptual perspective, if cortex can make predictions of sensory input, predicting visual input given locomotion is a particularly straightforward prediction – locomotion always results in visual flow. The two are inexorably coupled such that the speed of the visual flow always scales with locomotion speed, which makes such self-generated visual flow readily predictable. From an anatomical perspective, V1 neurons integrate thalamic and cortical input (Binzegger, 2004; Miller and Vogt, 1984; Vogt and Miller, 1983), which means they sample on-going cortical processing as well as the current sensory input. If the dense projections from motor-related areas to V1 neurons convey predictions of visual flow, then V1 neurons could compare those predictions to current sensory input to detect deviations from predictions. Furthermore, the visual system exhibits hierarchical organization (Riesenhuber and Poggio, 1999), which provides a possible anatomical substrate for predictive hierarchies (Friston, 2005). Finally, from a physiological perspective, motor-related signals have been shown to drive activity in V1 (Fu et al., 2014; Keller et al., 2012; Lee et al., 2014; Niell and Stryker, 2010; Pinto et al., 2013; Polack et al., 2013; Saleem et al., 2013). Given that visual flow predictions are coupled to locomotion, such predictive signals would appear motor-related, which are precisely the type of signals that have been observed in V1.

As demonstrated by a number of studies (Eliades and Wang, 2008; Keller et al., 2012; Keller and Hahnloser, 2009), the entry point to studying predictive mechanisms in a sensory system is to violate the animal's expectations by perturbing the sensory feedback during motor output. If the purpose of the system is to optimize the internal

model based on current sensory input (Friston, 2010), the error signals (sensorimotor mismatch) used to update the internal model should exhibit stimulus specificity – a hallmark of feedforward, hierarchical representation models of cortex (Riesenhuber and Poggio, 1999). Additionally, if such mismatch signals are useful for sensorimotor learning (Wolpert et al., 1995), these signals must contain the information necessary to refine the subsequent motor output for improvement to occur. It is currently not known whether mismatch signals recorded from primary sensory areas carry stimulus specific information that could in principle inform upstream predictive machinery, such as location of mismatch in sensory space. An alternative explanation for mismatch signals recorded thus far could be that the perturbations of sensory feedback used to study mismatch (Eliades and Wang, 2008; Keller et al., 2012; Keller and Hahnloser, 2009) are detected elsewhere in the brain. This signal is then fed back to primary sensory areas to elicit a startle response independently of the location of the perturbation in sensory space (see discussion of Saleem et al., 2013). In order to examine the receptive field structure (and thereby stimulus specificity) of mismatch neurons in V1 of mice navigating a virtual tunnel, self-generated visual flow can be perturbed in local areas of the visual field. By doing so, the following study confirms the existence and describes the properties of mismatch receptive fields in primary visual cortex of mice.

Mismatch receptive fields in mouse primary visual cortex

To determine if mismatch responses in layer 2/3 neurons of mouse primary visual cortex have receptive fields, we used a virtual-reality system that allowed us to manipulate visual feedback selectively in small parts of the visual field. Mice were head-fixed on a spherical treadmill (Dombeck et al., 2007) while we recorded neuronal activity by two-photon imaging of the genetically encoded calcium indicator GCaMP6f (Chen et al., 2013). Experiments always commenced with a closed-loop session, during which the rotation of the treadmill was coupled to movement in a virtual tunnel. In these closed-loop sessions, mice would self-generate visual-flow feedback during locomotion. To probe for receptive fields of mismatch signals, we perturbed visual-flow feedback at random times both globally, by halting visual flow throughout the entire visual field (full-field perturbation), and locally, by halting visual flow in one of 42 locations (local perturbation; **Figures 4A-C**). We used halts of visual flow as perturbation stimuli to minimize any potential direct visual drive of the perturbation. Passively observed full-field visual flow halts have been shown to lead to no measurable visual response in primary visual cortex (Keller et al., 2012). The walls and ceiling of the tunnel were lined with a random-brick pattern that allowed us to halt visual flow locally without changing the texture of the visual stimulus. To quantify visually evoked responses, we replayed the visual stimulus generated in the preceding closed-loop session including local and full-field perturbations, in subsequent open-loop sessions. In these open-loop sessions, the treadmill was uncoupled from the visual environment, but mice were free to run. We refer to local and full-field perturbations that the mouse passively viewed during open-loop sessions as local and full-field playback-halts. Note that for the analysis that follows, quantification of visual responses in open-loop sessions was restricted to periods when the animals did not run to enable quantification of visual responses to these stimuli in a passively viewing animal.

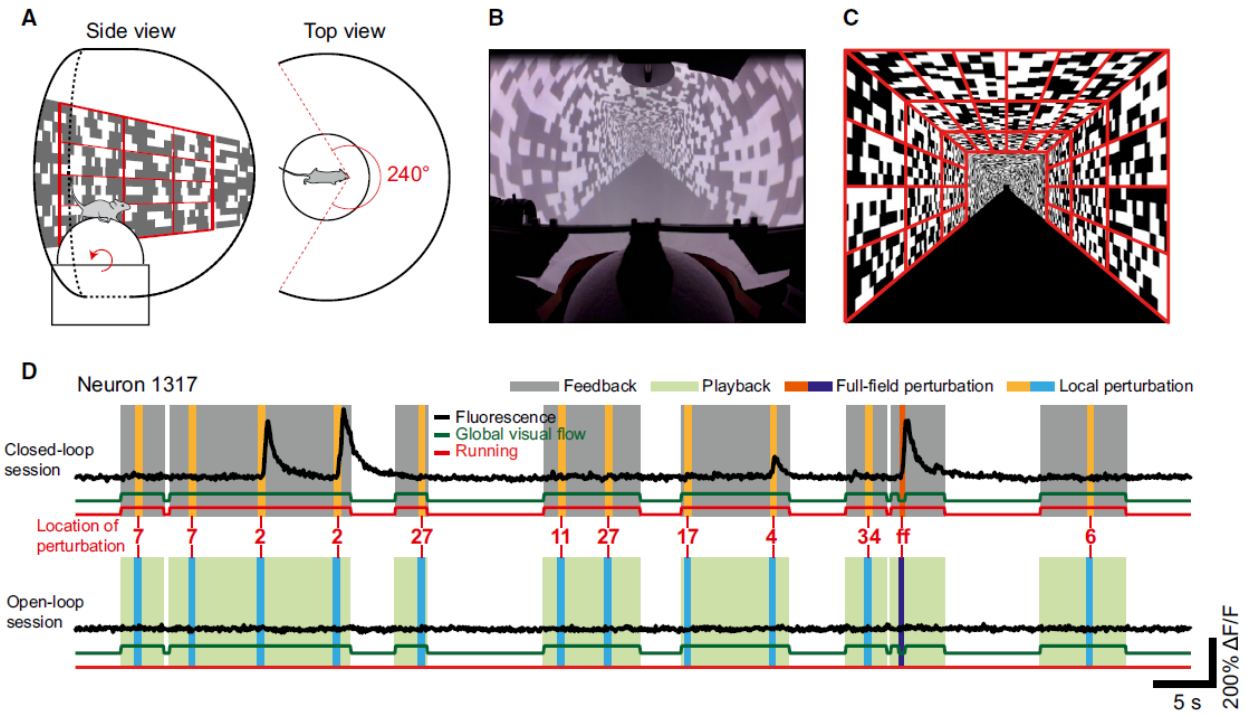

Figure 4. Experimental setup and paradigm. (A) Schematic of the virtual-reality system. Walls and ceiling of the virtual tunnel were textured with a random-brick pattern. (B) Photograph of the virtual tunnel from a point of view behind the mouse. (C) Schematic illustration of the random-brick pattern highlighting the 42 locations (4 by 3 on both walls, and 6 by 3 on the ceiling) sampled with local perturbations. (D) Top: Fluorescence response of one local mismatch neuron to local (light orange shading) and full-field (dark orange shading) perturbations. Gray shading indicates times of feedback. Red and green lines are binarized running and visual flow. Bottom: Fluorescence response of the same neuron to a playback of the same visual stimulus (including local and full-field playback-halts in light blue and dark blue respectively). Green shading indicates times of playback. Numbers label local perturbation identity as shown in (**Figure 5A**). Full-field perturbations are indicated as ff. Note that responses to mismatch were quantified on closed-loop data only. Playback-halts in open-loop sessions were analyzed only when mice were not running.

Mismatch neurons have receptive fields

Consistent with previous reports (Keller et al., 2012; Saleem et al., 2013) (however, see the discussion of Saleem et al. for a different interpretation of their results), we found that a subset of neurons in layer 2/3 responded selectively to feedback mismatch (242 of 2179 neurons, 11.1%). We classified a neuron as selective for mismatch if its response to mismatch is significant compare to baseline and at least twice as large as any other response (see Methods; in the following we refer to these mismatch-selective neurons as mismatch neurons). Analyzing the responses to local perturbation, we found that 89.7% (217 of 242 neurons) of mismatch neurons had a significant response to local mismatch in at least one location (**Figures 4D and 7A**). We refer to the part of the visual field in which a neuron responds to local mismatch as the mismatch receptive-field of a neuron. Mismatch receptive-fields were typically confined to one region of the visual field and encompassed only a small number (3 ± 0.1, mean ± SEM; **Figures 5A and A1A** – figures marked with '**A**' can be found in the Appendix) of the 42 possible locations. In 45% (109 of 242 neurons) of mismatch neurons, responses were selective for local mismatch (defined as a local mismatch response at least three times larger than the corresponding visual response; see Methods; **Figure A1B**). On average, the amplitude of peak local mismatch responses was similar to full-field mismatch responses (mean local mismatch: 14.9% ΔF/F, mean full-field mismatch: 11.9% ΔF/F, $p = 0.22$, paired Student's t test; **Figures 5B and A1C**). Both local and full-field mismatch responses were significantly larger than local and full-field playback-halt responses respectively (mean local playback-halt: 2.1% ΔF/F, $p < 0.01$, paired Student's t test; mean full-field playback-halt: 0.1% ΔF/F, $p < 0.01$, paired Student's t test).

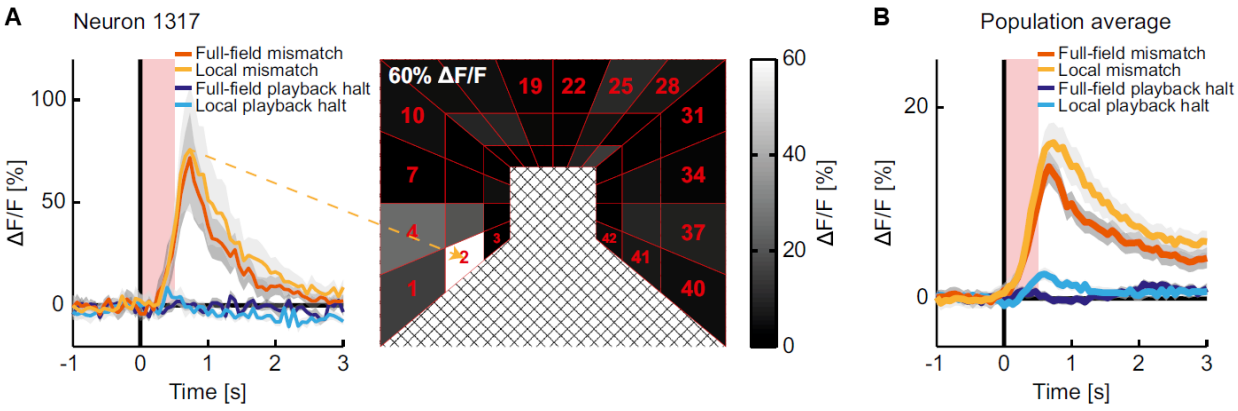

Figure 5. Selective local-mismatch detectors in mouse primary visual cortex. (A) Left: Average full-field (dark orange: mismatch, dark blue: playback-halt) and peak local (light orange: local mismatch, light blue: local playback-halt) perturbation responses of the neuron shown in (**Figure 4D**). Arrow indicates tunnel location of the peak local perturbation responses shown on the left. Pink shading indicates the duration of full-field and local mismatch. Grey shading indicates SEM. Right: Response map of the average local mismatch response of the example neuron. Numbers indicate local perturbation identity. Note that numbering is partially omitted but continues as indicated. Peak response amplitude is indicated in the top left corner of the response map. (B) Average full-field and peak local-mismatch responses, as well as full-field and local playback-halt responses of selective local-mismatch neurons (n = 109 neurons).

