

The role of cortical layer six in the perception and laminar representation of sensory change

by

Jakob Voigts

Vordiplom, Universität Heidelberg (2008)

Submitted to the Department of Brain and Cognitive Sciences
in partial fulfillment of the requirements for the degree of

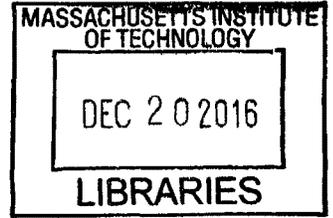
Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2016 [February 2017]

© Massachusetts Institute of Technology 2016. All rights reserved.



Signature redacted

Author

Department of Brain and Cognitive Sciences
September 2016

Signature redacted

Certified by

Christopher I. Moore
Professor of Neuroscience, Brown University
Thesis Supervisor

Signature redacted

Certified by

Emery N Brown, M.D., Ph.D.
Professor of Computational Neuroscience, Edward Hood Taplin Professor of Medical Engineering, and Warren M. Zapol Professor of Anesthesia
Thesis Supervisor

Signature redacted

Accepted by

Matthew A. Wilson, Ph.D.
Sherman Fairchild Professor of Neuroscience
Director of Graduate Education for Brain and Cognitive Sciences



77 Massachusetts Avenue
Cambridge, MA 02139
<http://libraries.mit.edu/ask>

DISCLAIMER NOTICE

Due to the condition of the original material, there are unavoidable flaws in this reproduction. We have made every effort possible to provide you with the best copy available.

Thank you.

The images contained in this document are of the best quality available.

The role of cortical layer six in the perception and laminar representation of sensory change

by
Jakob Voigts

Submitted to the Department of Brain and Cognitive Sciences
on September 2016, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Neocortex learns predictive models of sensory input, allowing mammals to anticipate future events. A fundamental component of this process is the comparison between expected and actual sensory input, and the layered architecture of neocortex is presumably central to this computation. In this thesis, I examine the role of laminar differences, and specifically the role of layer 6 (L6) in the encoding and perception of stimuli that deviate from previous patterns. In awake mice, layer 4 neurons encode current stimulus deviations with a predominantly monotonic, faithful encoding, while neurons in layer 2/3 encode history dependent change signals with heterogeneous receptive fields. Corticothalamic (CT) cells in Layer 6 respond sparsely, but faithfully encode stimulus identity. Weak optogenetic drive of L6 CT cells disrupted this encoding in layer 6 without affecting overall firing rates. This manipulation also caused layer 2/3 to represent only current stimuli. In a head-fixed stimulus detection task, small stimulus deviations typically make stimuli more detectable, and the L6 manipulation removed this effect, without affecting detection of non-changing stimuli. Analogously, in free sensory decision making behavior, the manipulation selectively impaired perception of deviant stimuli, without affecting basic performance. In contrast, stronger L6 drive reduced sensory gain and impaired tactile sensitivity. These results show an explicit laminar encoding of stimulus changes, and that L6 can play a role in the perception of sensory changes by modulating responses depending on previous, or expected input. This finding provides a new perspective on how the layered cortical architecture can implement computations on hierarchical models of the world.

Thesis Supervisor: Christopher I. Moore
Title: Professor of Neuroscience, Brown University

Thesis Supervisor: Emery N Brown, M.D,Ph.D.
Title: Professor of Computational Neuroscience, Edward Hood Taplin Professor of Medical Engineering, and Warren M. Zapol Professor of Anesthesia

Acknowledgments

I am deeply grateful to the many people who have been part of the fun journey of which the work described in this thesis was a part. I owe thanks to my advisor Chris Moore, for his infectious curiosity, for never being able to say no to a new idea for too long, and also for always pushing me to look for the next questions and the deeper answers. I also owe thanks to my other committee members, Nancy Kopell, Sebastian Seung and Matt Wilson, and my co-advisor Emery Brown for their extremely helpful insights.

I also owe special thanks to Emery Brown, Francisco Flores, Mike Halassa, Lukas Schmitt and the Brown and Wilson labs for the help and support on the project on the role of TRN in local sleep (joint work with Laura Lewis, not included in this thesis).

Chris Deister for being an amazing labmate, teaching me so many things. To Tyler Brown, and Dominique, Ulf, Katie, Nathan, Jason, and all of Moore lab for all the enlightening discussions and for making the lab a fun place to be in. Tim Buschman for the advice, and for making the state machine code used in chapter 4.2. Kate, Cassie, and Kia for keeping the lab running and dealing with the general chaos that we generated.

Scott and Shane for many helpful late-night discussions of the finer points of thalamo-cortical architecture.

Josh Siegle for being a great labmate, and for helping me get started with psychophysics experiments. More importantly, Open Ephys wouldn't have happened without him - our work together could not have been more effortless and productive.

Jon Newman for his work on open ephys, especially for making (and donating) the LED driver that made chapter 3.4 possible.

All others who have contributed to Open Ephys by supporting us and contributing designs and code, Caleb Kemere, Chris and Matt for their help as board members, D. Meletis, M. Carlen, J. Goldberg, M. Jazayeri, O. Yizhar, D. Huber, and D. Moorman, and Reid Harrison for supporting us early on, and Aaron for all of his incredible work.

All the mice without whom none of this work would have been possible.

David Herman, for pulling off some very hard experiments for our gap-crossing study, for being an overall amazing labmate, and perhaps most notable for making *even* LA a fun place. Tansu Celikel for looking for students at the Heidelberg math department and getting me interested in neuroscience and in doing hands-on experiments, and Winfried Denk for the support and for encouraging me to go to grad school for neuroscience.

All friends that have made the time at MIT fun, Chris, Rodrigo, Sangyu, McCoy, Kean, Meg, Danielle, and all my BCS 2008 classmates.

My brother Simon, who I had way too little time to spend with over the last 7 years. My Parents, Liane und Manfred, danke für all die Unterstuetzung und Ermutigung.

Finally, my amazing and brilliant wife Laura without who none of this would have happened.

Contents

1	Introduction	7
1.1	Cortical computation and stimulus change encoding	7
1.1.1	Stimulus change encoding	7
1.1.2	Predictive coding	9
1.1.3	Cortical hierarchies	10
1.1.4	Error signals and receptive fields	12
1.2	Role of layer 6	13
1.3	Overview of the thesis	16
2	Neocortical representation of stimulus changes	29
2.1	Chronic high channel count microdrive recording in mice	29
2.1.1	Introduction	30
2.1.2	Materials and methods for drive implant construction	32
2.1.3	Drive mechanism	33
2.1.4	Optical fibers	35
2.1.5	Results	36
2.1.6	Discussion	36
2.2	Methods	40
2.2.1	Animal subjects	40
2.2.2	Viral injections	40
2.2.3	Surgical procedures	41
2.2.4	Experimental design for electrophysiology and 2-photon imaging	41
2.2.5	Analysis of electrophysiology data	42
2.3	Laminar specific encoding of stimulus changes	44
2.3.1	Layer 4 neurons encode stimulus deviations faithfully	47
2.3.2	Layer 2/3 neurons encode stimulus deviations heterogeneously	48
2.3.3	Ideal observer decoding of L2/3 and L4 responses	50
2.4	Stimulus encoding in Layer 6	51
2.4.1	2-Photon Imaging setup	52
2.4.2	2-Photon Imaging in deep tissue	52
2.4.3	Layer 6 CT neurons encode stimulus amplitude	53
3	Optogenetic manipulation of stimulus representation in L6	77
3.1	Methods	77
3.1.1	Optogenetic stimulation	77
3.2	Gain modulation by L6 CT cells	78
3.2.1	Simultaneous 2-Photon Imaging and Optogenetic stimulation	79

3.2.2	2-Photon data analysis	80
3.3	Weak drive of L6 CT cells	80
3.4	Effect on stimulus encoding in L6	84
3.5	Effect on stimulus encoding in other layers	85
3.6	Role of fast spiking interneurons	89
4	Role of L6 in the perception of stimulus changes	97
4.1	Saliency of deviant stimuli	97
4.2	Head-fixed stimulus detection	98
4.2.1	Animal subjects	98
4.2.2	Behavioral training and stimulus design	98
4.2.3	Head-fixed stimulus detection	101
4.2.4	Head-fixed behavioral analysis	101
4.2.5	Effect of weak L6 CT drive on change induced stimulus saliency	101
4.3	Untrained free behaviour	104
4.3.1	Gap-crossing	104
4.3.2	Gap-crossing analysis	105
4.3.3	Unsupervised approximate whisker tracking	107
4.3.4	Assessing stimulus novelty perception in Gap crossing	109
4.3.5	Effect of weak L6 CT drive on change preception in gap-crossing	109
4.3.6	Effect of sensory gain modulation on change perception	110
5	Conclusions	117
5.1	Summary	117
5.2	What do these findings teach us about the role of L6	118
5.3	Implications for models of canonical cortical computation	122
5.4	Rodent S1 specific computations	126
5.5	Conclusions and future directions	128
	Appendices	138
A	Open source tools for neuroscience	139

Chapter 1

Introduction

1.1 Cortical computation and stimulus change encoding

A fundamental aspect of cognitive function is the ability of the brain to build models of the sensory environment (von Helmholtz, 1867), which entails comparing expected to actual sensory input (Mumford, 1992). In this thesis I am examining the role of a specific neocortical layer and cell type in this computation. The neocortex, the evolutionarily most recent addition to the mammalian nervous system (Lui et al., 2011), is the principal substrate for higher sensory function, generation of complex motor sequences, cognitive functions and language. Neocortex is characterized by a repetition of stereotyped modules of a six-layered architecture (Mountcastle, 1957), hinting at the existence of generalized mechanisms of cortical function that could apply across brain areas serving different functions.

1.1.1 Stimulus change encoding

From an evolutionary point of view, it is valuable to anticipate the immediate future and to dedicate resources to unexpected stimuli. For almost all animals, unexpected sensory inputs occur when the animal's internal model of its environment, and with that its possible motor plans, are not matched to reality, which can result in anything from non-adaptive behaviours to existential risks to life.

Neocortex represents stimuli that deviate from an internal model of the expected stimulus differently from expected or repeated stimuli (Chater et al., 2006; Rao and Ballard, 1999). Signatures of this computation are seen in the characteristic responses to deviations from patterned stimuli in visual (Courchesne et al., 1975), auditory (Tiitinen et al., 1994; Ulanovsky et al., 2003), and language processing (Kutas and Hillyard, 1980; Pulvermüller and Shtyrov, 2003) areas, underlining the importance of change detection as a fundamental neocortical operation.

To study this process in neocortex, and to examine its possible contributions to hierarchical models and predictive coding in neocortex, a careful consideration of the contributions of neocortex to basic change encoding is required. Specifically, it is useful to distinguish between *bottom-up* change encoding relative to preceding states that are represented in the local circuit, or may even be inherited from preceding levels of sensory processing, and *top-down* change encoding that compares a local input to an expectation that is derived from a computation that was carried out in a distal brain area that encodes higher order stimulus

features. This distinction likely maps to different implementation mechanisms: stimulus changes that are computed in a bottom-up process can be detected by a local brain area, with a variety of mechanisms. Stimulus change encoding that relies on top-down input must, by definition, rely on input from distant cortical (or thalamic) regions.

Change detection is typically studied as the increase in firing rates elicited by deviant stimuli following stimulus-specific adaptation 'SSA' (Movshon and Lennie, 1979). In this process, stimulus-tuned neurons along the afferent pathway and at thalamocortical synapses adapt to repeated stimulation (Chung et al., 2002; Katz et al., 2006), and subsequent deviant stimuli activate new pools of less adapted neurons at higher rates, leading to increased neocortical drive. SSA has been observed for simple stimuli in rodents (Khatri and Simons, 2007), as well as in areas ranging from low level auditory (Inferior Colliculus) (Ayala and Malmierca, 2013), to high order visual feature encoding areas (encoding of face features in anterior medial temporal neocortex) (Furl et al., 2007) in humans, and for multiple time-scales (Ulanovsky et al., 2004) showing that adaptation to specific stimulus features is a generalized property of neocortex.

In addition to SSA that occurs in individual brain regions, the transfer of adaptation to optical flow, leading to motion aftereffects (Addams, 1834; Purkinje, 1820; Barlow and Hill, 1963) between tactile and visual modalities (Konkle et al., 2009) underlines that adaptation occurs at the level of a multi-tiered internal model of the sensory input model (cross-modal in the case of motion aftereffects, or across many levels of visual processing in the case of facial features). These findings clearly indicate the necessity of disambiguating the contributions of bottom-up and local adaptation from possible top-down effects related to the encoding of hierarchical models in the brain (Grill-Spector et al., 2006).