To compare mismatch responses to visual responses, we analyzed onset responses to feedback, playback, full-field mismatch and full-field playback-halts. Neurons were classified as visual if visual flow onset responses were on average larger than 3% ΔF/F, significant as measured by Student's t-test, and the neurons were not classified as mismatch selective (1433 out of 2179 neurons; see Methods; **Figure 6 and Table A1**).

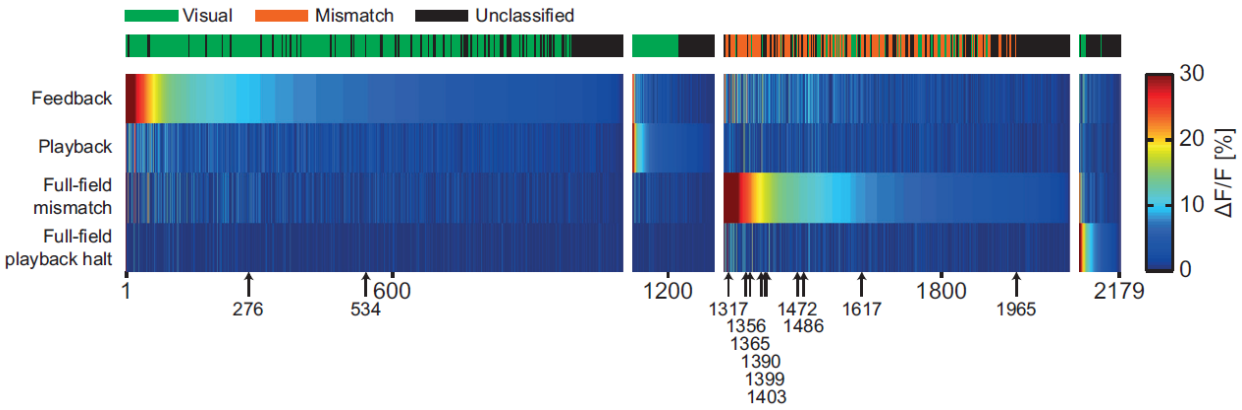

Figure 6. Visual and mismatch neuron classification. Average responses to feedback, playback, full-field mismatch and full-field playback-halt onsets for all neurons (n = 2179). Neurons were sorted based on their peak responses to these stimuli and classified as visual, mismatch, or unclassified (see Methods; green: visual neurons; orange: mismatch neurons; black: unclassified). Arrows indicate example neurons from **Figures 4D, 7A, 9A, A1C and A2**.

We found that 84.3% (1208 out of 1433 neurons) of these neurons also had significant responses to local stimuli (local perturbations and local playback-halts were pooled; see Methods; **Figures 7A and A2**). As expected, these local responses were confined to small portions of the visual field in both closed-loop and open-loop sessions consistent with retinotopic organization (**Figures 7A and A2**). On average, peak local responses of visual neurons were significantly larger than responses to the corresponding full-field stimuli (mean local mismatch: 12.1% ΔF/F; mean full-field mismatch: 3.4% ΔF/F; $p < 0.01$, paired Student's t test; mean local playback-halt: 7.2% ΔF/F; mean full-field playback-halt: -0.5% ΔF/F; $p < 0.01$, paired Student's t test; **Figure 7B**).

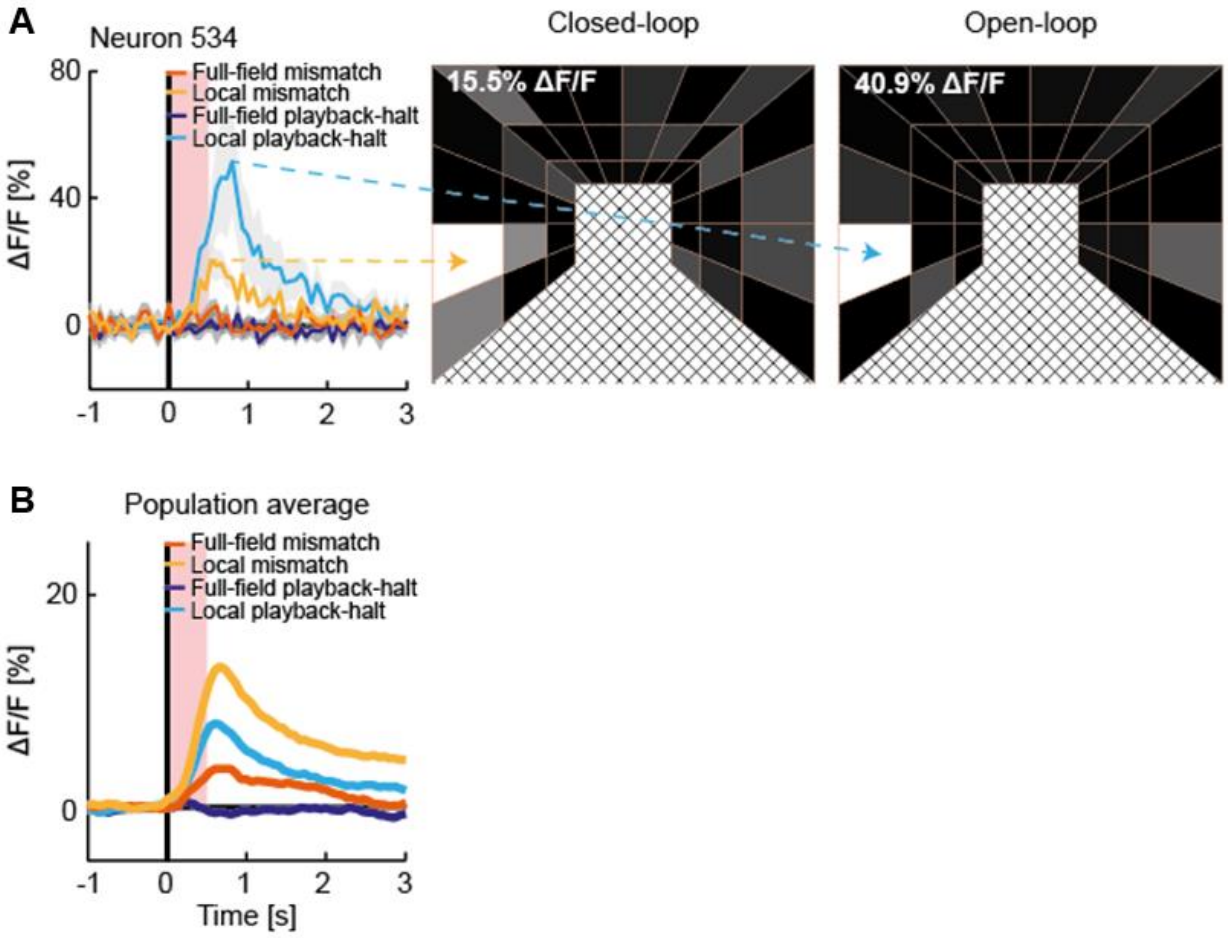

Figure 7. Visual neurons respond to local mismatch and local playback-halts in confined regions of the visual field. (A) Example visual neuron with local responses. Response maps show the spatial location of the peak local response during closed-loop (middle panel) and open-loop sessions (right panel). (B) Average full-field and peak local-mismatch responses, as well as full-field and local playback-halt responses of visual neurons with significant local responses (n = 1208 neurons; colors as in (A)). Note, SEM is smaller than the line width of the curves plotted and is thus occluded.

Note that the classification of neurons as *mismatch* or *visual* was done merely to simplify analysis and presentation of the results. We found no evidence for a bimodal distribution that would justify classification of distinct

functional cell types. Neurons were distributed continuously between the two extremes of being purely mismatch-selective and purely visual (**Figure 6**).

Mismatch receptive-fields are of the same size as visual receptive-fields and are aligned to the retinotopic map

Using local perturbations to quantify receptive field sizes of visual responses (see Methods), we found that visual receptive-fields had a median half-width-at-half-maximum (HWHM) of 11.1 degrees, consistent with previous reports (Bonin et al., 2011; Niell and Stryker, 2008). Similarly, using responses to local perturbations during visual-flow feedback to quantify mismatch receptive-field size, we found that the distribution of mismatch receptive-field sizes was comparable to that of visual receptive-fields with a median of 10.6 degrees (**Figure 8A**). Thus, the spatial resolution of local mismatch responses matches that of visual responses.

To test if local mismatch responses are arranged in a topographic map and aligned with the retinotopy of visual responses, we compared mismatch-response maps to both intrinsic optical imaging maps of visual responses as well as visual-response maps measured by local perturbations. The location of the peak of the average mismatch-response map of all mismatch neurons within one imaging region (250 µm x 300 µm) corresponded well to the retinotopic location of the imaging region within primary visual cortex (**Figure 8B**). Moreover, for each imaging region, the average mismatch-response map of mismatch neurons correlated better with the average visual-response map of visual neurons in the same imaging region than with visual-response maps of different imaging regions (8 imaging regions; $p < 0.01$, paired Student's t test), or randomly shifted visual-response maps (**Figures 8C and A3**). Thus, mismatch receptive-fields are arranged topographically, aligned to the map of visual space in visual cortex.

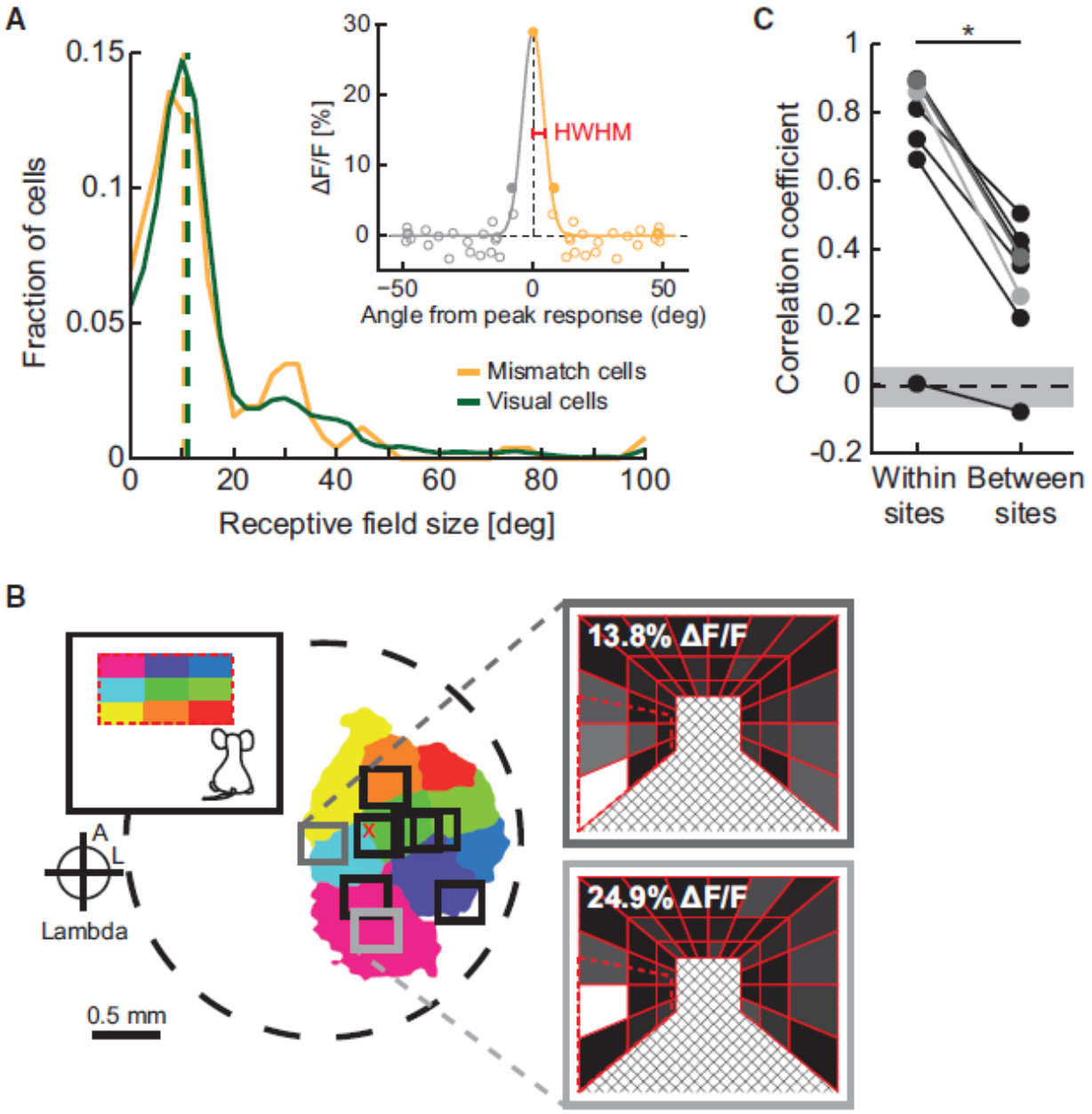

Figure 8. Size and organization of mismatch receptive fields. (A) Distribution of receptive field sizes for mismatch neurons (n = 109, light orange line) and visual neurons (n = 1208, green line) measured as half-width-half-maximum (HWHM). Dashed lines indicate median receptive field sizes (10.6 and 11.1 degrees HWHM for mismatch and visual neurons, respectively). Inset: Average local mismatch responses (light orange) as a function of visual angle from location of peak mismatch response for an example mismatch neuron. To fit a Gaussian curve (solid line), data are reflected along the y-axis (in grey). HWHM of the Gaussian fit is used to quantify size of the

receptive field. Filled (open) circles indicate significant (not significant) local responses (see Methods). (B) Left: Schematic of a cranial window (dashed circle) indicating location of all imaged regions in relation to an average intrinsic optical imaging map (see Methods). Marker on the left indicates the approximate position of lambda, and anterior (A) and lateral (L) directions. Red 'x' indicates the location of neuron 1315 shown in **Figures 4D and 7A**. Inset: color coding of intrinsic optical imaging responses by stimulus location. Right: Response maps of local mismatch neurons from two imaging regions at different retinotopic locations. Peak responses are indicated in the top left corner of the response maps. Note that intrinsic optical imaging was done using a monitor and only covered part of the visual space mapped with visual flow perturbations in the virtual reality setup (dashed red line marks approximated correspondence). (C) Correlations of average mismatch- and visual-response maps within the same imaging regions and between different imaging regions. Note that, as our imaging regions were confined to a part of visual cortex easily accessible for imaging, correlations between different regions were larger than expected by chance ((B) and **Figure A3**). Shades of grey correspond to the two imaging regions marked in (B). The dashed line is chance correlation based on randomly shifted visual-response maps (see Methods); shading marks standard deviation; $p < 0.01$ paired Student's t test.

Mismatch responses signal a difference between predicted and actual visual flow

Locomotion has been shown to amplify visual responses by a factor of roughly two to three (Niell and Stryker, 2010; Keller et al., 2012; Saleem et al., 2013). This effect can be described as a running-induced gain of visual responses. To test if such a running-induced gain could account for the difference between mismatch and playback-halt responses in mismatch neurons, we quantified average running-induced gain for perturbation stimuli and visual stimuli in active versus passive conditions (mismatch versus playback-halts, and feedback versus playback; see Methods; **Figure 9**). We found that the running-induced gain associated with visual stimuli in our data was comparable to previous reports (Niell and Stryker, 2010; Saleem et al., 2013) (**Figure 9A**), both on a

population level and when only looking at visual neurons. However, we found that the average running-induced gain associated with a full-field perturbation (full-field mismatch versus full-field playback-halt) was almost five times larger than the average running-induced gain of a visual-flow stimulus on a population level (feedback versus playback; mismatch gain: 9.9; visual gain: 2; p = 0.019, one-tailed paired Student's t test; **Figure 9A**). Similarly, for local mismatch neurons, we found that the average running-induced gain between local mismatch and local playback-halt responses (light orange and light blue responses respectively in **Figure 5B**) was significantly higher than that observed for visual responses (mismatch gain: 7.3; visual gain: 1.7; p = 0.015 one-tailed paired Student's t test). By contrast, for visual neurons we observed that the average gain between local mismatch and local playback-halt responses (light orange and light blue responses respectively in **Figure 7B**) was not significantly different to that observed between feedback and playback responses (mismatch gain: 1.7; visual gain: 2; p = 0.15, paired Student's t test). Thus, mismatch responses cannot be explained by a uniform running-induced response gain.

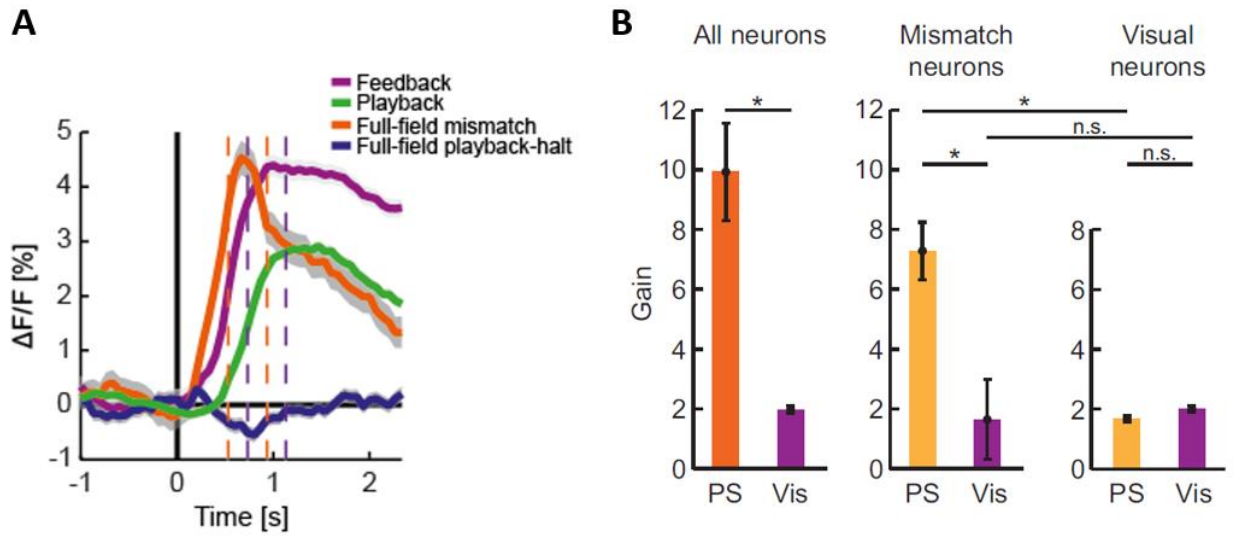

Figure 9. Mismatch responses cannot be explained by a uniform running-induced gain. (A) Responses of all neurons (n = 2179) to the onsets of feedback (purple), playback (green), full-field mismatch (dark orange), and full-field playback-halts (dark blue). Dashed lines indicate response windows used to measure running-induced

gain for (B) (feedback and playback response window (purple): +700 ms to + 1150 ms; mismatch and playback-halt response window (dark orange): +450 ms to +900 ms; see Methods). (B) Left: Average running-induced gain of a full-field perturbation response (dark orange: full-field mismatch vs. full-field playback-halt; PS) and of a visual response (purple: feedback vs. playback; Vis) for all neurons (n = 5; p = 0.019, one-tailed paired Student's t test; see Methods). Middle and right: Average running-induced gain of a local perturbation response (light orange: local mismatch vs local playback-halt; PS) and of a visual response (purple: feedback vs playback; Vis) for mismatch neurons (n = 5, middle; p = 0.015, one-tailed paired Student's t test), and visual neurons (n = 5, right; p = 0.15 paired Student's t test). The difference in running-induced gain of a local perturbation response between mismatch and visual neurons was significant (n = 5, p = 0.003, paired Student's t test), while the difference in gain of a visual response was not (n = 5, p = 0.305, paired Student's t test).

To assess if mismatch responses could be the consequence of a perturbation induced behavioral change, we analyzed pupil dilation, pupil position, and running speed changes triggered on either full-field or local mismatches. Consistent with previous findings (Keller et al., 2012), we found that full-field mismatch led to a decrease in running speed and an increase in pupil diameter. Both of these responses however, lagged neuronal responses to mismatch by 1540 ms and 1176 ms respectively (neuronal mismatch response latency: 150 ms; running mismatch response latency: 1690 ms; pupil mismatch response latency: 1326 ms). Moreover, even though neuronal responses to local mismatch were comparable to full-field mismatch, local mismatch did not lead to any detectable behavioral response (**Figures 10A and 12B**). The presence of behavioral responses to full-field and not local mismatch is consistent with the idea that a full-field mismatch is a more salient stimulus, in that it activates a larger retinotopic area of visual cortex.