Adaptation to specific stimuli can serve a role not only in compensating for persistent states of the external world, but also for dynamics of neural systems. Adaptation at fast (Tsodyks and Markram, 1997) and slow (Dragoi et al., 2000) timescales has also been suggested as an internal compensation mechanism for maintenance of computational properties of circuits in the presence of unavoidable fluctuations in neural excitability (Stevenson et al., 2010), and to optimize the encoding of rapid stimulus variations (Fairhall et al., 2001). In this context, the computational role of the encoding of stimulus changes in neocortex needs to be carefully considered in how it performs a specific, interpretable computational role in stimulus processing versus its role in maintaining computational properties. This ambiguity between a computationally relevant process and its implementation (Marr and Poggio, 1976) is especially problematic in reduced experiments that rely on a narrow range of stimuli, and abstract away from the context that any given neural system usually operates in (Gomez-Marin et al., 2014).

In sum, there are many parallel 'bottom-up' mechanisms, both pre-cortical, and cortical, that contribute to the detection and encoding of stimulus changes. Therefore, in order to specifically investigate mechanisms that contribute to the encoding of deviations relative to 'top-down' expectations that are set across different levels of representations, and across different cortical areas, we need to employ a stimulus design that maximizes the involvement of cortical mechanisms, while minimizing the contribution of pre-cortical SSA, and compare change encoding across neural populations that receive predominantly direct thalamic input, and ones that receive, or are modulated by, cortico-cortical input.

1.1.2 Predictive coding

In addition to adaptation-mediated bottom-up mechanisms that lend increased saliency to deviant stimuli, and change the encoding and consequently interpretation of sensory input, there are a large class of cortically represented, ‘top-down’ factors such as stimulus context, history, or expectation (Knierim and van Essen, 1992; Chelazzi et al., 1993; Reynolds et al., 2000; Maunsell and Treue, 2006) that significantly influence stimulus responses, indicating the existence of cortical mechanisms that compare expectations to sensory data. This view is deeply linked with the framework of theories that posit that neocortex builds hierarchical models in which top-down and bottom-up flows of information are reconciled to build predictive models of the world.

More generally, the ability to anticipate future events, which in a complex environment entails building predictive models that match the complexity of this environment, and can enable the animal to infer causes of sensory inputs (Friston and Stephan, 2007), has general utility. Internal models of the world allow for more complex behaviours than is possible with slowly adapting stimulus-response mappings, and enable computationally complex behaviours such as language and abstract reasoning. It has long been proposed that the brain solves this problem by constructing internal, generative models that attempt to anticipate and match sensory inputs to predictions (von Helmholtz, 1867; Barlow, 1961; Gregory, 1968; Dayan et al., 1995; Hawkins and Blakeslee, 2005).

In the simplest case, anticipatory processing of sensory inputs can compensate for delays in sensory and motor processing, and even relatively simple cases of predictive coding have clear utility. If a complex sensorimotor contingency is constrained by its delay, executing a motor command that is not targeted to the current state of the world, but instead to an extrapolated estimate of the state of the world at the time that the motor command is executed is needed. The idea that actions are driven not by the current state of the world as it is available to the sensory system, but by an anticipation of the future state of the body and world dates back to as early as the 19th century (James, 1890; Prinz, 1990; Hommel et al., 2001). One well documented example of a psychophysical effect that is consistent with anticipatory encoding is the flash-lag effect (Nijhawan, 2002; Cavanagh, 1997), where a briefly visible static stimulus that is aligned to a continuously visible moving object is perceived as displaced from the moving one. This finding suggests that the visual system continuously extrapolates the position of moving objects to compensate for delays in visual processing. Similar extrapolation processing has been proposed for all systems that handle sensorimotor relationships (Bridgeman, 2007; Grüsser, 1995), and has been proposed as a general mechanism for action planning (Bubic et al., 2010).

Regardless of the utility of the ability of a neural system to learn to predict future input, doing so could also be a viable general strategy for any system trying to learn properties of a time-varying input. The errors made in predicting future states of a time-varying signal provide a readily available and informative training signal for learning generative models of the signal (Kulkarni et al., 2015). This line of reasoning suggests that regardless of whether any given brain region is computationally involved in explicit prediction, it might still be expected to compute an ongoing prediction error.

1.1.3 Cortical hierarchies

The most biologically plausible implementation of such generative models is via hierarchical *predictive coding*, where cortical areas aim to anticipate incoming sensory data, building more invariant representations in the process, while lower areas reconcile their input with the prediction made by the higher areas, and provide information about the mismatch between this prediction and the actual data to the higher areas to mark unexpected input and improve the model (Mumford, 1992). The view that these internal models of the world arise naturally from predictive coding is particularly attractive because it provides a biologically plausible link between the computation and its implementation in the neocortical architecture (Friston and Kiebel, 2009; Rao and Ballard, 1999; Bastos et al., 2012). By examining the role of the layered cortical architecture in the encoding of sensory change (specifically, an aspect of this architecture that lends itself well to integration of descending corticocortical with ascending sensory information), this thesis aims to advance our understanding of the biological substrate that implements hierarchical cortical models.

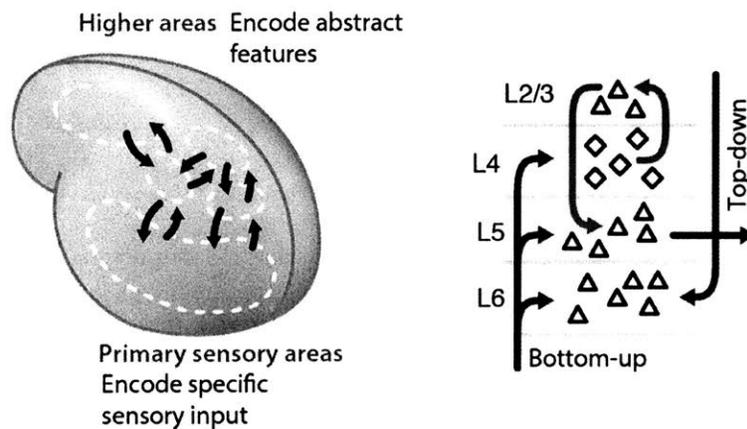


Figure 1-1: **The flow of information across cortical areas involves the layered cortical architecture.** **Left**, Neocortex represents sensory information, and its interpretation in multiple regions, in different levels of abstraction. Information is communicated between, and within these areas, and this flow of information can be broadly categorized as bottom-up, from ‘lower’, posterior, sensory areas, to ‘higher’ anterior areas encoding more abstract, invariant information, and top-down in the opposite direction, as well as ‘lateral’ for information flow across regions at the same level. **Right**, The layered cortical circuit has distinct input and output connectivity across layers. Studying the integration of different inputs in this architecture is therefore crucial to the understanding of neocortical function.

The view of the brain as computing hierarchical models of the world can be thought of in the context of Bayesian probabilistic inference when top-down or local expectations for input into any given brain area are interpreted as a prior that shapes the interpretation of new data, and is updated based on the data (von Helmholtz, 1867; Mach, 1897; Knill and Pouget, 2004; Tenenbaum et al., 2006; Griffiths et al., 2008; Tenenbaum et al., 2011). This computational view allows the study of the computation that is carried out by animals,

regardless of the implementation (Marr and Poggio, 1976), and has led to a wealth of observations showing that animals behave optimally from a Bayesian point of view in a variety of contexts (Freeman and Pasztor, 1999; Weiss et al., 2002; van Beers et al., 2001; Knill, 1998; Crowell and Banks, 1996; Wolpert et al., 1995). Bayesian framework theories also make testable predictions about the encoding of sensory information at the circuit level. Bayesian contextual effects on the perception of stimuli, even in simple stimulus detection tasks (Zhaoping and Jingling, 2008), make specific predictions about how stimulus expectations and uncertainty can be encoded in neural circuits (Rao, 2004; Ma et al., 2006). These findings underline the importance of integration of top-down information flow in the local cortical circuit as a fundamental aspect of brain function.

There are many lines of evidence that suggest neocortex builds hierarchical models of the world. In the visual system, sensory information is processed along streams of distinct, initially retinotopically organized, but increasingly high-order cortical areas (Essen and Maunsell, 1983; Felleman and Van Essen, 1991), that encode increasingly selective, and invariant representations of the visual scene (Riesenhuber and Poggio, 1999) that relate to models of how visual scenes could be broken down into generative primitives in multiple levels of abstraction (Marr and Poggio, 1976; Marr, 1982).

One central aspect of top-down processing that highlights the importance of top-down control of cortical information processing is the selective allocation of resources to behaviourally relevant stimuli. In visual scenes, for example, multiple objects compete for the allocation of neuronal resources, and this process is biased by complex conjunction of stimulus features and context (Desimone and Duncan, 1995). In many behavioural contexts, only selected aspects of a sensory input are relevant to further processing, and can affect behaviour (Duncan, 1984; Kosslyn and Andersen, 1995). In almost all models of hierarchical cortical models, information from sensory areas is selectively routed to higher areas. This process can be thought of as a serial allocation of ‘selective attention’ to relevant stimulus features by a top-down input (Graboi and Lisman, 2003), that could be implemented by a variety of cortical mechanisms, ranging from selective modulation of sensory evoked firing rates (Maunsell and Cook, 2002; Maunsell and Treue, 2006) to temporal correlation of groups of neurons (Singer et al., 1995; Engel et al., 1999; Siegle et al., 2014), or even implemented at the level of thalamic gating of information flow to, and between, cortical areas (Crick, 1984; McAlonan et al., 2008; Lewis et al., 2015).

One example of selective bottom-up flow of information is the ‘change blindness’ effect of selective attention in the visual system, where even salient stimuli are not detected by subjects if embedded in a complex scene that otherwise engages attention (James, 1890; Simons and Levin, 1997; O’Regan et al., 1999). Conversely, top-down facilitation of object detection (Bar, 2003) by modulation of lower visual areas has been linked to high order areas including orbitofrontal cortex (Bar et al., 2006), and the finding that saliency of visual objects is computationally determined by top-down input (Oliva et al., 2003) highlight that the predictions made by models at higher levels of the hierarchy lead to behaviourally relevant changes in how information is interpreted at low levels of processing.

In addition to the context-dependent modulation of information flow to higher brain areas, top-down processes also affect what information is encoded in lower sensory areas. A particularly clear example of the local neuronal effects of information that is encoded

non-locally is the modulation of responses of visually responsive neurons by inputs outside their first-order receptive fields (Bolz and Gilbert, 1986; Grieve and Sillito, 1991; Cavanaugh et al., 2002). This type of modulation can be broken down into ‘lateral’ inputs that originate from information that is encoded at the same level of the cortical hierarchy, such as retinotopically distant columns of the same area (Angelucci and Bressloff, 2006; Bolz and Gilbert, 1986), or ‘top-down’ input originating from brain areas thought to encode more invariant, abstract features (Gómez-Laberge et al., 2016; Bar, 2003; Graboi and Lisman, 2003; Zhang et al., 2014). This modulation can take the form of modulation of the gain of sensory responses (Carandini and Heeger, 2012), or modify their local receptive field properties (Crook et al., 1998; Martínez-Trujillo and Treue, 2004), or variability (Gómez-Laberge et al., 2016) to increase their information content in a task or context specific manner (Hupé et al., 1998). In the temporal domain, unpredicted stimuli elicit higher firing rates than predicted ones, both for simple cross-modal cases (Sutton et al., 1965), and for stimuli that are unpredictable in terms of their motion pattern (Alink et al., 2010). These findings provide links between the framework theories and findings on hierarchical, and predictive neocortical coding, and local receptive field properties of individual neurons. Elucidating a potential mechanism behind this influence of non-locally encoded information on local stimulus encoding was a key motivation for the work described in this thesis.

In addition to providing a way to think about general brain function, hierarchical models provide a structured way to think about brain disorders that are characterized by relatively minor changes in local circuit function but affect behaviour, such as schizophrenia (Bassett et al., 2008; Vogels and Abbott, 2007; Fletcher and Frith, 2009; Dima et al., 2009; Cook et al., 2012), anxiety disorders (Browning et al., 2015), and autism spectrum disorders (Turi et al., 2016; Manning et al., 2015). These findings suggest that a better understanding of how intra-cortical relay implements predictive models could inform a better understanding of how subtle, locally defined, and neurophysiologically tractable at the level of individual neurons or small circuits, changes in local circuit function can lead to complex psychiatric phenotypes (Huys et al., 2016).

In summary, the learning of hierarchical models of sensory input by neural systems, and learning to predict future sensory inputs are deeply related aspects of the same process. Neural systems, especially in mammals, have evolved to represent and operate on highly complex models of their environment, and understanding the basic mechanisms by which this process is implemented is one of the central questions of neuroscience.

1.1.4 Error signals and receptive fields

Mismatch signals are important from a behavioural point of view, as well as for almost all theories of how deep cortical hierarchies function. Even at a local level of the receptive fields (RFs) of individual neurons, there are therefore reasons to believe that neocortex should implement an explicit mechanism for comparing predicted to actual input.

Olshausen and Field (1996) observed that training a simple encoding of natural images as linear combination of basis functions under the additional constraint of (L_1) sparsity of the coefficients can result in basis functions that resemble oriented, spatially localized bandpass filters, strikingly similar to the gabor-filter like RFs of cells in mammalian primary visual cortex.

Similarly, enforcing sparsity in simple neural networks trained to reproduce image segments (this type of architecture is termed auto-encoder (Ballard, 1987; Pascal Vincent, 2010)) learn gabor-like filters when sparsity is enforced by choosing a hidden layer of significantly fewer neurons than are in the input and output layers. In relatively simple models of the visual system, introduction of a top-down expectation that enables computation of prediction errors in lower areas results in neurons with similar classes of extra-classical receptive-fields (end-stopping) as observed in V1 (Rao and Ballard, 1999).

More recently, similar unsupervised methods have been shown to be useful for unsupervised ‘pre-training’ of neural networks, which allows fast training of extremely deep networks (Hinton et al., 2006). This approach has become extremely influential (LeCun et al., 2015) and was one of the starting points (Krizhevsky et al., 2012) for the recent re-emergence of neural networks in machine learning.

These results showed that some of the non-trivial properties of sensory encoding that we observe in the nervous system may arise from comparatively simple principles, and provide a framework for studying the basic properties and computational requirements that neural systems need to implement in order to be able to perform complex computations. One of these requirements is the existence of a local error, or mismatch signal: To learn receptive fields that result in efficient stimulus encoding, and to achieve symmetry breaking between neurons that could potentially represent the same information, there has to be a representation of each neuron’s contribution to an error signal computed on more than just this neuron’s output. If this basic property was not fulfilled by neocortex, neurons would not be able to form receptive fields that, as an ensemble, result in sparse and precise encoding of complex stimuli. Even though a wide variety of mechanisms of lateral inhibition and competitive mechanisms between cells encoding different stimuli have been observed in almost all sensory systems (von Bekešy, 1967; Fiorentini, 1972), similar mechanisms are expected to exist for the representation of prediction errors, rather than for external stimulus properties.

1.2 Role of layer 6

In this thesis, I am investigating a fundamental mechanism by which top down input, or local history, could be integrated with bottom up, or sensory input. This process has been considered at many possible levels of the cortical circuit, and it is still not known how processes at these different levels interact and work together in implementing cortical computations.

Studying the neural implementation of computations performed by neocortex in the context of this laminar organization provides a starting point for the discovery of general mechanisms that should be applicable across diverse brain areas (Mountcastle, 1998; Hawkins and Blakeslee, 2005; Carandini and Heeger, 2012; Miller, 2016). Here, I consider the contribution of inter-laminar differences, especially the contribution of modulation by layer 6 in this computation, but it is useful to consider the broad range of neuronal mechanisms by which change-detection could be integrated.

At the most local level mismatch between different afferent information streams could be integrated in individual cells via coincidence detection in local dendritic segments (Stuart and Häusser, 2001), or via gating between inputs to basal and apical dendritic regions (Larkum et al., 1999). At the level of neuronal populations, this process could be implemented by competing ensembles (Fukai and Tanaka, 1997; Maass, 2000), or shifts in attractor dynamics (Miller, 2000; Mante et al., 2013; Stokes et al., 2013), or temporal firing patterns (Uhlhaas et al., 2009) across neural populations. Regardless of these mechanisms, information integration across brain areas is expected to exhibit a laminar component, owed to the laminar segregation of inputs and outputs of cortical circuits.

Here I examine the integration of descending input with previous or descending information at a more global level, that of the laminar cortical circuitry. The mammalian neocortex is made up of a stereotyped architecture that is shared across almost all cortical areas, regardless of their function (Mountcastle, 1957; Hubel and Wiesel, 1974; Rockel et al., 1980), in which a majority of neuronal connections are made among neighboring (<1mm) neurons (Jiang et al., 2015). A central aspect of this architecture is the presence of a laminar architecture, typically divided into six major layers, L1-6 (Harris and Shepherd, 2015; Douglas et al., 1989). The canonical view (Felleman and Van Essen, 1991; Callaway, 1998) of this architecture is that ‘bottom-up’ input arrives predominantly at L4 (though there is also significant thalamo-cortical drive of deeper layers (Douglas et al., 1989; Constantinople and Bruno, 2013)), is then relayed to cells in L2/3, that then projects to other cortical areas, as well as to local L5 cells that provide the main long-range output from a cortical circuit. Layer 1 contains mostly inhibitory neurons as well as dense axonal tracts and dendrites of cells in deeper layers.

Neocortical layer 6 (L6) is well positioned to contribute to this process, as it integrates lemniscal thalamic, long-range cortico-cortical, and modulatory inputs (Martin and Whitteridge, 1984; Zhang and Deschênes, 1998; Thomson, 2010; Zhang et al., 2014). Layer 6 is comprised of a variety of neuron types (Ferrer et al., 1986b,a; Wiser and Callaway, 1996, 1997). The main class of pyramidal neurons examined in this study are corticothalamic L6 neurons (CT) which project both to thalamus (Briggs and Usrey, 2008), as well as to superficial layers (Bortone et al., 2014). In primary visual neocortex, neurons are sparsely sensory driven with highly selective receptive fields (Vélez-Fort et al., 2014), are morphologically well positioned to integrate information with dendritic nonlinear gating mechanisms (Marx and Feldmeyer, 2013) and can modulate sensory gain through an intracortical pathway (Olsen et al., 2012; Bortone et al., 2014), suggesting that L6 could regulate sensory responses depending on specific stimulus content, context and history. This prediction is supported by L6 mediated modulation of visual receptive fields by stimulus context (Bolz and Gilbert, 1986). Further, in humans, deep cortical layers are specifically activated in top-down sensory processing (Kok et al., 2016) of illusory figures that require figure completion. In secondary visual cortex, L6 is involved in object recognition memory (López-Aranda et al., 2009) suggesting a role in the relay of information between layers of the visual hierarchy.

However, whether L6 plays a role in the representation and perception of stimulus changes is not known. Specifically, little is known about the role of the highly specific stimulus encoding in these cells. If L6 encoded a stimulus expectation, as is suggested by these prior findings, then manipulations of this encoding should result in informative changes of

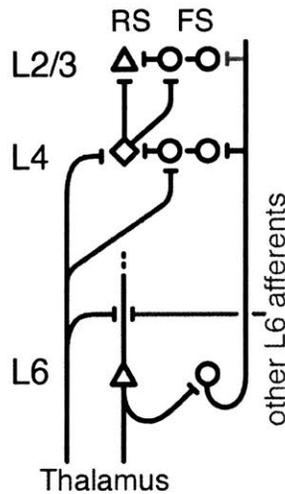


Figure 1-2: **Basic L6 circuit.** Sketch of the basic circuit surrounding L6 (cortico-thalamic, ‘CT’) pyramidal neurons that is considered in this thesis. L6 cells receive input (sensory and other) from thalamic afferents (red) as well as from other long-range cortico-cortical projections (blue), local input from other cortical layers within the same and neighboring columns (not shown). L6 CT cells project back to thalamus, to local L6 inhibitory interneurons, as well as to other neurons in superficial layers.

both neuronal encoding in superficial layers, and of perception of stimuli that deviate from previous expectation, but have little effect on stimuli that can be equivalently processed without engaging any change-detection mechanisms. Here, we tested this hypothesis using extensive single-neuron recordings across neocortical layers, calcium imaging, behavior, and simultaneous optogenetic perturbations in awake mice.

Here, I study the role of the laminar organization of neocortex, specifically the contribution of specific stimulus representations in layer 6 in the outlined processes in a series of change-encoding (Ch.2,3) and perception (Ch.4) experiments using mouse barrel cortex (Welker, 1964; Vincent, 1912; Woolsey and Van der Loos, 1970). Barrel cortex is well suited to this study because of its well characterized computational properties (Ritt et al., 2008; Kleinfeld and Deschenes, 2011), network architecture (Feldmeyer, 2012) with easy access to genetic tools (Gong et al., 2007), and its involvement in easily trainable (Hutson and Masterton, 1986) and quantifiable (Voigts et al., 2008) sensory behaviour that requires integration of information over experimentally convenient timescales (Celikel and Sakmann, 2007; Voigts et al., 2015).

1.3 Overview of the thesis

The implementation of the basic computational principles that underlie cognitive processes has been a subject of decades of research. A lot of progress has been made in characterizing the computations that underlie complex behaviour, and equivalently, much is now known about the basic molecular components that make up neurons, and eventually neural circuits. However, there still remains a massive gap in our understanding of even the simplest neural computations at the level of neurons and neural circuits.

In this thesis, I attempt to shed light on a simple but possibly fundamental aspect of cortical processing, the adaptation of local circuits to specific inputs, and the subsequent explicit representation of deviations, at the level of the basic laminar neocortical circuit. I use the encoding and perception of simple and small stimulus changes as a tool to probe the implementation of this basic computation, specifically the role of stimulus encoding by cortical layer 6 (L6), in the context of sensory behaviours in mice. This approach does not conclusively test specific hypotheses about the wiring diagram of the neocortex, but provides insights into the differentiated roles of cortical layers, and specifically into the role of stimulus encoding in L6 in this computation.

The first part of this thesis focuses on the representation of stimulus changes in the barrel cortex of mice. In Chapter 2, I describe the advances in extracellular recording technology that I developed to record large populations of identified neurons across all layers of neocortex in awake behaving mice. Using this method, I describe a laminar difference between the encoding of stimulus changes in Layers 4 and 2/3, that recapitulates similar findings in visual cortex of higher species and provides an accessible model system for studying the interaction between cortical layers in the representation of stimulus changes.

I also describe the approach I took for imaging large populations of neurons in cortical L6. Using this method I found that even though L6 CT cells are sparsely active, they encode stimulus content and could therefore play a role in the observed stimulus change encoding in superficial layers.

In Chapter 3, I describe the finding that weak optogenetic drive of L6 CT cells results in a specific re-weighting of neural activity in L6, where a subset of cells are driven to be more active, while others are suppressed, most likely due to the recruitment of FS interneurons in L6. This manipulation therefore does not change overall firing rates in L6, but reduces the stimulus information encoded by these cells. Further, firing rates in other layers are largely not affected. In L2/3, the same manipulation selectively removes the encoding of stimulus deviations, making these cells instead represent the current stimuli with a faithful encoding that more closely mirrors their thalamic input.

In Chapter 4, I describe two independent behavioural experiments that show that the manipulation of L6 encoding results in specific perceptual and behavioural deficits: In a head-fixed threshold stimulus detection task, small stimulus deviations typically make stimuli more detectable. The weak L6 manipulation removed this benefit, without affecting detection of non-changing stimuli. Analogously, in a task that tests free sensory decision making behavior, the manipulation selectively impaired perception of small, sudden stimulus changes, without affecting basic sensory performance. I also show that induction of

sensory gain reduction with stronger L6 drive impaired tactile sensitivity.

In Chapter 5 I discuss the results in the context of other work, and their implications for models of cortical circuits. I also outline some directions for future studies that build on questions raised by this thesis.

In the Appendix, I briefly outline some work on open-source tools for neuroscience that I worked on throughout my thesis work simultaneously with the projects described here.

Bibliography

- Addams, R. (1834). An account of a peculiar optical phenomenon seen after having looked at a moving body. *Philosophical Magazine: A Journal of Theoretical, Experimental and Applied Physics*.
- Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., and Muckli, L. (2010). Stimulus Predictability Reduces Responses in Primary Visual Cortex. *The Journal of Neuroscience*, 30(8):2960–2966.
- Angelucci, A. and Bressloff, P. C. (2006). Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. *Progress in Brain Research*, 154:93–120.
- Ayala, Y. A. and Malmierca, M. S. (2013). Stimulus-specific adaptation and deviance detection in the inferior colliculus. *Frontiers in Neural Circuits*, 6.
- Ballard, D. H. (1987). Modular learning in neural networks. In *Proceedings of the Sixth National Conference on Artificial Intelligence-Volume 1*, pages 279–284. AAAI Press.
- Bar, M. (2003). A Cortical Mechanism for Triggering Top-Down Facilitation in Visual Object Recognition. *Journal of Cognitive Neuroscience*, 15(4):600–609.
- Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hämäläinen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., and Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America*, 103(2):449–454.
- Barlow, H. B. (1961). Possible Principles Underlying the Transformations of Sensory Messages. *Sensory Communication*, 1.
- Barlow, H. B. and Hill, R. M. (1963). Evidence for a Physiological Explanation of the Waterfall Phenomenon and Figural After-effects. *Nature*, 200(4913):1345–1347.
- Bassett, D. S., Bullmore, E., Verchinski, B. A., Mattay, V. S., Weinberger, D. R., and Meyer-Lindenberg, A. (2008). Hierarchical Organization of Human Cortical Networks in Health and Schizophrenia. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 28(37):9239–9248.
- 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(4):695–711.
- Bolz, J. and Gilbert, C. D. (1986). Generation of end-inhibition in the visual cortex via interlaminar connections. *Nature*, 320(6060):362–365.