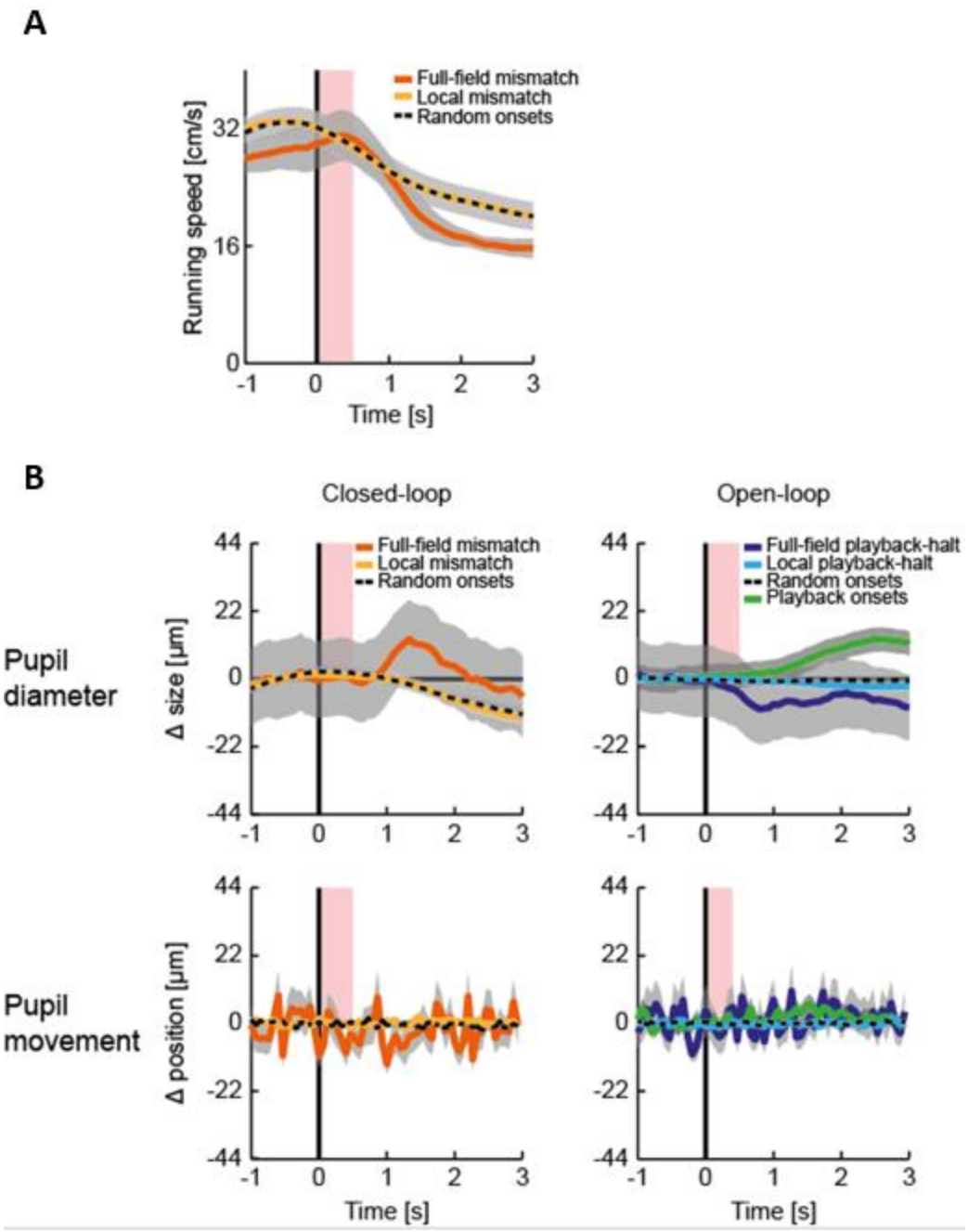

Figure 10. Mismatch responses cannot be explained by behavioral changes induced by local or full-field perturbations. (A) Average running profiles triggered on full-field mismatch (dark orange), local mismatch (light orange), and randomly selected onsets during running (dashed black line). (B) Left: Changes in average pupil diameter (top) and position (bottom) during closed-loop sessions triggered on full-field mismatch (dark orange line), local mismatch (light orange line), and randomly selected onsets during running (dashed black line). Right:

Average pupil diameter (top) and position (bottom) changes during open-loop sessions, triggered on full-field playback-halts (dark blue line), local playback-halts (light blue line), playback onsets (green line), and randomly selected onsets during stationary periods (dashed black line). Note that only full-field stimuli elicited significant pupil responses, although in opposite directions depending on the condition (full-field mismatch led to an increase while full-field playback-halts led to a decrease in pupil diameter). This is consistent with the idea that full-field visual flow halts are particularly salient during running (generating full-field mismatch), and not during passive viewing (as is the case for full-field playback-halts).

To signal the magnitude of the difference between predicted and actual visual flow, mismatch responses should scale with the difference between running and visual flow speed. We quantified this in two separate ways by analyzing mismatch responses as a function of running speed, and by analyzing responses of mismatch neurons during open-loop sessions, during which running and visual flow are not coupled. First, to investigate whether local-mismatch responses scale with the magnitude of the difference between running and visual flow speeds, we quantified local-mismatch responses as a function of running speed prior to local perturbations. We found that the average amplitude of local mismatch responses increased with increasing running speed of the mouse just prior to the perturbation onset (**Figure 11**). This was true for both full-field and local mismatch responses.

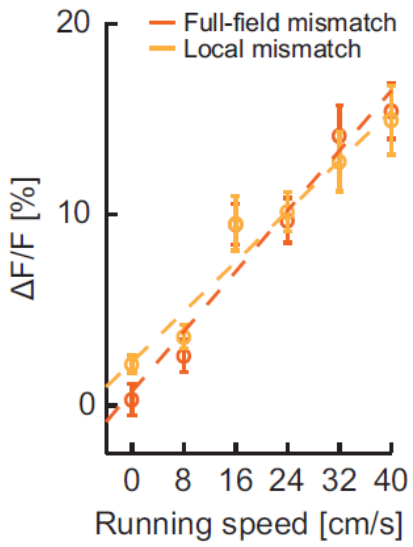

Figure 11. Local and full-field mismatch responses scale with running speed. Average local (light orange) and full-field (dark orange) mismatch responses as a function of the running speed of the mouse in a 1 s time window prior to mismatch onset. Error bars represent SEM. Dashed lines are linear fits to the data.

Second, to test if mismatch responses report a difference between predicted (*P*) and actual (*V*) visual flow (mismatch response = *P*-*V*) when both running speed and visual flow speed vary continuously, we measured responses of mismatch neurons during open-loop sessions (**Figure 12A**). To quantify the interaction between running and visual flow speeds that best explained mismatch responses, we fit a first-degree polynomial under a sigmoid nonlinearity of the form $s(aP + bV)$ to individual mismatch neuron responses (n = 242; see Methods). The sigmoid nonlinearity was used to account for the saturation of calcium responses at high activity levels, and significantly improved the quality of the fits. We minimized the squared error between data and model with a grid search over the parameters *a*, and *b*, and found that mismatch responses were on average best fit by a model that scaled positively with running speed and negatively with visual flow speed (a = 0.012 ± 0.003, b = -0.009 ± 0.003, mean ± SEM; **Figures 12B and 14C**). This shows that mismatch responses scale with the difference between predicted and local visual flow when predicted visual flow exceeds actual visual flow. Note that this is consistent

with our classification of mismatch neurons based on responses to full-field perturbations during closed-loop sessions.

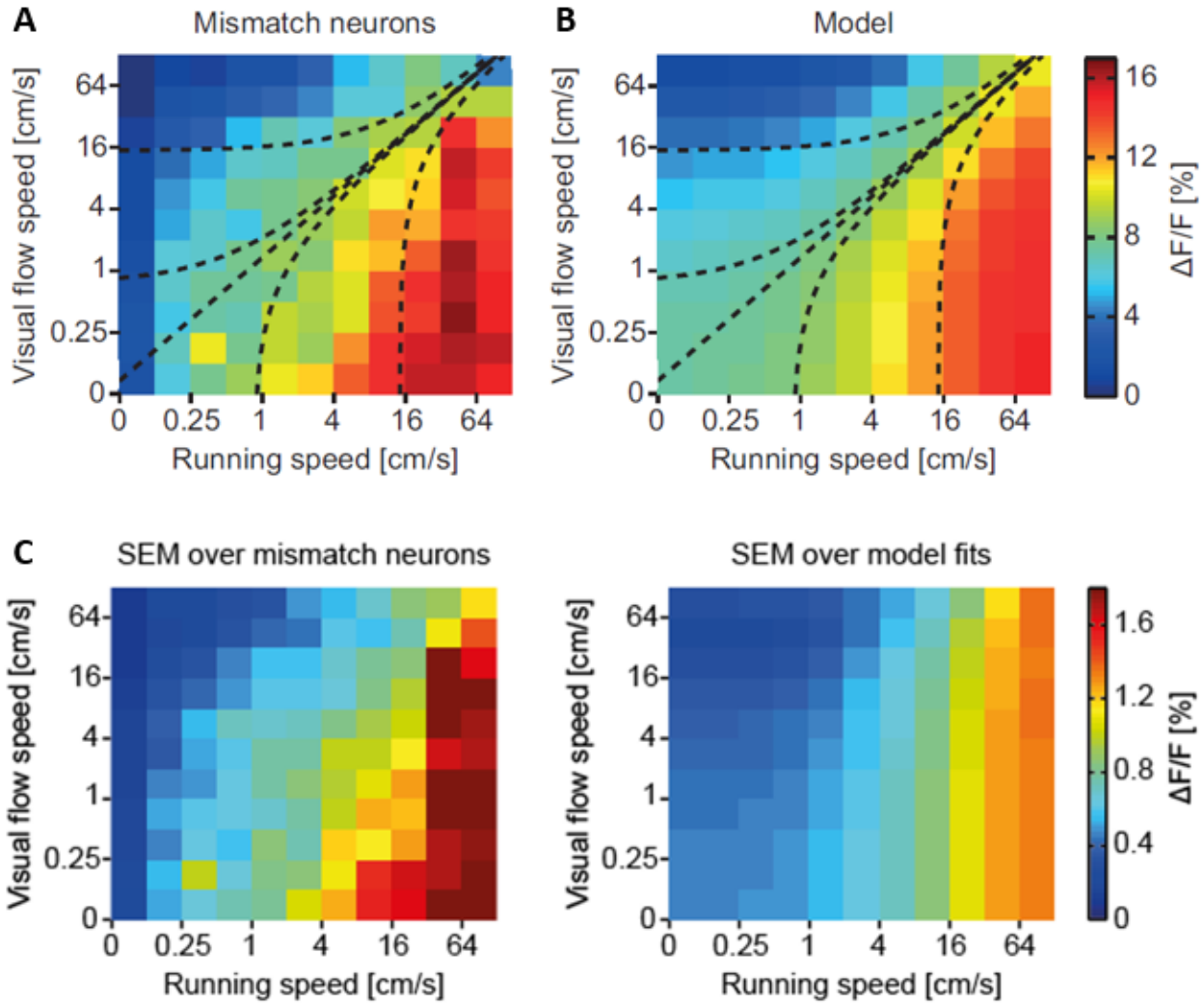

Figure 12. Mismatch neurons signal local deviations between predicted and actual visual flow. (A) Average mismatch neuron responses (n = 242; averaged over individual mismatch neurons) as a function of combinations of visual flow (*V*) and running speeds (predicted visual flow, *P*) during open-loop sessions. Dashed lines are isocontour lines of the average model fit ($s(aP + bV)$) with the parameters a = 0.012 ± 0.003 and b = -0.009 ± 0.003 (mean ± SEM), as shown in (D). This is consistent with a subtractive interaction of the form *P*-*V*. (B) Average model responses averaged over fits to the responses of individual mismatch neurons. (n = 242; see Methods). (C)

Standard error of the mean (SEM) of mismatch neuron (left) and model (*s(aP + bV)*; right) responses to combinations of visual flow (*V*) and running speeds (predicted visual flow, *P*) during open-loop sessions (see Methods). Note that the SEM is a factor of 10 smaller than the average mismatch responses.

Discussion

Predictive coding theories of cortical function stipulate the existence of an internal model that is used to generate 'top-down' predictions of sensory stimuli that can be compared to 'bottom-up' sensory input to detect deviations from predictions. For example, an internal model of the sensory consequences of locomotion could be used to generate predictions of self-generated visual flow that could be relayed to early visual areas for comparison with feedforward visual input to detect sensorimotor mismatch (Keller et al., 2012). This work has shown that such a comparison occurs at the level of receptive fields in primary visual cortex (**Figure 5A**), thereby integrating classical, feedforward visual processing schemes based on receptive fields in the framework of predictive coding. The topographic organization of local mismatch signals matches that of visual signals (**Figures 8B and C**), suggesting that local mismatch signals are computed based on local visual signals (e.g. visual neurons that respond to visual flow in a retinotopicaly matched area of the visual field). Functionally, such local mismatch signals could be used to detect externally moving objects during self-motion or provide visual feedback of limb movements during sensorimotor learning, which is supported by the fact that local mismatch signals match local visual signals in magnitude (compare **Figures 5 and 9**) and acuity (**Figure 8A**).

Since local mismatch signals are spatially restricted and do not respond to mismatch events outside of their receptive fields mismatch computation cannot be an outcome of a global 'startle' response to unexpected stimuli. This is further supported by the fact that on average local perturbations did not elicit detectable behavioral responses (**Figure 10**), yet the magnitude of local mismatch responses could fully account for responses to full field mismatch (**Figure 5B**) that did elicit a behavioral response (**Figure 10**). A running-induced gain of visual responses that has been suggested to account for increased responses of visual neurons during locomotion (Niell and Stryker, 2010) could also not account for the gains observed for full field or local mismatch responses (**Figure 9B**). This is consistent with the heterogeneous effects of locomotion on V1 neuron responses observed before (Keller et al., 2012; Saleem et al., 2013). Furthermore, mismatch responses to combinations of continuously

varying visual flow and running speed were best explained by a model whereby an excitatory predictive drive (P) is inhibited by actual visual input (V; mismatch response = P-V; **Figure 12**) consistent with the proposed mismatch computation these neurons perform based on responses to local visual flow halts in the coupled condition. These data suggest that primary visual cortex topographically integrates predictions of visual flow and actual sensory input to signal local deviations from expected sensory input.