- Bortone, D. S., Olsen, S. R., and Scanziani, M. (2014). Translaminar Inhibitory Cells Recruited by Layer 6 Corticothalamic Neurons Suppress Visual Cortex. *Neuron*, 82(2):474–485.
- Bridgeman, B. (2007). Efference copy and its limitations. *Computers in Biology and Medicine*, 37(7):924–929.
- Briggs, F. and Usrey, W. M. (2008). Emerging views of corticothalamic function. *Current Opinion in Neurobiology*, 18(4):403–407.
- Browning, M., Behrens, T. E., Jocham, G., O’Reilly, J. X., and Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature neuroscience*, 18(4):590–596.
- Bubic, A., von Cramon, D. Y., and Schubotz, R. I. (2010). Prediction, Cognition and the Brain. *Frontiers in Human Neuroscience*, 4.
- Callaway, E. M. (1998). Local Circuits in Primary Visual Cortex of the Macaque Monkey. *Annual Review of Neuroscience*, 21(1):47–74.
- Carandini, M. and Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1):51–62.
- Cavanagh, P. (1997). Predicting the present. *Nature*, 386(6620):19–21.
- Cavanaugh, J. R., Bair, W., and Movshon, J. A. (2002). Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *Journal of Neurophysiology*, 88(5):2530–2546.
- Celikel, T. and Sakmann, B. (2007). Sensory integration across space and in time for decision making in the somatosensory system of rodents. *Proceedings of the National Academy of Sciences*, 104(4):1395–1400.
- Chater, N., Tenenbaum, J. B., and Yuille, A. (2006). Probabilistic models of cognition: Conceptual foundations. *Trends in Cognitive Sciences*, 10(7):287–291.
- Chelazzi, L., Miller, E. K., Duncan, J., and Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, 363(6427):345–347.
- Chung, S., Li, X., and Nelson, S. B. (2002). Short-term depression at thalamocortical synapses contributes to rapid adaptation of cortical sensory responses in vivo. *Neuron*, 34(3):437–446.
- Constantinople, C. M. and Bruno, R. M. (2013). Deep Cortical Layers Are Activated Directly by Thalamus. *Science*, 340(6140):1591–1594.
- Cook, J., Barbalat, G., and Blakemore, S.-J. (2012). Top-down modulation of the perception of other people in schizophrenia and autism. *Frontiers in Human Neuroscience*, 6.
- Courchesne, E., Hillyard, S. A., and Galambos, R. (1975). Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalography and Clinical Neurophysiology*, 39(2):131–143.

- Crick, F. (1984). Function of the thalamic reticular complex: The searchlight hypothesis. *Proceedings of the National Academy of Sciences*, 81(14):4586–4590.
- Crook, J. M., Kisvárdy, Z. F., and Eysel, U. T. (1998). Evidence for a contribution of lateral inhibition to orientation tuning and direction selectivity in cat visual cortex: Reversible inactivation of functionally characterized sites combined with neuroanatomical tracing techniques. *The European Journal of Neuroscience*, 10(6):2056–2075.
- Crowell, J. A. and Banks, M. S. (1996). Ideal observer for heading judgments. *Vision Research*, 36(3):471–490.
- Dayan, P., Hinton, G. E., Neal, R. M., and Zemel, R. S. (1995). The Helmholtz machine. *Neural Computation*, 7(5):889–904.
- Desimone, R. and Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, 18(1):193–222.
- Dima, D., Roiser, J. P., Dietrich, D. E., Bonnemann, C., Lanfermann, H., Emrich, H. M., and Dillo, W. (2009). Understanding why patients with schizophrenia do not perceive the hollow-mask illusion using dynamic causal modelling. *NeuroImage*, 46(4):1180–1186.
- Douglas, R. J., Martin, K. A., and Whitteridge, D. (1989). A Canonical Microcircuit for Neocortex. *Neural Computation*, 1(4):480–488.
- Dragoi, V., Sharma, J., and Sur, M. (2000). Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron*, 28(1):287–298.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113(4):501–517.
- Engel, A. K., Fries, P., König, P., Brecht, M., and Singer, W. (1999). Temporal Binding, Binocular Rivalry, and Consciousness. *Consciousness and Cognition*, 8(2):128–151.
- Essen, D. C. V. and Maunsell, J. H. R. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in Neurosciences*, 6:370–375.
- Fairhall, A. L., Lewen, G. D., Bialek, W., and de Ruyter van Steveninck, R. R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature*, 412(6849):787–792.
- Feldmeyer, D. (2012). Excitatory neuronal connectivity in the barrel cortex. *Frontiers in Neuroanatomy*, 6:24.
- Felleman, D. J. and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, 1(1):1–47.
- Ferrer, I., Fabregues, I., and Condom, E. (1986a). A Golgi study of the sixth layer of the cerebral cortex. I. The lissencephalic brain of Rodentia, Lagomorpha, Insectivora and Chiroptera. *Journal of Anatomy*, 145:217–234.
- Ferrer, I., Fabregues, I., and Condom, E. (1986b). A Golgi study of the sixth layer of the cerebral cortex. II. The gyrencephalic brain of Carnivora, Artiodactyla and Primates. *Journal of Anatomy*, 146:87–104.

- Fiorentini, A. (1972). Mach Band Phenomena. In Jameson, D. and Hurvich, L. M., editors, *Visual Psychophysics*, number 7 / 4 in Handbook of Sensory Physiology, pages 188–201. Springer Berlin Heidelberg.
- Fletcher, P. C. and Frith, C. D. (2009). Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience*, 10(1):48–58.
- Freeman, W. T. and Pasztor, E. C. (1999). Learning low-level vision. In *The Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999*, volume 2, pages 1182–1189 vol.2.
- Friston, K. and Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1521):1211–1221.
- Friston, K. J. and Stephan, K. E. (2007). Free-energy and the brain. *Synthese*, 159(3):417–458.
- Fukui, T. and Tanaka, S. (1997). A Simple Neural Network Exhibiting Selective Activation of Neuronal Ensembles: From Winner-Take-All to Winners-Share-All. *Neural Computation*, 9(1):77–97.
- Furl, N., van Rijsbergen, N. J., Treves, A., and Dolan, R. J. (2007). Face adaptation after-effects reveal anterior medial temporal cortex role in high level category representation. *NeuroImage*, 37(1):300–310.
- Gómez-Laberge, C., Smolyanskaya, A., Nassi, J. J., Kreiman, G., and Born, R. T. (2016). Bottom-Up and Top-Down Input Augment the Variability of Cortical Neurons. *Neuron*.
- Gomez-Marin, A., Paton, J. J., Kampff, A. R., Costa, R. M., and Mainen, Z. F. (2014). Big behavioral data: Psychology, ethology and the foundations of neuroscience. *Nature Neuroscience*, 17(11):1455–1462.
- 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. *The Journal of Neuroscience*, 27(37):9817–9823.
- Graboi, D. and Lisman, J. (2003). Recognition by Top-Down and Bottom-Up Processing in Cortex: The Control of Selective Attention. *Journal of Neurophysiology*, 90(2):798–810.
- Gregory, R. L. (1968). Perceptual illusions and brain models. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 171(1024):279–296.
- Grieve, K. L. and Sillito, A. M. (1991). A re-appraisal of the role of layer VI of the visual cortex in the generation of cortical end inhibition. *Experimental Brain Research*, 87(3):521–529.
- Griffiths, T. L., Kemp, C., and Tenenbaum, J. B. (2008). Bayesian models of cognition.
- Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10(1):14–23.

- Grüsser, O. J. (1995). On the history of the ideas of efference copy and reafference. *Clio Medica (Amsterdam, Netherlands)*, 33:35–55.
- Harris, K. D. and Shepherd, G. M. G. (2015). The neocortical circuit: Themes and variations. *Nature Neuroscience*, 18(2):170–181.
- Hawkins, J. and Blakeslee, S. (2005). *On Intelligence*. St. Martin’s Griffin, New York, reprint edition edition.
- Hinton, G. E., Osindero, S., and Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7):1527–1554.
- Hommel, B., Müsseler, J., Aschersleben, G., and Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *The Behavioral and Brain Sciences*, 24(5):849–878; discussion 878–937.
- Hubel, D. H. and Wiesel, T. N. (1974). Uniformity of monkey striate cortex: A parallel relationship between field size, scatter, and magnification factor. *The Journal of Comparative Neurology*, 158(3):295–305.
- Hupé, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., and Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature*, 394(6695):784–787.
- Hutson, K. A. and Masterton, R. B. (1986). The sensory contribution of a single vibrissa’s cortical barrel. *Journal of Neurophysiology*, 56(4):1196–1223.
- Huys, Q. J. M., Maia, T. V., and Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, 19(3):404–413.
- James, W. (1890). *The Principles of Psychology*. Henry Holt and Company, United States.
- 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 (New York, N.Y.)*, 350(6264):aac9462.
- Katz, Y., Heiss, J. E., and Lampl, I. (2006). Cross-whisker adaptation of neurons in the rat barrel cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 26(51):13363–13372.
- Khatri, V. and Simons, D. J. (2007). Angularly nonspecific response suppression in rat barrel cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, 17(3):599–609.
- Kleinfeld, D. and Deschenes, M. (2011). Neuronal Basis for Object Location in the Vibrissa Scanning Sensorimotor System. *Neuron*, 72(3):455–468. WOS:000296826000004.
- Knierim, J. J. and van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology*, 67(4):961–980.
- Knill, D. C. (1998). Discrimination of planar surface slant from texture: Human and ideal observers compared. *Vision Research*, 38(11):1683–1711.
- Knill, D. C. and Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12):712–719.

- Kok, P., Bains, L. J., van Mourik, T., Norris, D. G., and de Lange, F. P. (2016). Selective Activation of the Deep Layers of the Human Primary Visual Cortex by Top-Down Feedback. *Current Biology*, 26(3):371–376.
- Konkle, T., Wang, Q., Hayward, V., and Moore, C. I. (2009). Motion Aftereffects Transfer between Touch and Vision. *Current Biology*, 19(9):745–750.
- Kosslyn, S. M. and Andersen, R. A. (1995). *Frontiers in Cognitive Neuroscience*. MIT Press.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In Pereira, F., Burges, C. J. C., Bottou, L., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc.
- Kulkarni, T. D., Whitney, W. F., Kohli, P., and Tenenbaum, J. (2015). Deep Convolutional Inverse Graphics Network. In Cortes, C., Lawrence, N. D., Lee, D. D., Sugiyama, M., and Garnett, R., editors, *Advances in Neural Information Processing Systems 28*, pages 2539–2547. Curran Associates, Inc.
- Kutas, M. and Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427):203–205.
- Larkum, M. E., Zhu, J. J., and Sakmann, B. (1999). A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature*, 398(6725):338–341.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444.
- Lewis, L. D., Voigts, J., Flores, F. J., Schmitt, L. I., Wilson, M. A., Halassa, M. M., and Brown, E. N. (2015). Thalamic reticular nucleus induces fast and local modulation of arousal state. *eLife*, 4:e08760.
- López-Aranda, M. F., López-Téllez, J. F., Navarro-Lobato, I., Masmudi-Martín, M., Gutiérrez, A., and Khan, Z. U. (2009). Role of Layer 6 of V2 Visual Cortex in Object-Recognition Memory. *Science*, 325(5936):87–89.
- Lui, J. H., Hansen, D. V., and Kriegstein, A. R. (2011). Development and Evolution of the Human Neocortex. *Cell*, 146(1):18–36.
- Ma, W. J., Beck, J. M., Latham, P. E., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, 9(11):1432–1438.
- Maass, W. (2000). On the Computational Power of Winner-Take-All. *Neural Computation*, 12(11):2519–2535.
- Mach, E. (1897). *Contributions to the Analysis of the Sensations*. Open Court Pub Co, University of Michigan.
- Manning, C., Tibber, M. S., Charman, T., Dakin, S. C., and Pellicano, E. (2015). Enhanced Integration of Motion Information in Children With Autism. *The Journal of Neuroscience*, 35(18):6979–6986.
- Mante, V., Sussillo, D., Shenoy, K. V., and Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, 503(7474):78–84.

- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Henry Holt and Co, New York, NY, USA.
- Marr, D. and Poggio, T. (1976). From Understanding Computation to Understanding Neural Circuitry. Technical report, Massachusetts Institute of Technology, Cambridge, MA, USA.
- Martin, K. A. and Whitteridge, D. (1984). Form, function and intracortical projections of spiny neurones in the striate visual cortex of the cat. *The Journal of Physiology*, 353:463–504.
- Martinez-Trujillo, J. C. and Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current biology: CB*, 14(9):744–751.
- Marx, M. and Feldmeyer, D. (2013). Morphology and physiology of excitatory neurons in layer 6b of the somatosensory rat barrel cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, 23(12):2803–2817.
- Maunsell, J. H. R. and Cook, E. P. (2002). The role of attention in visual processing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 357(1424):1063–1072.
- Maunsell, J. H. R. and Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, 29(6):317–322.
- McAlonan, K., Cavanaugh, J., and Wurtz, R. H. (2008). Guarding the gateway to cortex with attention in visual thalamus. *Nature*, 456(7220):391–394.
- Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience*, 1(1):59–65.
- Miller, K. D. (2016). Canonical computations of cerebral cortex. *Current Opinion in Neurobiology*, 37:75–84.
- Mountcastle, V. (1998). *Perceptual Neuroscience: The Cerebral Cortex*. Harvard University Press, Cambridge, MA.
- Mountcastle, V. B. (1957). Modality and topographic properties of single neurons of cat's somatic sensory cortex. *Journal of Neurophysiology*.
- Movshon, J. A. and Lennie, P. (1979). Pattern-selective adaptation in visual cortical neurones. *Nature*, 278(5707):850–852.
- Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biological Cybernetics*, 66(3):241–251.
- Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends in Cognitive Sciences*, 6(9):387–393.
- Oliva, A., Torralba, A., Castelhano, M. S., and Henderson, J. M. (2003). Top-down control of visual attention in object detection. In *2003 International Conference on Image Processing, 2003. ICIP 2003. Proceedings*, volume 1, pages I–253–6 vol.1.

- Olsen, S. R., Bortone, D. S., Adesnik, H., and Scanziani, M. (2012). Gain control by layer six in cortical circuits of vision. *Nature*, 483(7387):47–52.
- Olshausen, B. A. and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609.
- O’Regan, J. K., Rensink, R. A., and Clark, J. J. (1999). Change-blindness as a result of ‘mudsplashes’. *Nature*, 398(6722):34–34.
- Pascal Vincent, H. L. (2010). Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. *Journal of Machine Learning Research*, 11(12):3371–3408.
- Prinz, W. (1990). A Common Coding Approach to Perception and Action. In Neumann, D. O. and Prinz, P. D. W., editors, *Relationships Between Perception and Action*, pages 167–201. Springer Berlin Heidelberg.
- Pulvermüller, F. and Shtyrov, Y. (2003). Automatic processing of grammar in the human brain as revealed by the mismatch negativity. *NeuroImage*, 20(1):159–172.
- Purkinje, J. E. (1820). Beiträge zur näheren Kenntnis des Schwindels aus heautognostischen Daten.
- Rao, R. P. N. (2004). Bayesian Computation in Recurrent Neural Circuits. *Neural Computation*, 16(1):1–38.
- Rao, R. P. N. and Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1):79–87.
- Reynolds, J. H., Pasternak, T., and Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron*, 26(3):703–714.
- Riesenhuber, M. and Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2(11):1019–1025.
- Ritt, J. T., Andermann, M. L., and Moore, C. I. (2008). Embodied information processing: Vibrissa mechanics and texture features shape micromotions in actively sensing rats. *Neuron*, 57(4):599–613.
- Rockel, A. J., Hiorns, R. W., and Powell, T. P. S. (1980). The Basic Uniformity in Structure of the Neocortex. *Brain*, 103(2):221–244.
- Siegle, J. H., Pritchett, D. L., and Moore, C. I. (2014). Gamma-range synchronization of fast-spiking interneurons can enhance detection of tactile stimuli. *Nature Neuroscience*, 17(10):1371–1379.
- Simons, D. J. and Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, 1(7):261–267.
- Singer, W., Gray, and M., C. (1995). Visual Feature Integration and the Temporal Correlation Hypothesis. *Annual Review of Neuroscience*, 18(1):555–586.

- Stevenson, I. H., Cronin, B., Sur, M., and Kording, K. P. (2010). Sensory Adaptation and Short Term Plasticity as Bayesian Correction for a Changing Brain. *PLoS ONE*, 5(8).
- Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., and Duncan, J. (2013). Dynamic Coding for Cognitive Control in Prefrontal Cortex. *Neuron*, 78(2):364–375.
- Stuart, G. J. and Häusser, M. (2001). Dendritic coincidence detection of EPSPs and action potentials. *Nature Neuroscience*, 4(1):63–71.
- Sutton, S., Braren, M., Zubin, J., and John, E. R. (1965). Evoked-Potential Correlates of Stimulus Uncertainty. *Science*, 150(3700):1187–1188.
- Tenenbaum, J. B., Griffiths, T. L., and Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in cognitive sciences*, 10(7):309–318.
- Tenenbaum, J. B., Kemp, C., Griffiths, T. L., and Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *science*, 331(6022):1279–1285.
- Thomson, A. M. (2010). Neocortical Layer 6, A Review. *Frontiers in Neuroanatomy*, 4.
- Tiitinen, H., May, P., Reinikainen, K., and Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature*, 372(6501):90–92.
- Tsodyks, M. V. and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Sciences*, 94(2):719–723.
- Turi, M., Karaminis, T., Pellicano, E., and Burr, D. (2016). No rapid audiovisual recalibration in adults on the autism spectrum. *Scientific Reports*, 6:21756.
- Uhlhaas, P., Pipa, G., Lima, B., Melloni, L., Neuenschwander, S., Nikolić, D., Singer, W., Uhlhaas, P. J., Pipa, G., Lima, B., Melloni, L., Neuenschwander, S., Nikolić, D., and Singer, W. (2009). Neural synchrony in cortical networks: History, concept and current status. *Frontiers in Integrative Neuroscience*, 3:17.
- Ulanovsky, N., Las, L., Farkas, D., and Nelken, I. (2004). Multiple Time Scales of Adaptation in Auditory Cortex Neurons. *The Journal of Neuroscience*, 24(46):10440–10453.
- Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience*, 6(4):391–398.
- van Beers, R. J., Wolpert, D. M., and Haggard, P. (2001). Sensorimotor integration compensates for visual localization errors during smooth pursuit eye movements. *Journal of Neurophysiology*, 85(5):1914–1922.
- Vélez-Fort, M., Rousseau, C. V., Niedworok, C. J., Wickersham, I. R., Rancz, E. A., Brown, A. P. Y., Strom, M., and Margrie, T. W. (2014). The Stimulus Selectivity and Connectivity of Layer Six Principal Cells Reveals Cortical Microcircuits Underlying Visual Processing. *Neuron*, 83(6):1431–1443.
- Vincent, S. B. (1912). *The Functions of the Vibrissae in the Behavior of the White Rat ...* University of Chicago.

- Vogels, T. and Abbott, L. (2007). Gating Deficits in Model Networks: A Path to Schizophrenia? *Pharmacopsychiatry*, 40(S 1):S73–S77.
- Voigts, J., Herman, D. H., and Celikel, T. (2015). Tactile object localization by anticipatory whisker motion. *Journal of neurophysiology*, 113(2):620–632.
- Voigts, J., Sakmann, B., and Celikel, T. (2008). Unsupervised whisker tracking in unrestrained behaving animals. *Journal of neurophysiology*, 100(1):504–515.
- von Bekeésy, G. (1967). *Sensory Inhibition*. Princeton University Press, Princeton.
- von Helmholtz, H. (1867). *Handbuch der physiologischen Optik*. Verlag Leopold Voss, Leipzig.
- Weiss, Y., Simoncelli, E. P., and Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, 5(6):598–604.
- Welker, W. I. (1964). Analysis of Sniffing of the Albino Rat. *Behaviour*, 22(3/4):223–244.
- Wiser, A. K. and Callaway, E. M. (1996). Contributions of individual layer 6 pyramidal neurons to local circuitry in macaque primary visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 16(8):2724–2739.
- Wiser, A. K. and Callaway, E. M. (1997). Ocular dominance columns and local projections of layer 6 pyramidal neurons in macaque primary visual cortex. *Visual Neuroscience*, 14(2):241–251.
- Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science (New York, N.Y.)*, 269(5232):1880–1882.
- Woolsey, T. A. and Van der Loos, H. (1970). The structural organization of layer IV in the somatosensory region (SI) of mouse cerebral cortex. The description of a cortical field composed of discrete cytoarchitectonic units. *Brain Research*, 17(2):205–242.
- Zhang, S., Xu, M., Kamigaki, T., Do, J. P. H., 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*, 345(6197):660–665.
- Zhang, Z. W. and Deschênes, M. (1998). Projections to layer VI of the posteromedial barrel field in the rat: A reappraisal of the role of corticothalamic pathways. *Cerebral Cortex (New York, N.Y.: 1991)*, 8(5):428–436.
- Zhaoping, L. and Jingling, L. (2008). Filling-In and Suppression of Visual Perception from Context: A Bayesian Account of Perceptual Biases by Contextual Influences. *PLoS Computational Biology*, 4(2).

Chapter 2

Neocortical representation of stimulus changes

In this chapter, I describe some technical developments that allowed me to record large-scale spike trains from multiple cortical layers in awake mice (Voigts et al., 2013), and to quantify population encoding in deep layers of cortex using calcium imaging. I will then describe results on how the layered neocortical architecture encodes stimulus changes, specifically how pyramidal cells in layers 2/3 and 4, and corticothalamic cells in layer 6 differ in their change encoding.

2.1 Chronic high channel count microdrive recording in mice

The findings in this section were previously published in Voigts J., Siegle J.H., Pritchett D.L. and Moore C.I. The flexDrive: An ultra-light implant for optical control and highly parallel chronic recording of neuronal ensembles in freely moving mice. (2013) *Front. Syst. Neurosci.* 7:8. doi: 10.3389/fnsys.2013.00008.

To characterize layer-specific encoding of these stimulus deviations, we required stable recordings from large populations of stimulus driven neurons in awake, behaving mouse S1. While optical imaging methods are now capable of exceeding the yields of electrophysiological methods, the ability to record unambiguous individual action potentials from multiple depths at once, for long periods, under simultaneous optogenetic manipulation, together with the very high quality of the data made chronic electrophysiology the only method capable of collecting the data required for this study.

In this study, we implanted newly developed, high-density arrays of movable multi-contact electrodes (Voigts et al., 2013) in vibrissal primary somatosensory neocortex (S1). This method yielded ~ 25 identified neurons per session (5 mice), and made it possible to reliably target regions of S1 associated with specific vibrissae, with a recording quality that permitted clear distinction between regular spiking and fast-spiking neurons. Together, these advantages made it possible to characterize the laminar representation of stimulus changes in awake behaving mice.

Electrophysiological recordings from ensembles of neurons in behaving mice are a central

tool in the study of neural circuits. Despite the widespread use of chronic electrophysiology, the precise positioning of recording electrodes required for high-quality recordings remains a challenge, especially in behaving mice. The complexity of available drive mechanisms, combined with restrictions on implant weight tolerated by mice, limits current methods to recordings from no more than 4-8 electrodes in a single target area. We developed a highly miniaturized yet simple drive design that can be used to independently position 16 electrodes with up to 64 channels in a package that weighs approximately 2g. This advance over current designs is achieved by a novel spring-based drive mechanism that reduces implant weight and complexity. The device is easy to build and accommodates arbitrary spatial arrangements of electrodes. Multiple optical fibers can be integrated into the recording array and independently manipulated in depth. Thus, our novel design enables precise optogenetic control and highly parallel chronic recordings of identified single neurons throughout neural circuits in mice.

2.1.1 Introduction

Neuroscience increasingly relies on ensemble recordings that characterize not only individual neurons, but also the complex interplay of neurons within local circuits and across different brain areas (Miller and Wilson, 2008). Recently, the development of optogenetic tools (Boyden et al., 2005; Zhang et al., 2007; Cardin et al., 2010; Kravitz and Kreitzer, 2011; Pastrana, 2011) facilitated the precise, cell-type specific optical manipulation of neural circuits in behaving animals. The availability of transgenic mouse lines that allow the expression of light-gated ion channels in specific cell types makes it especially desirable to probe the interactions between cell types in neural circuits with large simultaneous recordings in behaving mice.