Where do visual flow predictions originate from?

Neurons in superficial layers of primary visual cortex receive diverse inputs from cortical and sub-cortical regions as well as from deeper layers within V1 (Harris and Mrsic-Flogel, 2013; Harris and Shepherd, 2015)(**Figure 2**) that could potentially relay visual flow predictions to V1. One possible source of visual flow predictions in V1 are long range cortical projections from anterior cingulate cortex (ACC) and secondary motor cortex. ACC inputs have been shown to modulate surround suppression (Zhang et al., 2014) in V1 – an effect that is not inconsistent with predictions of visual flow since responses to predictable input would be suppressed by predictions (Rao and Ballard, 1999). ACC projections to V1 have been shown to convey visual stimulus predictions for learned spatial sequences of stimuli (Fiser et al., 2016). Additionally, unpublished data from our lab has shown that inactivation of ACC leads to a decrease of mismatch responses in V1 (data from Marcus Leinweber).

However, other sources of input may also relay predictions to V1. Retrosplenial cortex that has been shown to harbor representations of self-motion (Cho and Sharp, 2001) also sends direct projections to V1 (**Figure 2**). Thalamic inputs to V1 from the lateral posterior nucleus (LP, pulvinar equivalent in mice) and LGN are modulated by locomotion, with the LP projection enriched for sensorimotor mismatch signals (Roth et al., 2016). Finally, neuromodulatory inputs have been implicated in mediating locomotion influences in V1 (Fu et al., 2014; Lee et al., 2014; Pinto et al., 2013; Polack et al., 2013). Given that predictions of visual flow need to have a high temporal

and spatial resolution (Figure 5A) neuromodulatory signals seem unlikely to have sufficient resolution, although their modes of action may be faster than previously thought (Howe and Dombeck, 2016; Pinto et al., 2013).

Do stimuli that move faster than the observer generate mismatch signals?

The local mismatch neurons described in this study signal deviations form expectations when predicted visual flow (P; based on the running speed) exceeds actual visual flow (V; **Figure 12**). Are there mismatch neurons that perform the symmetric computation when actual visual flow exceeds predicted visual flow (V-P)? The current study lacks the appropriate stimuli to address this question since although such neurons would be expected to respond to playback (but not to any other) onsets (**Figure 6**) these would be indistinguishable from purely visual responses driven by feedforward input (where P = 0). Such responses are never the less expected to exist under predictive coding schemes (Rao and Ballard, 1999), and a paradigm that could probe for such signals would require speeding up the visual flow during locomotion. Neurons that would respond to such perturbations and scale with difference between predicted and actual visual flow when actual visual flow exceeds predicted visual flow would fulfil the criteria for a V-P computation.

Further interesting questions arise when considering predictions of visual flow in more directions than forward. In the head fixed preparation used for this study (**Figure 4A**), the problem of not knowing what the mouse could be predicting is constrained by having the mice only run in one direction (mice rarely run backwards on the treadmill). However, how does the mismatch system handle predictions of visual flow when generated by locomotion in any direction, or by unrestrained head and body movements? Are mismatch neurons in V1 invariant to the orientation of predicted visual flow and hence only concerned with their preferred mismatch direction relative to predictions (i.e. V-P versus P-V)? Since error propagation in models of predictive coding has a hierarchical structure (Rao and Ballard, 1999), it is possible that mismatch computations in V1 are limited in their information content to only signal the location of mismatch (Figure 4A) and direction of mismatch relative to the

prediction (**Figure 12**), so that 'mismatch orientation' is computed upstream, perhaps in secondary visual areas. These questions remain to be addressed.

How does the visual system deal with depth?

In this study, the average spatio-temporal frequency of the visual stimulus was kept constant throughout the experiments – however, in the real world, irregularities of objects as well as perspective in the visual scene will generate a range of local visual flow speeds that will not match the predicted visual flow speed given locomotion. These effects could in principle lead to local mismatches for stationary objects – since a distant stationary object will generate a different visual flow to a nearby stationary object. How can the mismatch system avoid saturation in naturalistic environments? A purely speculative implementation could be that local mismatch is in fact used to compute depth of objects in the visual scenes in the first place. Consider moving along a side of a building – the building will generate a range of visual flow speeds that will predictably be faster at the periphery than towards the vertical meridian. This would generate a 'gradient' of local mismatch signals that sample the retinotopic part of the visual field that the building is located in. In this simple example, such a gradient can be used to measure distance of the object from the observer. Ultimately, since the building is stationary, the predictions of visual flow will likely account for the depth of the object, but these low level mismatch representations in V1 could be crucial in constructing the internal model necessary to make such predictions. Given the fact that we are aware which objects are stationary in a given visual scene very quickly, this whole process would have to also be very fast. An interesting speculation is that this could be accomplished if the internal model was housed locally in V1, for example in deeper layers, and the entire computation was performed within the connections between superficial and deep V1 neurons. Of note is also the fact that temporal correlations of the visual scene (Huang and Rao, 2011) during movement are necessary to establish whether an object is stationary or not. Interestingly humans

continuously scan the visual scene, even during fixation (Rolfs, 2009). Such microsaccadic movements could underlie the sampling of the visual scene that is necessary for fast computing of depth in the visual scene.

Conclusion

This study proposes that neurons in superficial layers of mouse primary visual cortex topographically integrate predictions of visual flow with actual sensory input to signal local deviations of predictions. Given this computation, the classical concept of feedforward representations based on receptive fields can be expanded into the wider theoretical framework of predictive coding. However, it remains to be seen whether this interpretation also holds for other sensory areas, and whether cortical function as a whole can be understood in the framework of predictive coding.

Methods

Subjects and cranial window implantation surgery

All experiments were performed on five 57-72 day old female C57BL/6J mice, and were conducted in accordance with the guidelines of the Veterinary Department of the Canton Basel-Stadt. For the window implantation surgery, animals were deeply anesthetized with FMM, a mixture of Fentanyl (0.05 mg/kg; Actavis), Midazolam (5.0 mg/kg; Dormicum, Roche) and Medetomidine (0.5 mg/kg; Domitor, Orion). An approximately 4 mm diameter craniotomy was made over right visual cortex. Four to six injections of AAV2/1-*ef1a*-GCaMP6f (approximately 200 nl per injection; AAV titer: 1.8×10^{12} GC/ml) were made in the medial/anterior part of visual cortex. A 3 or 4 mm diameter glass coverslip was implanted using superglue (Pattex). The skull surface was made rough by scratching with a hypodermic needle and subsequently covered with histoacrylic (Braun). A custom made headplate was attached to the skull using dental acrylic (Paladur, Heraeus). Mice were then returned to their home cage and housed in groups of three to five animals with continuous access to a running wheel.

Intrinsic optical imaging

Twelve to twenty-one days after implantation animals were anaesthetized with FMM (half surgical dose) for window cleaning and intrinsic optical imaging. For intrinsic optical imaging, mice were head fixed and presented with drifting sinusoidal gratings of random orientations in 18 locations (3x3 per monitor) on two monitors (145 cm diagonal, one monitor per eye) positioned 30 cm away from the eyes. Each stimulus location was probed 5 or 10 times for 4 s each with an inter stimulus interval (gray screen) of 5 s in random sequence. The brain was illuminated with a red LED, and the emission was filtered with a 610/30 nm bandpass filter (Semrock). Images were acquired with a 12bit ORCA-05G CCD camera (Hamamatsu). Data were analyzed as previously described (Schuett et al., 2002), and an intrinsic optical imaging map was generated for each animal. For each map,

responses to gratings at individual locations were divided by the most responsive pixel. In order to calculate the average map across all animals (**Figure 8B**), all normalized maps were aligned to each other using the response to the center of the 3x3 stimulus pattern (most responsive location across all animals). A winner-take-all algorithm was employed to calculate boundaries between areas responding to different spatial locations.

Two photon microscope

All calcium imaging experiments were performed on a custom built two-photon microscope. Illumination source was a tunable femtosecond laser (Insight, Spectra Physics) tuned to 930 nm. Emission light was band-pass filtered using a 525/50 filter (Semrock) and detected using a GaAsP photosensor (H7422, Hamamatsu). PMT 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 microscope was based on a 12 kHz resonant scanner (Cambridge Technology). Images were acquired at a resolution of 750 x 400 pixels (60 Hz frame rate), and a piezo-electric linear actuator (Physik Instrumente) was used to acquire images in 4 different imaging planes in alternation. This resulted in an effective frame rate of 15 Hz. The field of view was 300µm x 250µm.

Experimental paradigm

All calcium imaging experiments were performed in the time window of 22 to 31 days after implantation and viral transduction. Animals were imaged for 3 or 7 sessions in a virtual reality setup as previously described (Leinweber et al., 2014). During imaging sessions, animals were head-fixed on a spherical treadmill and surrounded by a toroidal screen. A virtual tunnel was projected onto this screen. Walls and ceiling of the tunnel were lined with a random-brick pattern (**Figures 4B and 6C**). This was done to enable local halting of visual flow in one of 42 predefined locations without perturbing the statistics of the visual stimulus. Local halts of visual flow varied in size

depending on location in the tunnel (vertical/horizontal visual angle; largest: 15°/60°, smallest: 6.25°/10°). Each imaging session consisted of one to three 8 or 13 minutes closed-loop sessions (22.6 ± 0.7 minutes, mean ± SEM) during which movement of the virtual tunnel (visual flow) was coupled to the locomotion of the mouse on the treadmill. To generate mismatch events, visual flow was halted either locally in one of the 42 predefined locations or full-field throughout the visual field. These visual flow perturbations lasted 0.5 s and were interspersed randomly at an inter-perturbation interval range of 2.5 to 7.5 s. Subsequently there were one to five 8 or 13 minute open-loop sessions (27.1 ± 1.1 minutes, mean ± SEM) during which the visual flow generated during a previous closed-loop session, including local and full-field flow halts, was replayed. Note, that to minimize potential habituation effects, individual closed- and open-loop sessions were interleaved. On average, this resulted in 2 local perturbations per location and imaging session. In total, the average recording time per imaging region was 4.1 ± 0.7 hours (mean ± SEM). During all experiments, running speed of the mouse, visual flow speed and videos of both pupils were recorded. Pupil position and diameter were measured as previously described (Leinweber et al., 2014).

Image analysis

Raw images from eight regions of layer 2/3 monocular visual cortex (**Figure 8B**) were full frame registered to correct for brain motion. Neurons were manually selected based on mean and maximum fluorescence images. A total of 2179 neurons were selected with an average of 272 ± 28 neurons per imaging region (mean ± SEM). Raw fluorescence traces were corrected for slow drift in fluorescence using an 8^{th}-percentile filtering with a 15 s window (Dombeck et al., 2007). ΔF/F traces were calculated as mean florescence in a selected region of every imaging frame and subsequently subtracted and normalized by the median fluorescence.

Data analysis

To functionally classify neurons, average responses to the onset of feedback, playback, full-field mismatch and full-field playback-halts were calculated as mean fluorescence in a response window minus mean fluorescence in a baseline window for each neuron (response window for feedback and playback onsets was +700 ms to +1150 ms, and the baseline subtraction window was -450 ms to 0 ms; for full-field mismatch and full-field playback-halts the response window was +450 ms to +900 ms, and the baseline subtraction window was -450 ms to 0 ms; **Figure 9A**). Feedback and mismatch onsets were measured in closed-loop sessions. Feedback onsets were defined as a 2 s period of sustained running (and therefore visual flow) in the absence of perturbations, following a 4 s stationary period. Full-field and local mismatch onsets were defined as a full-field or local perturbation, respectively, that occurred within a 2 s window of running respectively. Playback and playback-halt onsets were measured in open-loop sessions. Playback onsets were defined as a 2 s period of replayed visual flow in the absence of perturbations following a 4 s period of no visual flow. Full-field and local playback-halt onsets were defined as a full field or local perturbation, respectively, that occurred during a 2 s window of replayed visual flow. Data were excluded if animals ran during playback, or full-field and local playback-halt onsets, to enable quantification of visual responses to these stimuli in a passively viewing animal.