While the use of large-scale recordings in behaving animals has been highly successful in primates (Serruya et al., 2002; Lebedev and Nicolelis, 2006; Nicolelis et al., 2003; Buschman and Miller, 2007; Feingold et al., 2012), and to some degree in rats (Nicolelis et al., 1993), adapting the approach to mice has been difficult due to their smaller size. During experiments, implant weight can be offset by a pulley system with a counter weight (Yamamoto and Wilson, 2008), or by placing the animal in a headpost (Dombeck et al., 2007). However, a mouse's comfort and post-op survival depends critically on its ability to move, eat, and drink in its home cage, imposing a weight limit of approximately 4g on implants. For applications that combine recording with behavioral phenotyping (Crawley, 2007), an implant weight closer to 2g is required so that the mice can move freely or so that the implant can be tilted in a way that does not occlude the field of view in experiments requiring videography (Ritt et al., 2008; Voigts et al., 2008). Studies that address the distributed development of neural systems in adolescent animals also require lighter implants.

One approach to minimizing implant weight while keeping channel count high is to use static electrode arrays, thereby relieving the added weight of multiple independent drives (Bragin et al., 2000). While useful for many applications, there are obvious drawbacks to this approach. First, if drive placement is in any way inaccurate initially, electrode position cannot be corrected. Second, a more subtle but equally important concern is the need to move electrodes to continue to obtain high-quality units. One of the main constraints on the duration that any electrode can yield high quality single-unit data is gliosis, the process

of successive encapsulation of foreign materials by glial cells that insulate the electrode from surrounding neurons (Polikov et al., 2005; Turner et al., 1999). Even though stable recording conditions can be maintained over months in optimal conditions (Tseng et al., 2011; Freire et al., 2011), the process of glial encapsulation begins as early as one day post-implant (Fujita et al., 1998) and can lead to a progressive deterioration in the experimenters ability to identify and discriminate individual neurons (Dickey et al., 2009; Williams et al., 1999; Muthuswamy et al., 2011; Vetter et al., 2004). Similarly, small movements of electrodes relative to the surrounding tissue can damage the neuropil and lead to a decline in unit yield over time.

Consequently, obtaining and maintaining high-quality recordings over many days requires the ability to precisely reposition the recording electrodes in the neural tissue long after the initial implant surgery. By lowering electrodes after the date of implantation, their recording sites can be repeatedly moved out of the zone of neural degradation or glial migration (Jackson et al., 2010; Muthuswamy et al., 2011). (See Fig.2-2e for an example where lowering an electrode by $\sim 60\mu\text{m}$ restores the recording quality of a stereotrode that previously lost the ability to resolve units.) Such motions also make it possible to delay the onset of the electrode-related tissue reactions altogether until after the completion of behavioral training, for example. While there are mouse-specific commercially available implants with up to 8 individually movable electrodes (VersaDrive, Neuralynx, Bozeman MT, see Dobbins et al. (2007) for protocol), as well as custom-built designs that can combine optical fibers with tetrodes (Fee and Leonardo, 2001; Korshunov, 2006; Lansink et al., 2007; Kloosterman et al., 2009; Haiss et al., 2010; Anikeeva et al., 2012), no current method combines the key features of low weight, high number of individually movable electrodes, high placement stability, and independently adjustable optical fibers. These features are essential for obtaining high-quality, parallel, and distributed recordings within and across neural circuits while providing precise optogenetic control of target neurons. While solutions based on microscale motors are starting to become feasible (Muthuswamy et al., 2011), it will be several years before they surpass the established microwire electrode methods in terms of channel count, availability of multi-site electrodes, ease of use, and cost.

To overcome these limitations in current methods, we have developed a simple, highly miniaturized drive design that replaces the drive mechanisms found in current implants with a one-piece spring design. This results in a significant reduction in drive size and weight, without sacrificing channel count. Our *flexDrive* fits up to 16 individually movable electrodes or electrode bundles and can maintain stable recording conditions for months. The design integrates guides for two or more optical fibers, which can either be static or adjustable). Due to its small size and low weight ($\sim 2\text{g}$, $\sim 2\text{cm}$ height, $\sim 1.5\text{cm}$ diameter), the *flexDrive* is well tolerated by mice with only minimal impact on natural behavior.

The design includes support for 16, 32 or 64 channels and interfaces with standard amplifier connectors. The drive can be assembled in about one day after little training, and can be customized to fit specific experimental requirements such as different electrode types, spatial arrangement of the electrodes and optical fibers, or amplifier interface. The custom-made parts of the *flexDrive* can be ordered from online vendors or workshops with the provided design specifications. All design files are available under the TAPR Open Hardware license, which requires others to make adaptations of the design freely available as well (see www.github.com/open-ephys/flexdrive and www.open-ephys.org/flexdrive).

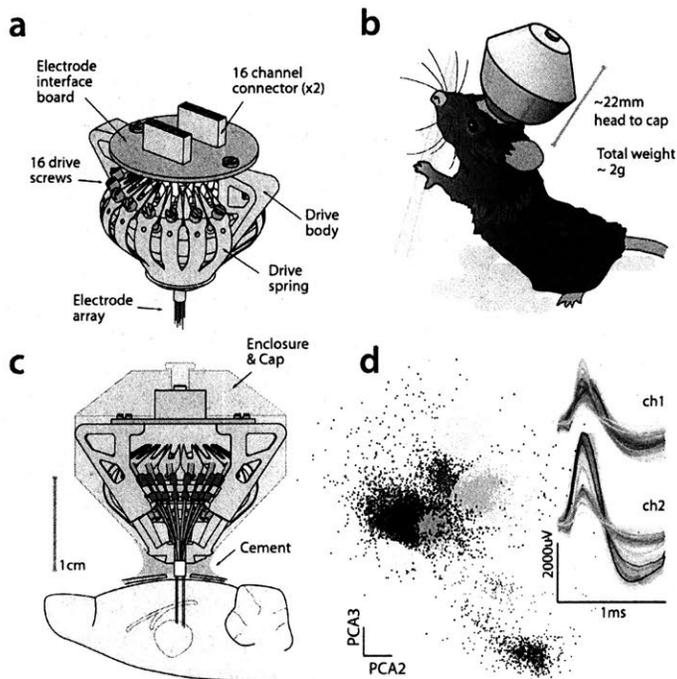


Figure 2-1: The flexDrive provides a low-weight and high-yield method for chronic electrophysiology. **a**, Isometric view of the flexDrive showing the one-piece spring (blue) that acts as the drive mechanism. **b**, Illustration of the flexDrive implanted in a 6 month old C57/bl6 mouse. Due to the low implant weight ($\sim 2\text{g}$), the impact of the drive on natural behavior is minimal. **c**, Cross section of the drive and its placement on the mouse skull. In this example, electrodes target the thalamus. **d**, Cortical action potentials recorded from a stereotrode ($12\mu\text{m}$ nichrome wire, gold plated to $\sim 300\text{K}\Omega$) on a flexDrive showing eight clusters (color coded clusters, non clustered spikes in grey) and average and 95% percentiles of the waveforms on the two electrode contacts.

The design described here presents a significant improvement in the quality and quantity of the data that can be obtained in experiments using optogenetic circuit manipulations in mice, enabling the study of the concerted function of large neural circuits, rather than local neurons.

2.1.2 Materials and methods for drive implant construction

Given that large-scale recordings of neural activity rely on precise positioning of many electrodes, we designed the flexDrive around a method that allow the experimenter to arrange electrodes in a variety of patterns. In our design, electrodes are positioned by an array of flexible polyimide tubes. By placing individual guide tube arrays at different locations within the drive body, multiple brain regions can be targeted precisely. This control gives researchers the ability to adapt the design to fit their specific experimental needs, such as recording from elongated but narrow target regions (or from bilateral targets) with a single implant (Fig.2-2a).

The array of guide tubes is assembled by building up layers of polyimide tubes and fixing them with cyanoacrylate glue. Alternatively, arranging the guide tubes within a larger guide cannula makes this process faster, but sacrifices some flexibility. By using only a subset of the guide tubes to hold electrodes, or by introducing placeholders and optical fibers into the array, any spatial pattern of electrode and optical fiber positions can be fabricated with high repeatability and precision. The electrodes are free to move laterally within the guide tubes. Such laterally flexible anchoring of electrodes has been shown to decrease adverse tissue reactions (Biran et al., 2007).

The closest lateral spacing between electrodes that can be accomplished with this method is dictated by the outer diameter of the guide tubes. We recommend a distance of $\sim 250\mu\text{m}$ or larger for the guide tubes (using 33 gauge), but higher densities of $\sim 125\mu\text{m}$ are possible by using smaller diameter guide tubes. However, tests conducted with dense electrode arrays of pitches of $125\mu\text{m}$ failed to yield usable recordings, possibly due to an increased inflammatory response.

The array of guide tubes is attached to a plastic drive body (Fig.2-1a, Fig.2-2a) that is manufactured from an ABS-like material using stereolithography (Accura55 American Precision Prototyping, proprietary material). This drive body supports all components of the drive and facilitates fast and precise assembly. While most components are eventually fixed with an epoxy glue, the design features snap-fit grooves, facilitating the alignment of the guide tubes and the spring. By customizing the locations of the guide tubes in the drive body CAD file, precise targeting of separate recording sites are readily achieved.

2.1.3 Drive mechanism

A central constraint on data collection in chronic electrophysiology is the difficulty of recording the activity of identified, individual neurons (termed units). While the use of tetrodes (Gray et al., 1995; Wilson and McNaughton, 1993) or stereotrodes (McNaughton et al., 1983) have made it possible to reliably identify individual neurons in recordings, obtaining sufficiently clear data from large numbers of electrodes remains a challenge. The presented drive design addresses this constraint by enabling highly parallel recordings in mice without sacrificing the ability to precisely reposition many individual electrodes.

To enable the precise positioning and re-positioning of electrodes in the awake mouse brain, we replaced the traditional multi-part drive design (Kloosterman et al., 2009) with a simplified mechanism in which a single spring and one screw per electrode form the adjustment mechanism. Each electrode (or electrode bundle) is inserted through a piece of polyimide tubing, called a shuttle tube, that can move up and down in its guide tube. The linear motion that drives the electrodes is provided by the 16 arms of a single steel spring that are each held down by an adjustment screw (Fig.2-2a). Each shuttle tube is attached to one of the arms of the spring. By turning the screw, the spring arm is either pressed down or released, which translates to a linear motion of the shuttle tube within the guide tube, moving the electrode in the brain. Each turn of a screw corresponds to $\sim 250\mu\text{m}$ in electrode motion. The tension and lateral stability of the spring arm ensures that there is no sideways travel or twisting of the shuttle tube when the screw is adjusted, and allows the

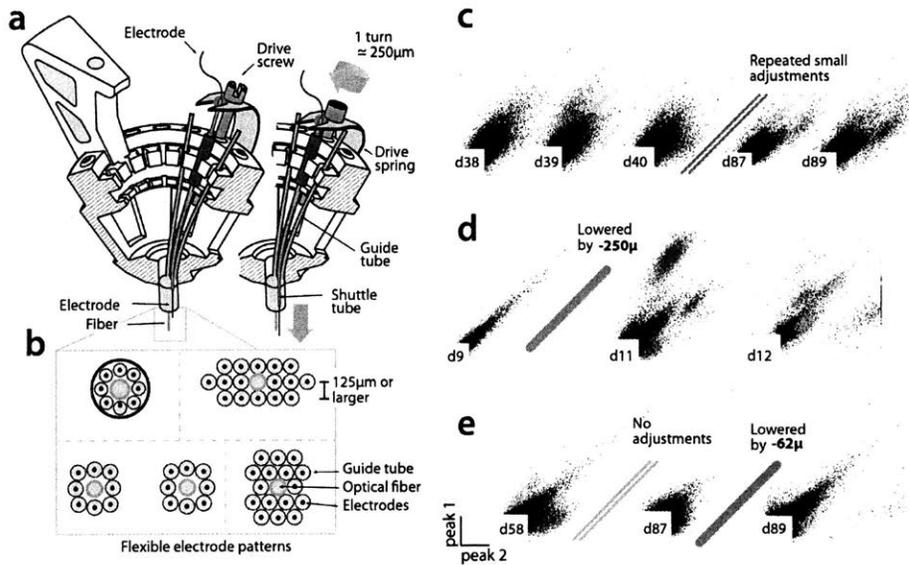


Figure 2-2: **The drive mechanism of the flexDrive.** **a**, Isometric view of the spring loaded drive mechanism. The pattern of electrodes is defined by an array of guide tubes (blue). Electrodes (black) are fixed inside shuttle tubes (orange) that can move up or down inside the guide tubes. The top of each shuttle tube is glued to a spring arm that is moved up or down by a drive screw. **b**, Examples of electrode patterns that can be fabricated by arrangement of the guide tubes and optical fibers. (c-e) Examples of identified units on stereotrodes, all plots peak/peak. **c**, Units can be maintained on an electrode for > 100 days by repeated small increments in electrode depth. **d**, Example of an electrode that was not penetrating the cortex at surgery, but is lowered into the brain later. **e**, Example of an electrode that loses the ability to discriminate units over time, but is reactivated by a small depth adjustment ~ 3 months after surgery.

electrode to be moved smoothly both up and downwards. By adjusting the drive screws in small increments, fine grained control over electrode depth is possible. In practice, we found that quarter turns ($\sim 62\mu\text{m}$) are a useful step size that can be used to recover recording quality on electrodes that have lost the ability to resolve units due to tissue degeneration (Fig.2-2e). This step size is larger than the minimal movement required for recovering units (Yamamoto and Wilson, 2008) but is easy to achieve by manual adjustment.