A neuron was classified as selective for mismatch if the average response to full-field mismatch was greater than 3% ΔF/F, significant ($p < 0.05$, paired Student's t test), and was at least twice as large as the maximum response to any of the other conditions (242 of 2179 neurons; see **Table A1** for the number of neurons per animal). To probe whether mismatch neurons responded to local stimuli, average responses to local perturbations during closed-loop and open-loop sessions for each of the 42 locations were calculated using the same response window criteria as for full-field mismatch (see above). A mismatch neuron was classified as responsive to local mismatch if its response to a local perturbation during closed-loop sessions was greater than 3% ΔF/F and significant for at least one of the 42 locations ($p < 0.05$ paired Student's t test; 217 of 242 neurons). A mismatch neuron was

classified as selective for local mismatch if its peak average response to local mismatch was three times the size of its response to local playback-halt at the same spatial location (109 of 242 neurons).

A neuron was classified as visual if its response to the onset of at least one among feedback, playback, full-field mismatch or full-field playback-halts was greater than 3% ΔF/F, the response was significant ($p < 0.05$, paired Student's t test), and the neuron was not classified as mismatch-selective as per criteria above (1433 out of 2179 neurons). To quantify local responses in visual neurons, responses to local perturbations during closed-loop and open-loop sessions were pooled for each of the 42 locations and average responses were calculated using the same response window criteria as for mismatch neurons. A visual neuron was classified as responsive to local stimuli if its average response to a local stimulus in at least one of the 42 locations was greater than 3% ΔF/F and significant ($p < 0.05$ paired Student's t test; 1208 out of 1433 neurons).

To visualize the spatial distribution of responses to local stimuli for individual neurons or groups of neurons, response profiles for all 42 local perturbation locations were projected onto a schematic of the virtual tunnel, a representation we refer to as "response map" (**Figures 5A, 9A, 10B, A1C, A2 and A3**). To generate average response maps for groups of neurons (**Figures 8B and A3**), the response maps of individual neurons were averaged. For comparison of response maps between mismatch and visual neurons (**Figures 8C and A3**), Pearson's linear correlation coefficients were calculated between average response maps of mismatch neurons and visual neurons from the same imaging region (**Figure A3**). As a control we computed the average correlation of the mismatch-response map of one imaging region with the visual-response map of all other imaging regions, and with randomly shifted variants of the visual-response maps.

To calculate receptive field sizes (**Figure 8A**) of mismatch and visual neurons that responded to local stimuli (mismatch neurons n = 109, visual neurons n = 1208), responses to local perturbations were plotted as a function of difference in visual angle from the location of peak response. Data were reflected along the y axis and a Gaussian fit was computed for each neuron individually. Quality of the fits were assessed by a coefficient of

determination (R^2) above 0.01 and alignment of the peak of the fit to the position of the peak response. Using these criteria, we could measure the receptive field size in 78.9% (86 out of 109 neurons) of mismatch neurons (R^2 = 0.2 ± 0.02, mean ± SEM) and 80.4% (971 out of 1208 neurons) of visual neurons (R^2 = 0.2 ± 0.01, mean ± SEM). Half-width half-maximum (HWHM) of the Gaussian fit was used to quantify the size of the receptive field. Receptive field size distributions in **Figure 8A** were smoothed with a window of 3 degrees.

To quantify average running-induced gains (**Figure 9B**), data were pooled from all animals and average responses to the onsets of feedback, playback, full-field and local mismatch and full-field and local playback-halts were calculated (using the same response window criteria as above; **Figure 9A**). An average running-induced gain was defined as (active response/passive response), and was calculated for (feedback/playback), (full-field mismatch/full-field playback-halt), and (local mismatch/local playback-halt) responses. Running-induced gains were calculated for all neurons (n = 2179; **Figure 9A**), local mismatch neurons (n = 109; **Figure 5B**) and visual neurons with significant local responses (n = 1208; **Figure 7B**). In the case of the average response of all neurons to a full-field playback-halt, the absolute value of the response was taken to avoid a negative gain. Standard errors of the mean for each gain and paired Student's t test between gains were calculated across individual animals (n = 5).

To calculate local and full-field mismatch responses as a function of the running speed of the animal (**Figure 11**), responses were binned by average velocity in a window 1 s prior to perturbation onset. Bins were centered on 0 cm/s to 40 cm/s in intervals of 8 cm/s with a width of ± 8 cm/s (apart from the first and last bins, which were defined as <8cm/s and >32cm/s, respectively). Linear fits were generated by a least squares fit to the data.

To quantify responses of mismatch neurons as a function of combinations of visual flow and running speeds, data from all open-loop sessions were pooled for individual neurons (n = 242; **Figure 12A** is the average over individual mismatch neurons). To avoid potential confounds, responses during full-field and local playback-halts were excluded. Visual flow and running speeds were binned, and mismatch neuron traces were averaged for various

combinations of these bins. Bin limits were chosen from a log scale distribution ranging from 0 cm/s to 128 cm/s, as sampling of visual flow and running speeds was consistently more uniform across bins chosen from a log scale than if chosen from a linear scale distribution (data not shown). Not all animals covered the entire range of speed bins, missing responses were interpolated by averaging responses from all adjacent bins. We used a first-degree polynomial under a sigmoid non-linearity of the form:

$$r = s(aP + bV)$$

where r is the predicted neuronal response. P and V are predicted and actual visual flow speeds (predicted visual flow corresponds to the running speed of the animal) respectively and were sampled with the same bins used for sampling the open-loop data (**Figures 12A and 14B**). Responses were fit by a grid search over the parameters a and b that were varied between -0.1 and 0.1 in steps of 0.002. s is a sigmoid saturation function of the form:

$$s(x) = \frac{A}{1 + e^{\frac{-x}{sat}}} + d$$

where d is an offset constant and was varied between -2 and 2 in steps of 0.1, A is an amplitude constant that was varied between -2.5 and 2.5 in steps of 0.05, and sat is a saturation constant. The sigmoid nonlinearity was used to account for saturation of calcium responses at high activity levels (mismatch responses saturated at high running speeds in open-loop sessions; **Figure 12A**). To decrease computation time, the value of sat was optimized in a coarse grid search over all parameters using a fit to pooled average responses of all mismatch neurons (n = 242), and was fixed at the optimum of 0.08 for all grid searches for individual neurons. We assessed model fits to the responses of individual mismatch neurons by minimizing the squared error of the fits. Model parameters that best fit mismatch neuron responses were: a = 0.012 ± 0.003, b = -0.009 ± 0.003, d = 0.049 ± 0.012, and A = 0.049 ± 0.024 (mean ± SEM; average least squared error = 0.783 ± 0.142; mean ± SEM; see **Figure S4D** for SEM of the mismatch neuron and model responses). Isocontour lines of the best fit were then superimposed on average mismatch and model responses (**Figures 12A and 14B**).

To quantify average pupil responses and running profiles following full-field and local perturbations (**Figures 10A and 12B)**, data were pooled from all animals and closed-loop sessions. Data containing eye blinks were excluded from analysis.

By design, mismatch only occurs during periods of running, and thus any running speed profile will always peak at the time of mismatch. To control for this, we generated a running speed profile and pupil responses triggered on randomly selected onsets during running in closed-loop sessions, that we then compared to actual pupil and running profiles to calculate response latencies. Similarly, to calculate the neuronal response latency to full-field and local perturbations, we generated neuronal response profiles triggered on randomly selected onsets during running and compared these to the actual perturbation response profiles. In all cases, we calculated the time of significant departure of the actual perturbation response profiles from response profiles triggered on random onsets ($p < 0.05$, paired Student's t test).

Bibliography

Ahmed, B., Anderson, J.C., Douglas, R.J., Martin, K.A., Nelson, J.C., 1994. Polyneuronal Innervation of Spiny Stellate Neurons in Cat Visual Cortex. J Comp Neurol 341, 16–24. doi:10.1002/cne.903410103

Barlow, H., 1953. Summation and inhibition in the frog's retina. J. Physiol. 69–88.

Bell, C.C., 1981. An efference copy which is modified by reafferent input. Science 214, 450–453. doi:10.1126/science.7291985

Binzegger, T., 2004. A Quantitative Map of the Circuit of Cat Primary Visual Cortex. J. Neurosci. 24, 8441–8453. doi:10.1523/JNEUROSCI.1400-04.2004

Bonin, V., Histed, M.H., Yurgenson, S., Reid, R.C., 2011. Local diversity and fine-scale organization of receptive fields in mouse visual cortex. J. Neurosci. 31, 18506–21. doi:10.1523/JNEUROSCI.2974-11.2011

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., Looger, L.L., Svoboda, K., Kim, D.S., 2013. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499, 295–300. doi:10.1038/nature12354

Cho, J., Sharp, P.E., 2001. Head direction, place, and movement correlates for cells in the rat retrosplenial cortex. Behav. Neurosci. 115, 3–25. doi:10.1037/0735-7044.115.1.3

Crick, F.C., Koch, C., 1998. Constraints on cortical and thalamic projections : the no-strong-loops hypothesis. Nature 391, 245–250. doi:10.1038/34584

Cudeiro, J., Sillito, A.M., 2006. Looking back: corticothalamic feedback and early visual processing. Trends

Neurosci. 29, 298–306. doi:10.1016/j.tins.2006.05.002

Dombeck, D.A., Khabbaz, A.N., Collman, F., Adelman, T.L., Tank, D.W., 2007. Imaging Large-Scale Neural Activity with Cellular Resolution in Awake, Mobile Mice. Neuron 56, 43–57. doi:10.1016/j.neuron.2007.08.003

Douglas, R.J., Martin, K.A.C., 2011. What's black and white about the grey matter? Neuroinformatics 9, 167–179. doi:10.1007/s12021-011-9106-1

Douglas, R.J., Martin, K.A.C., Whitteridge, D., 1989. Microcircuit for Neocortex Tholomus.

Eliades, S.J., Wang, X., 2008. Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature 453, 1102–6. doi:10.1038/nature06910

Eliades, S.J., Wang, X., 2005. Dynamics of auditory-vocal interaction in monkey auditory cortex. Cereb. Cortex 15, 1510–1523. doi:10.1093/cercor/bhi030

Eliades, S.J., Wang, X., 2003. Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. J. Neurophysiol. 89, 2194–2207. doi:10.1152/jn.00627.2002

Erisken, S., Vaiceliunaite, A., Jurjut, O., Fiorini, M., Katzner, S., Busse, L., 2014. Effects of locomotion extend throughout the mouse early visual system. Curr. Biol. 24, 2899–2907. doi:10.1016/j.cub.2014.10.045

Felleman, D.J., Van Essen, D.C., 1991. Distributed hierachical processing in the primate cerebral cortex. Cereb. Cortex 1, 1–47. doi:10.1093/cercor/1.1.1

Fiser, A., Mahringer, D., Oyibo, H.K., Petersen, A. V, Leinweber, M., Keller, G.B., 2016. Experience-dependent spatial expectations in mouse visual cortex. Nat. Neurosci. 19, 1658–1664. doi:10.1038/nn.4385

Friston, K., 2010. The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 11, 127–138. doi:10.1038/nrn2787

Friston, K., 2005. A theory of cortical responses. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 360, 815–36. doi:10.1098/rstb.2005.1622

Fu, Y., Tucciarone, J.M., Espinosa, J.S., Sheng, N., Darcy, D.P., Nicoll, R.A., Huang, Z.J., Stryker, M.P., 2014. A cortical circuit for gain control by behavioral state. Cell 156, 1139–1152. doi:10.1016/j.cell.2014.01.050

Garey, L.J., Powell, T.P., 1971. An experimental study of the termination of the lateral geniculo-cortical pathway in the cat and monkey. Proc. R. Soc. London. Ser. B, Contain. Pap. a Biol. character. R. Soc. 179, 41–63. doi:10.1098/rspb.1971.0080

Goodale, M.A., Milner, A.D., 1992. Separate Visual Pathways for Perception and Action. Essent. Sources Sci. Study Conscious. 175.