Implant fabrication is greatly simplified by this spring loaded drive design. The array of guide tubes is glued to the drive body and individual tubes snap-fit into prepared grooves that ensure proper alignment. The spring is then attached in one step and its arms are moved under the screw heads. The shuttle tubes are then glued to these spring arms, completing the construction of the drive mechanism and making the drive ready for loading with electrodes.

By making it possible to individually adjust electrode depth, this design facilitates the targeting of small target regions, enables significantly higher unit yields over longer time spans than previously possible, and enables highly parallel recordings in awake, behaving

mice.

2.1.4 Optical fibers

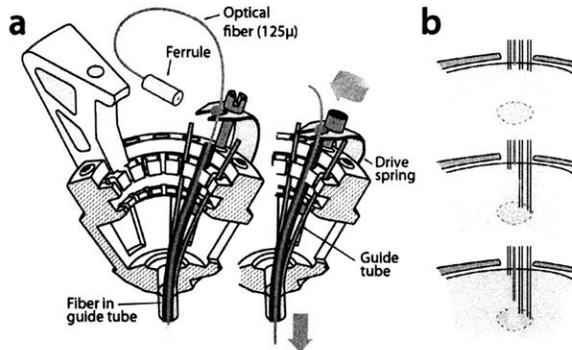


Figure 2-3: **Variant of the flexDrive in which an optical fiber is lowered in the brain by one of the 16 drive mechanisms.** **a**, The fiber is inserted through a guide tube and fixed to a drive spring, replacing a shuttle tube and electrode. The remaining 15 drives can be used for electrodes or more fibers. **b**, Sketch of the workflow of an experiment made possible through moveable fibers and electrodes. A target area (dashed lines) is localized by slowly lowering a subset of electrodes first, then the fiber can be brought into optimal position for localized activation of the area or for the collection of optical signals.

Experiments using optogenetic manipulation of neural circuits often require spatially distinct recording and stimulation sites. In the flexDrive, optical fibers can be built into the guide tube array at any desired position and depth (Fig.2-2b) and can remain static while any of the surrounding electrodes are lowered (Fig.2-2a). Such an arrangement limits the deformation of brain tissue during electrode adjustment compared to methods in which fibers and electrodes move together (Anikeeva et al., 2012), and ensures that optical stimulation parameters will be reproducible throughout the entire experiment. However, if independent adjustment of the fiber depth throughout the experiment is desired, one or more small diameter fibers ($\sim 125\mu\text{m}$) can be inserted in place of electrodes and can be lowered using the same spring-driven mechanism (Fig.2-3). In this case, the fiber is lowered into a guide tube and glued to the spring in place of a shuttle tube. The free upper end of the fiber with the ferrule connector (extending ~ 2 cm past the spring) is then looped around and fixed perpendicular to the electrode interface board using epoxy (Fig.2-3a). This free loop provides enough flexibility for the fiber to move up and down. If desired, electrodes can be glued to the fibers at constant depth offset (Anikeeva et al., 2012). In practice, we find that attaching 2 ferrules to the electrode interface board is straightforward, though in principle up to 16 fibers could be attached.

This variant of the flexDrive enables researchers to precisely position optical fibers to electrically identified target areas (Fig.2-3b), or to compare optical manipulation of neural activity in different positions in the same animal. Further, this method enables the collection of optical signals from fluorescent probes (O'Connor et al., 2009; Scanziani and Häusser,

2009) while simultaneously recording extracellularly.

2.1.5 Results

Recording from optically activated identified neurons and distributed, small targets To verify the utility of our drive design for optical activation of neurons in a chronic behaving mouse, we implanted parvalbumin (PV)-Cre+ mice with a double-floxed adeno-associated carrying the gene for channelrhodopsin-2 (ChR2) (Cardin et al., 2009, 2010; Wang and Carlén, 2012). All experimental procedures were in accordance with the guidelines published in the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the Brown University Institutional Animal Care and Use Committee. The virus was injected into the barrel field the primary somatosensory cortex (SI) during the same surgical procedure in which the flexDrive was implanted. Due to injection depth, the virus was expressed predominantly in the fast-spiking interneurons in cortical layers 2/3 and 4, as confirmed by post-mortem histology. We implanted the mice with drives constructed with 8 tetrodes surrounding a static 200 μ m optical fiber. The tetrodes were lowered individually while the fiber remained fixed on the surface of the neocortex. By adjusting the depth of individual electrodes, we were able to record neurons in SI (Fig 4a,b) over the span of the experiment (\sim 3 months) with an average yield of 2.50 cells/tetrode resulting in 20 ± 4.7 simultaneously recorded units (N = 8 tetrodes in the target region over 17 sessions). This approach presents a marked improvement in the quality of data compared to prior experiments in which we used static electrode implants.

In a separate experiment, we implanted mice with flexDrives designed to target both SI and the thalamic reticular nucleus (TRN) (Fig.2-4c). The TRN presents a thin target, and recording units from this brain region has proven challenging in the past (Halassa et al., 2011). Using the flexDrive, we were able to slowly advance electrodes until the electrophysiological signature of the recording indicated that we reached TRN (identified by elevated tonic firing rate and thin spike waveforms (Pinault, 2004)). In these targets regions, our method resulted in a unit yield of 1.43 cells/stereotrode (N = 32 electrodes over 16 sessions each) which enabled recordings from \sim 10–20 units per recording session (Fig.2-4d).

As shown by these examples, the precise positioning of 16 individual multi-site electrodes afforded by the flexDrive allows us to record spike trains from identified single neurons in regions that previously could not be reliably targeted with multiple electrodes.

2.1.6 Discussion

Combining high density parallel recordings of identified neurons throughout neural circuits with the specificity of optogenetic control is essential for experiments seeking to understand complex neural circuits. Recently, studies have demonstrated the great utility of simultaneous optogenetic interventions and single- and multi-unit recordings from awake, behaving mice (Halassa et al., 2011; Cardin et al., 2010). However, these studies lacked the ability to record neural activity from more than a few identified neurons at a time, mainly due to the use of static electrodes. In addition, recordings of identified neurons across multiple brain regions have been limited by the complexity and weight of the required implants.

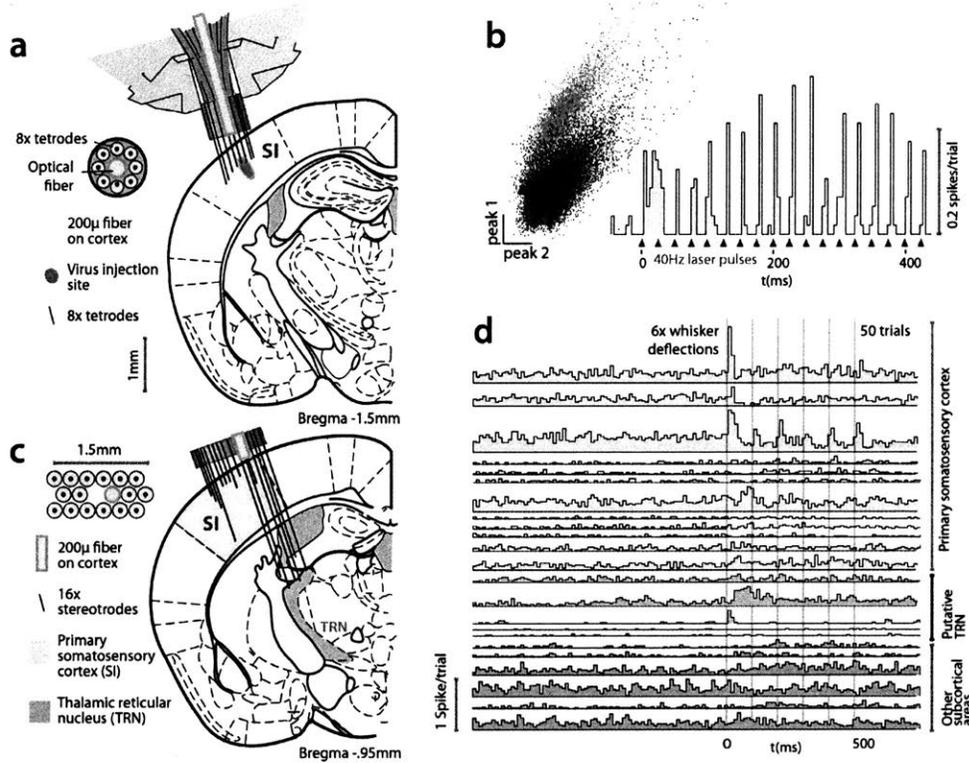


Figure 2-4: Example applications of the flexDrive for experiments that require stable optical excitation of neurons, or simultaneous recordings from distributed, small target regions. **a**, Activation of PV-positive neurons in layer 2/3 of mouse primary somatosensory cortex (SI) with ChR2. An array of 8 tetrodes arranged in a circular pattern around a static 200µm fiber (see insert) were slowly lowered into layer 2/3 of SI. **b**, Example trace of an identified PV neuron on one of the tetrodes for one session. **c**, Experiment in which an array of 16 stereotrodes was used to simultaneously record from SI and the thalamic reticular nucleus (TRN) in awake, behaving mice. The electrode positions are shown for the 3rd day after the first electrodes reached TRN. **d**, Example peri-stimulus time histograms of 23 simultaneously recorded single units. A subset of the recorded neurons in SI and TRN are modulated by vibrissa deflections induced with a piezoelectric stimulator.

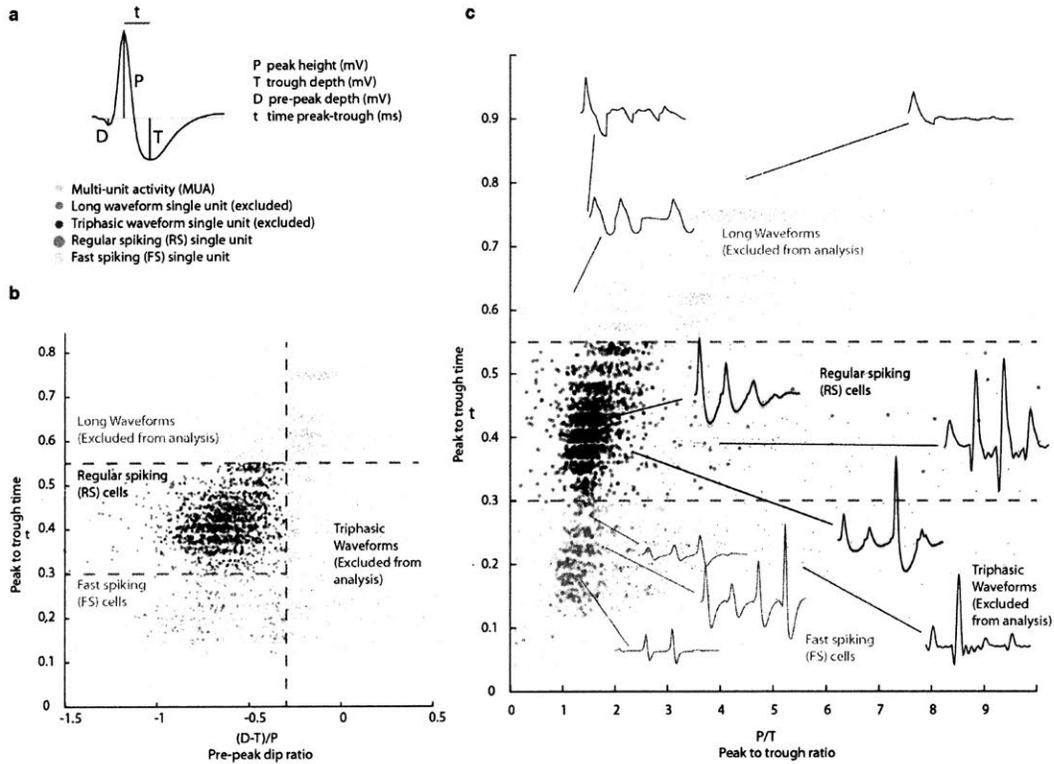


Figure 2-5: **Classification of spike waveforms into fast spiking (FS) and regular spiking (RS) neurons.** **a**, Waveform measurements used for unit classification. All waveforms were digitized at 30KHz and band-pass filtered offline using a zero-phase (acausal) FIR filter at 300-9000Hz to avoid introduction of phase distortion due to filtering. **b**, Scatter plot of all recorded putative neurons, pre-dip ratio vs. peak to trough. Dotted lines indicate the thresholds for classification of units. **c**, Scatter plot of peak to trough ratio vs. peak to trough time. Example waveforms are plotted with mean and std over concatenated waveforms from the 4 tetrode contacts. For illustration purposes, the discrete time variable is plotted with an additional uniform scatter of 0.9 samples. Triphasic waveforms, and waveforms with no clear hyper-polarization, 'long' waveforms we