Hafting, T., Fyhn, M., Molden, S., Moser, M., Moser, E.I., 2005. Microstructure of a spatial map in the entorhinal cortex. Nature 436, 801–806. doi:10.1038/nature03721

Harris, K.D., Mrsic-Flogel, T.D., 2013. Cortical connectivity and sensory coding. Nature 503, 51–8. doi:10.1038/nature12654

Harris, K.D., Shepherd, G.M.G., 2015. The neocortical circuit: themes and variations. Nat. Neurosci. 18, 170–181. doi:10.1038/nn.3917

Hartline, H.K., 1937. The Response of Single Optic Nerve Fibers of the vertebrate eye to illumination of the retina 400–415.

Heinzle, J., Hepp, K., Martin, K.A.C., 2007. A microcicuit model of the frontal eye fields. J. Neurosci. 27, 9341–9353. doi:10.1523/JNEUROSCI.0974-07.2007

Howe, M.W., Dombeck, D.A., 2016. Rapid signalling in distinct dopaminergic axons during locomotion and reward. Nature 535, 505–510. doi:10.1038/nature18942

Huang, Y., Rao, R.P.N., 2011. Predictive coding. Wiley Interdiscip. Rev. Cogn. Sci. 2, 580–593. doi:10.1002/wcs.142

Hubel, D.H., Wiesel, T.N., 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. 160, 106–154.2. doi:10.1523/JNEUROSCI.1991-09.2009

Jordan, M.I., Rumelhart, D.E., 1992. Forward Models: Supervised Learning with a Distal Teacher. Cogn. Sci. 16, 307–354. doi:10.1207/s15516709cog1603_1

Keller, G.B., Bonhoeffer, T., Hübener, M., 2012. Sensorimotor Mismatch Signals in Primary Visual Cortex of the Behaving Mouse. Neuron 74, 809–815. doi:10.1016/j.neuron.2012.03.040

Keller, G.B., Hahnloser, R.H.R., 2009. Neural processing of auditory feedback during vocal practice in a songbird. Nature 457, 187–90. doi:10.1038/nature07467

Lee, A.M., Hoy, J.L., Bonci, A., Wilbrecht, L., Stryker, M.P., Niell, C.M., 2014. Identification of a Brainstem Circuit Regulating Visual Cortical State in Parallel with Locomotion. Neuron 83, 455–466. doi:10.1016/j.neuron.2014.06.031

Leinweber, M., Zmarz, P., Buchmann, P., Argast, P., Hübener, M., Bonhoeffer, T., Keller, G.B., 2014. Two-photon calcium imaging in mice navigating a virtual reality environment. J. Vis. Exp. e50885. doi:10.3791/50885

Leopold, D.A., Logothetis, N.K., 1998. Microsaccades differentially modulate neural activity in the striate and extrastriate visual cortex. Exp. Brain Res. 123, 341–345. doi:10.1007/s002210050577

Marshel, J.H., Garrett, M.E., Nauhaus, I., Callaway, E.M., 2011. Functional specialization of seven mouse visual cortical areas. Neuron 72, 1040–1054. doi:10.1016/j.neuron.2011.12.004

Martinez-Conde, S., Macknik, S.L., Hubel, D.H., 2000. Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys. Nat Neurosci 3, 251–258. doi:10.1038/72961

Miller, M.W., Vogt, B.A., 1984. Direct connections of rat visual cortex with sensory, motor, and association cortices. J. Comp. Neurol. 226, 184–202. doi:10.1002/cne.902260204

Niell, C.M., Stryker, M.P., 2010. Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. Neuron 65, 472–479. doi:10.1016/j.neuron.2010.01.033

Niell, C.M., Stryker, M.P., 2008. Highly selective receptive fields in mouse visual cortex. J Neurosci 28, 7520–7536. doi:10.1523/JNEUROSCI.0623-08.2008

O'Keefe, J., Dostrovsky, J., 1971. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res. 34, 171–175. doi:10.1016/0006-8993(71)90358-1

Olshausen, B.A., Field, D.J., 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature. doi:10.1038/381607a0

Pinto, L., Goard, M.J., Estandian, D., Xu, M., Kwan, A.C., Lee, S.-H., Harrison, T.C., Feng, G., Dan, Y., 2013. Fast modulation of visual perception by basal forebrain cholinergic neurons. Nat. Neurosci. 16, 1857–63. doi:10.1038/nn.3552

Polack, P.-O., Friedman, J., Golshani, P., 2013. Cellular mechanisms of brain state-dependent gain modulation in visual cortex. Nat. Neurosci. 16, 1331–9. doi:10.1038/nn.3464

Pollen, D., 1999. On the neural correlates of visual perception. Cereb. Cortex 4–19. doi:10.1093/cercor/9.1.4

Poulet, J.F.A., Hedwig, B., 2006. The cellular basis of a corollary discharge. Science 311, 518–22. doi:10.1126/science.1120847

Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., Fried, I., 2005. Invariant visual representation by single neurons in the human brain. Nature 435, 1102–7. doi:10.1038/nature03687

Rao, R.P.N., 1999. An optimal estimation approach to visual perception and learning. Vision Res. 39, 1963–1989. doi:10.1016/S0042-6989(98)00279-X

Rao, R.P.N., 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. doi:10.1038/4580

Riesenhuber, M., Poggio, T., 1999. Hierarchical models of object recognition in cortex. Nat. Neurosci. 2, 1019–25. doi:10.1038/14819

Rolfs, M., 2009. Microsaccades: Small steps on a long way. Vision Res. 49, 2415–2441. doi:10.1016/j.visres.2009.08.010

Roth, M.M., Dahmen, J.C., Muir, D.R., Imhof, F., Martini, F.J., Hofer, S.B., 2016. Thalamic nuclei convey diverse contextual information to layer 1 of visual cortex 19. doi:10.1038/nn.4197

Saleem, A.B., Ayaz, A., Jeffery, K.J., Harris, K.D., Carandini, M., 2013. Integration of visual motion and locomotion in mouse visual cortex. Nat. Neurosci. 16, 1864–9. doi:10.1038/nn.3567

Schneider, D.M., Nelson, A., Mooney, R., 2014. A synaptic and circuit basis for corollary discharge in the auditory cortex. Nature 513, 189–94. doi:10.1038/nature13724

Schuett, S., Bonhoeffer, T., Hübener, M., 2002. Mapping retinotopic structure in mouse visual cortex with optical imaging. J. Neurosci. 22, 6549–59. doi:20026635

Sherman, S.M., Guillery, R.W., 1998. On the actions that one nerve cell can have on another: distinguishing "drivers" from "modulators". Proc. Natl. Acad. Sci. U. S. A. 95, 7121–7126. doi:10.1073/pnas.95.12.7121

Sherrington, C.S., 1906. Observations on the scratch-reflex in the spinal dog. J. Physiol. 34, 1–50. doi:10.1113/jphysiol.1906.sp001139

Simoncelli, E.P., Heegert, D.J., Heeger, D.J., 1998. A Model of Neuronal Responses in Visual Area MT. Vision Res. 38, 743–761. doi:S0042698997001831 [pii]

Sperry, R.W., 1950. NEURAL BASIS OF THE SPONTANEOUS OPTOKINETIC RESPONSE PRODUCED BY VISUAL INVERSION. J. Comp. Physiol. Psychol. 43, 482–489.

Tsao, D.Y., Freiwald, W.A., Tootell, R.B.H., Livingstone, M.S., 2006. A Cortical Region Consisting Entirely of Face-Selective Cells 311, 670–674. doi:10.1126/science.1206034

Vezoli, J., 2004. Quantitative Analysis of Connectivity in the Visual Cortex: Extracting Function from Structure. Neurosci. 10, 476–482. doi:10.1177/1073858404268478

Vogt, B.A., Miller, M.W., 1983. Cortical connections between rat cingulate cortex and visual, motor, and postsubicular cortices. J. Comp. Neurol. 216, 192–210. doi:10.1002/cne.902160207

von Holst, E., 1954. Relations between the central Nervous System and the peripheral organs. Br. J. Anim. Behav. 2, 89–94. doi:10.1016/S0950-5601(54)80044-X

Wolpert, D., Ghahramani, Z., Jordan, M., 1995. An internal model for sensorimotor integration. Science (80-.). 269, 1880.

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J.P., Chang, W.-C., Jenvay, S., Miyamichi, K., Luo, L., Dan, Y., 2014. Long-range and local circuits for top-down modulation of visual cortex processing 345, 660–664.

Appendix

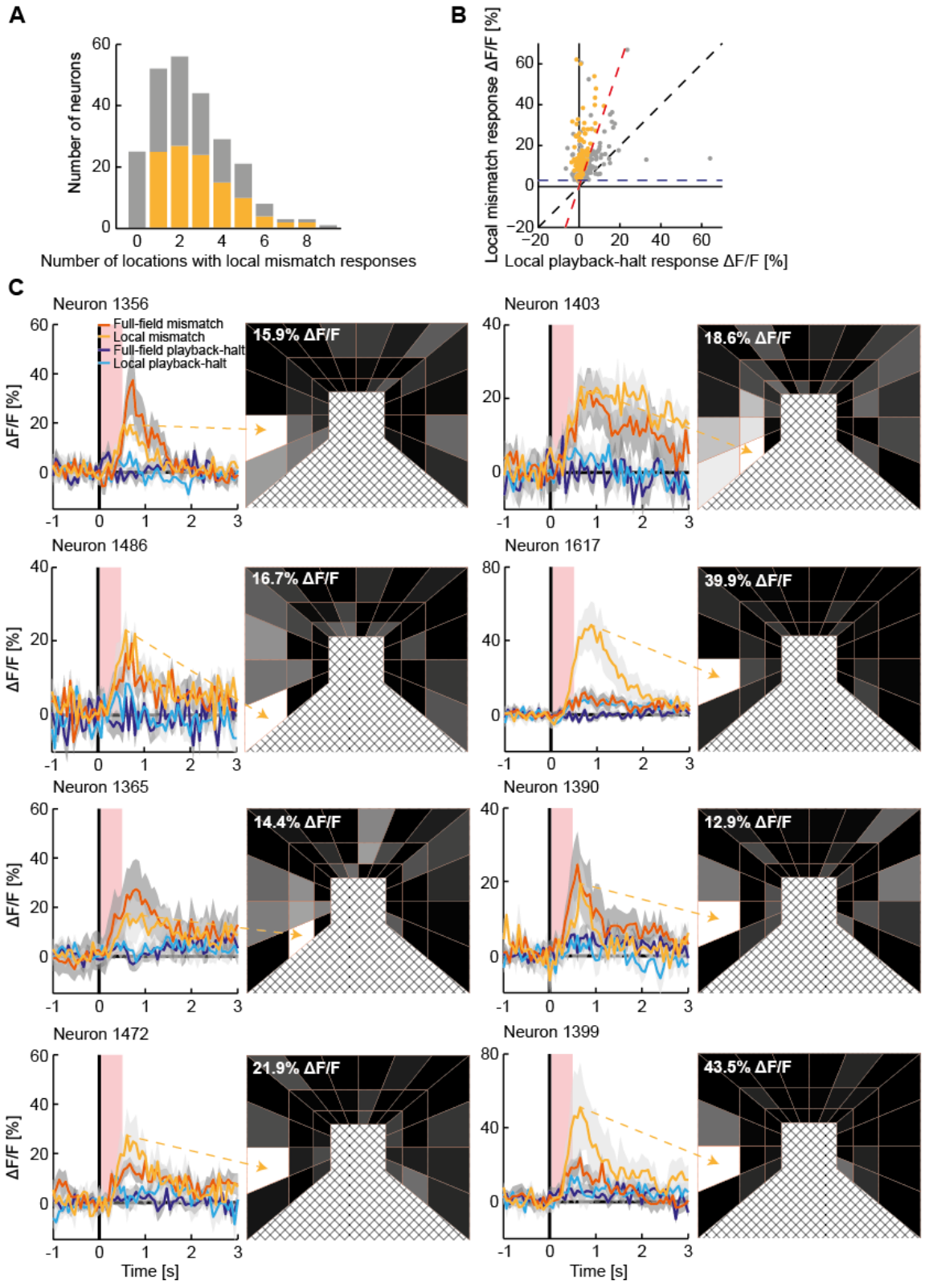
A
Number of neurons
Number of locations with local mismatch responses
B
Local mismatch response ΔF/F [%]
Local playback-halt response ΔF/F [%]
C
Neuron 1356
Full-field mismatch
Local mismatch
Full-field playback-halt
Local playback-halt
15.9% ΔF/F
Neuron 1403
18.6% ΔF/F
Neuron 1486
16.7% ΔF/F
Neuron 1617
39.9% ΔF/F
Neuron 1365
14.4% ΔF/F
Neuron 1390
12.9% ΔF/F
Neuron 1472
21.9% ΔF/F
Neuron 1399
43.5% ΔF/F
ΔF/F [%]
Time [s]

Figure A1. Mismatch neurons respond to local mismatch in confined regions of the visual field. (A) Number of locations eliciting significant local mismatch responses in all mismatch neurons (grey bars; n = 242) and mismatch neurons with selective local mismatch responses (light orange bars; n = 109; see (B) and Methods). (B) Criteria for selectivity of local mismatch neurons. Scatter plot of local playback-halt responses and local mismatch responses of mismatch neurons (n = 242). Dashed black line marks unity; horizontal dashed blue line marks 3% ΔF/F local mismatch response; dashed red line marks local mismatch responses that are three times as large as local playback-halt responses. Neurons above this line, with significant local mismatch responses (light orange dots) greater than 3% ΔF/F, were considered to have selective local mismatch responses. (C) Examples of mismatch neurons with selective local mismatch responses. Representation as in **Figure 5A** for eight different neurons from three different animals.

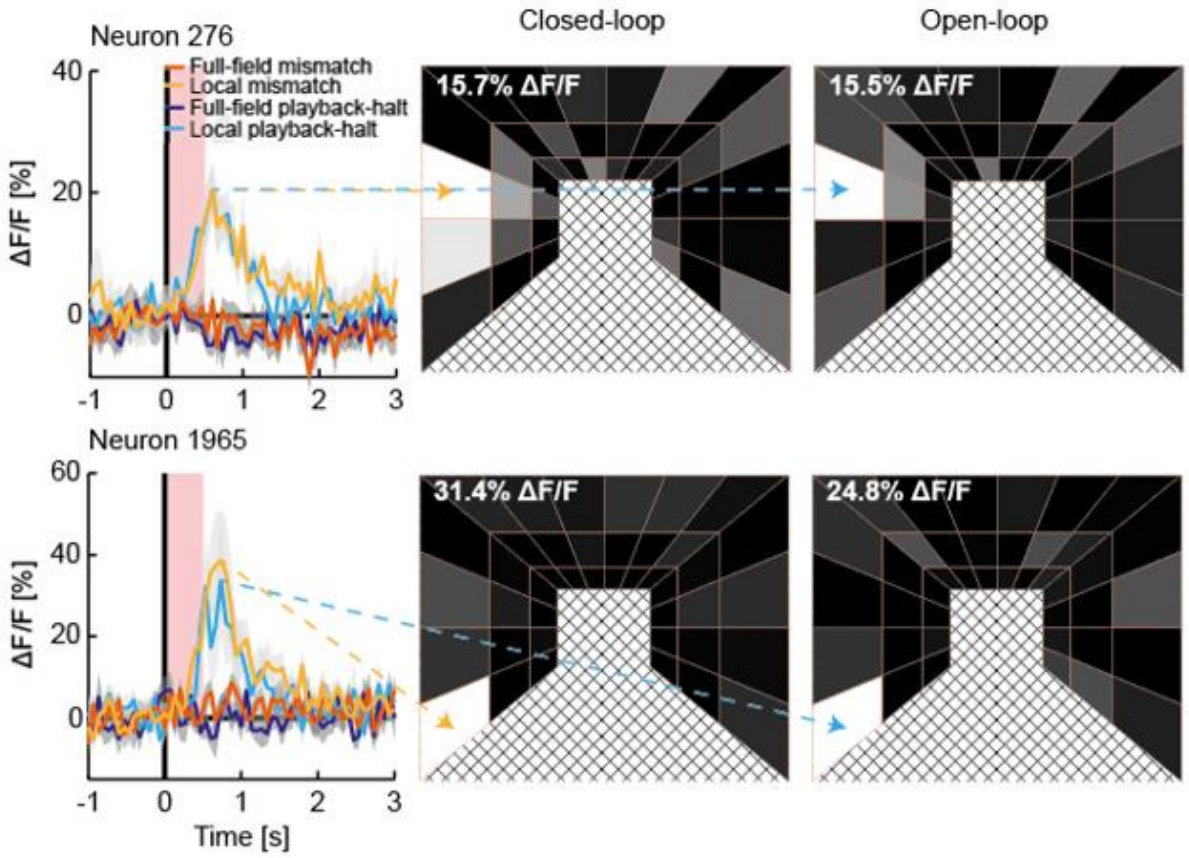

Figure A2. Examples of visual neurons with local responses. Response maps for each neuron show the spatial location of the peak local response during closed-loop (middle panel) and open-loop sessions (right panel). Representation as in **Figure 7A**

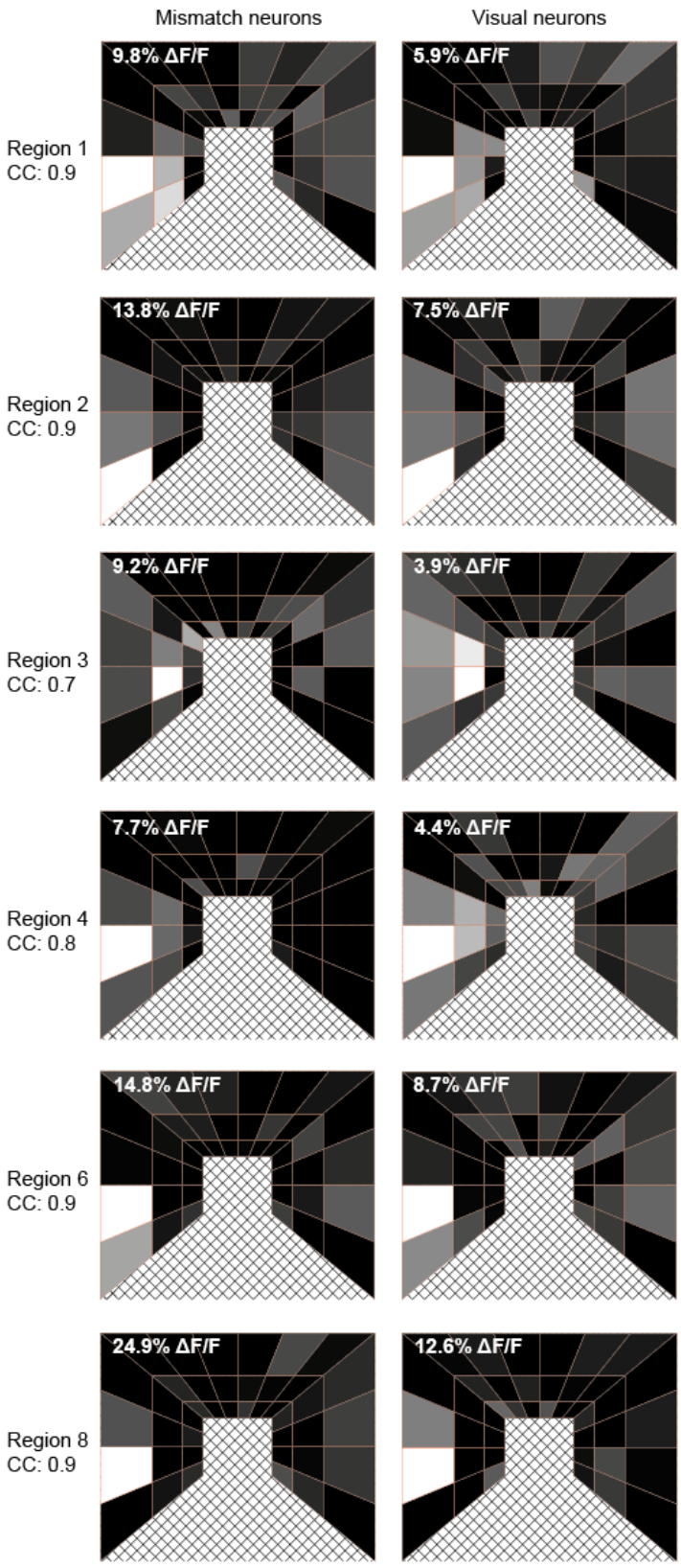
Mismatch neurons
Visual neurons
9.8% ΔF/F
5.9% ΔF/F
Region 1
CC: 0.9
13.8% ΔF/F
7.5% ΔF/F
Region 2
CC: 0.9
9.2% ΔF/F
3.9% ΔF/F
Region 3
CC: 0.7
7.7% ΔF/F
4.4% ΔF/F
Region 4
CC: 0.8
14.8% ΔF/F
8.7% ΔF/F
Region 6
CC: 0.9
24.9% ΔF/F
12.6% ΔF/F
Region 8
CC: 0.9

Figure A3. Spatial distribution of responses to local stimuli are highly similar between mismatch and visual neurons within the same imaging region. Correspondence of mismatch responses and visual responses by imaging region. Response maps of average local mismatch responses in mismatch neurons (left column) and average local perturbation responses in visual neurons (right column) for six example imaging regions. Peak responses are indicated in the upper left corner of each response map. Pearson's correlation coefficients (CC) between mismatch and visual neuron response maps for individual regions are indicated.

	Animal 1 (n = 384)	Animal 2 (n = 274)	Animal 3 (n = 204)	Animal 4 (n = 697)	Animal 5 (n = 620)
Mismatch neurons (n = 242)	73 (19.0%)	13 (4.7%)	86 (42.2%)	24 (3.4%)	46 (7.4%)
Mismatch neurons with significant local mismatch responses (n = 217)	71 (18.5%)	12 (4.4%)	75 (36.8%)	22 (3.2%)	37 (6.0%)
Mismatch neurons with selective local mismatch responses (n = 109)	36 (9.4%)	9 (3.3%)	36 (17.6%)	10 (1.4%)	18 (2.9%)
Visual neurons (n = 1433)	177 (46.0%)	128 (46.7%)	95 (46.6%)	543 (77.9%)	490 (79.0%)
Visual neurons with significant local responses (n = 1208)	166 (43.2%)	110 (40.1%)	87 (42.6%)	524 (75.2%)	321 (51.8%)

Table A1. Percentages of neuronal response types per animal. Indicated in the top row is the total number of recorded neurons per animal (n = 2179 across five animals). Of these, a subset was classified as mismatch neurons (n = 242). A subset of neurons in this group were further classified as mismatch neurons with significant local mismatch responses (n = 217), and mismatch neurons with selective local mismatch responses (n = 109). A separate subset of neurons was classified as visual neurons (n = 1433). A subset of neurons in this group were further classified as visual neurons with significant local responses (n = 1208). Indicated in the table are the number and percentage of neurons from each class out of the total number of recorded neurons per animal.