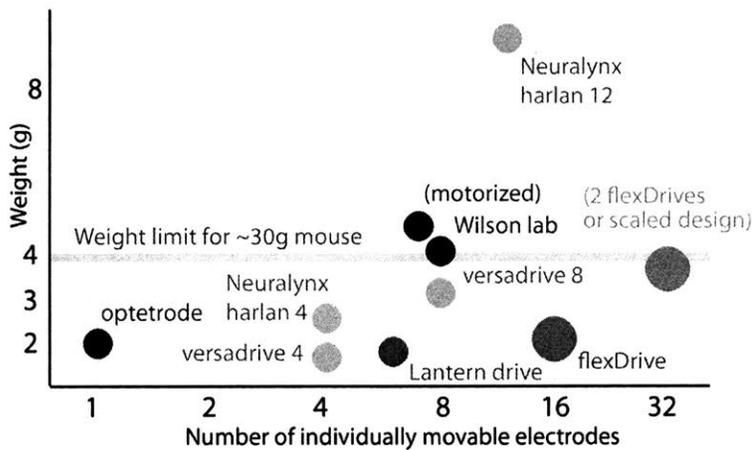


Figure 2-6: **Comparison between existing types of implants and the flexDrive.** Our novel design results in a higher number of individually movable electrodes at a reduced implant weight compared to existing methods (Yamamoto and Wilson (2008); Battaglia et al. (2009); Kloosterman et al. (2009); Anikeeva et al. (2012), Neuralynx- Bozeman MT). The drive weight of ~ 2 g enables experimenters to either implant two drives per mouse, or to scale the design to 32 driven electrodes per implant.

Here, we described the design of a drive implant that provides the ability to record from 16 individually adjustable multi-contact electrodes simultaneously for months in awake, behaving mice while optogenetically manipulating neural activity. By replacing the complex mechanisms employed in previous drive designs by a simple spring design (Fig.2-1,2-2) the flexDrive is significantly lighter (Fig.2-6) and easier to construct. The ability to independently adjust each electrode over months (Fig.2-2) allows for high flexibility in recording from small target areas and results in higher yields of well-isolated single-unit activity, over significantly longer time spans than would be possible with static implants.

By repeatedly lowering electrodes, multiple attempts at obtaining good recordings can be made. This is especially important when targeting specific cell types in small target regions (Fig.2-4b), thin laminated structures such as the cell layers of the hippocampus (Kloosterman et al., 2009), or specific cortical layers (Fig.2-4a). Delaying the lowering of electrodes until after surgery also increases the targeting precision because electrodes can be positioned after initial brain swelling has subsided (Cole et al., 2011). This procedural step is of increased importance for the large craniotomies required for distributed recordings over more than one target site. Finally, implantation of large electrode arrays with pitches below $250\mu\text{m}$ raises the risk of brain deformation during insertion due to the increased localized friction between the electrodes and brain tissue (Rennaker et al., 2005). This problem can be mitigated by individually lowering the electrodes one at a time.

We have demonstrated the ability of the flexDrive to record from light-driven PV-positive interneurons in layers 2/3 of the primary somatosensory cortex (SI) of awake mice using ChR2 (Fig.2-4a,b). The light was delivered to the neurons through an optical fiber positioned on the pial surface, showing the ability of the drive to record from optically driven neurons with independent positioning of the recording electrodes relative to the light

source. We further demonstrated the utility of the design by recording simultaneous single neurons from SI and the thalamic reticular nucleus (TRN) (Fig.2-4c,d) showing that the design can combine highly parallel recordings from 16 electrodes with the positioning accuracy required to observe neurons in small, deep targets such as the TRN.

To conclude, the flexDrive presents a straightforward method for obtaining stable and high-quality electrophysiological data from multiple target sites in awake, behaving mice. This permits researchers to make full use of the precision and specificity of optogenetic methods by directly probing the concerted function of neural circuits, rather than individual neurons.

In addition to the work described in this thesis, the drive design is already been adopted by multiple labs and has already been used in published studies including in mice (Siegle and Wilson, 2014), and lizards/dragons (Shein-Idelson et al., 2016).

2.2 Methods

2.2.1 Animal subjects

NTSR1-Cre mice (strain B6.FVB(Cg)-Tg(Ntsr1-cre)GN220Gsat/Mmcd, stock number 030648-UCD) (Gong et al., 2007) were used. For some experiments, NTSR1-Cre mice were crossed with a floxed Chr2 reporter line (strain B6;129S-Gt (ROSA) 26Sortm32 (CAG-COP4*H134R/EYFP) Hze/J, stock number 012569). For electrophysiology, 5 mice of either sex, 4 NTSR1-Cre mice using viral injections and 1 reporter line cross was used. For 2-photon imaging, 6 NTSR-1 Cre mice with viral injections were used.

2.2.2 Viral injections

For virus mediated Chr2 expression (Chapter 3), we targeted the caudal region of the barrel field (1.5 mm posterior to bregma and 3.5 mm lateral to the midline). Injections were performed through a burr-hole with a glass micropipette (pulled and beveled, tip diameter of 20-35 μ m) attached to a stereotaxic-mountable syringe pump (QSI Stoelting). 300 nl of virus (AAV DIO Chr2-mCherry, $\sim 2 * 10^{12}$ viral molecules per ml) was injected at 0.05 μ l /min at $\sim 800\mu$ m below the dura.

All experiments requiring viral transfection were performed > 4 weeks after injection. For 2 photon imaging, ~ 300 nl of AAV2/1-hSyn-Flex-GCaMP6s (HHMI/Janelia Farm, GENIE Project; $\sim 1 * 10^{13}$ viral molecules per ml) (Chen et al., 2013), or in a subset of mice a 1:1 mixture of floxed GCaMP6 and floxed Chr2, all produced by the U. Penn Vector Core, was injected at $\sim 750\mu$ m targeting the posterior c-row barrels, identified by vascular landmarks and confirmed using intrinsic imaging to restrict expression to NTSR1+ neurons. Mice were tested for aberrant expression outside of L6 CT cells either by histology (for behaviour and electrophysiology), or by collecting z-stacks (for 2-photon imaging). Mice with fluorescent non-L6 cells were excluded from the study.

2.2.3 Surgical procedures

Mice were 8–14 weeks old at the time of surgery. Animals were individually housed and maintained on a 12-h reversed cycle. All procedures and animal care protocols conformed to guidelines established by the National Institutes of Health, and approved by Brown Universitys Institutional Animal Care and Use Committee. Mice were anesthetized with isoflurane (2% induction, 0.75–1.25% maintenance in 1 l/min oxygen) and secured in a stereotaxic apparatus. The scalp was shaved, wiped with hair-removal cream and cleaned with iodine solution and alcohol. After intraperitoneal (IP) injection of Buprenex (0.1 mg/kg) and dexamethasone (4 mg/kg) and local application of lidocaine, the skull was exposed. For some mice, AAV was injected as described. The skull was cleaned with ethanol, and a base of adhesive luting cement (C&B Metabond) was applied. A 0.5 mm diameter area of the skull over left primary somatosensory cortex was thinned. A stub of fiber-optic cable (0.22 NA, inner diameter of 200 μ m, 1.25 mm OD metal ferrule) was glued into place at the side of the craniotomy using transparent luting cement. Custom head posts (www.github.com/open-ephys/headposts_etc) were affixed with luting cement, the incision was closed with VetBond (3M), and mice were removed from isoflurane. Mice were given 310 d to recover before the start of water restriction. For electrophysiology, we implanted flexDrives (Voigts et al., 2013) with 16 stereotrodes (N=2, 17 sessions), or tetrodes (N=3, 58 sessions) made from 12.5 μ m coated OD polyimide-coated nichrome wire (Kanthal), twisted, heated and gold-plated to 200 – 400k Ω impedance (Nguyen et al., 2009). Lateral electrode spacing was 250 μ m .

For acute recordings (Fig. 3-2), the same procedure as for 2-photon imaging was used, but the window was omitted. Mice remained on 0.75–1.25% isoflurane for maintenance, the craniotomy was kept covered in warm saline and a laminar silicone probe (Neuronexus A1x16-3mm-50-177) was lowered into SI.

Two stainless-steel screws were implanted anterior to bregma to serve as ground. For some mice we injected AAV as described. A craniotomy was drilled over left S1 (\sim 1.5 mm posterior, 3.5 mm lateral, \sim 2.5 mm diameter). A fiberoptic stub was added as for behavioral testing, and a large durotomy was opened. A layer of bacteriostatic surgical lubricant was added, and the drive was lowered at an angle of \sim 15 and fixed in place using dental cement. After recovery ($>$ 3 d), mice were habituated to the setup and electrodes were lowered into the brain (\sim 2hrs between individual electrodes) while noting when each electrode penetrated the brain. Mice were water restricted as described. During the experimental life time of mice, electrodes were advanced to target neocortical layers and maintain recording quality (Voigts et al., 2013).

2.2.4 Experimental design for electrophysiology and 2-photon imaging

Mice rested on a styrofoam ball supported by an air cushion (Fig.1a). Mice were water restricted and monitored as described, and licked a spout to indicate stimulus detection for reward, but no time-outs, or catch trials were used. Sessions were stopped at signs of animal distress and were increased over the first 2–3 weeks of acclimatization, resulting in sessions of \sim 2000 trials over \sim 2h. Stimuli were delivered as described, with deviants of relative amplitude of the deflections (Fig.2-17). In a subset of sessions, changes in direction

were employed, as in the head-fixed detection task. Amplitudes were calibrated to the range that correspond to 80-100% hit rate in the behavioral detection task. The interval between stimuli was 3–5s. For 2-photon imaging, mice were not water restricted and stimulus amplitudes were sampled from two baseline stimuli and two deviant conditions (increase to 120% or decrease to 80%) in order to increase statistical power for the sparse L6 responses.

2.2.5 Analysis of electrophysiology data

Unless indicated, we used non-parametric Wilcoxon rank sum / Mann-Whitney U-test tests for comparing groups (non-paired), or Wilcoxon signed rank tests for testing medians versus zero or comparing paired measurements. Extracellular voltage traces were band-pass filtered to 1-10000Hz at acquisition (3rd order butterworth filter), and band-pass filtered to 300-9000Hz (zero-phase acausal FIR) for analysis of spiking. Sessions in which vibrissae had slipped out of the stimulator were excluded. Spikes were sorted into single units using Simple Clust (<http://github.com/open-ephys/simpleclust>). The 90% quantiles of neuron count/session were 17 and 39 over all 75 sessions. We recorded the depth at which electrodes penetrated the brain, marked by emergence of off-diagonal peak-to-peak amplitudes in the MUA activity (presumably from L1 axons) as the 0mm position to estimate the depth of electrodes. We combined this information with the drive screw position and angle of the drive to estimate depth. In deeper layers, we additionally used the depth at which electrodes entered the white matter (loss of cortical activity) as a further reference point.

The mapping from depth to layers was approximately :

Layer	depth
L2/3	100–350 μm
L4	350–450 μm
L5	450–650 μm
L6	>650 μm

The depth estimates were adjusted to take electrode angle and curvature of cortex and white matter borders into account. Drive depth estimates were verified at the L3 /L4 boundary of primary somatosensory cortex (SI), via the stimulus evoked LFP signature (Castro-Alamancos and Connors, 1996). We classified neurons as regular spiking (RS) and fast-spiking (FS) by spike waveform (Bortone et al., 2014), (Fig.2-5). Stimulus driven neurons were classified by fitting a generalized linear model (Truccolo et al., 2005) (GLM) to the PSTH. We classified cells as phasically driven if either of two conditions were met:

(i) An offset term and 6 bins (basis functions) spanning the first 100ms of the first vibrissa deflection were fit, coefficients for at least 2 bins were significantly nonzero at a P level of 0.03 and any coefficients other than the offset term had a lower standard error bound > 0.002.

(ii) A constant term and 6 repeated bins for over first 100ms of the first 3 deflections were fit, these coefficients were shared between deflections capturing cases of weaker but sustained stimulus drive. Additionally, one parameter for each 100ms vibrissa deflection period after the first one was used to avoid false positives due to slower firing rate drifts.

Cells were classified as driven if the coefficients for the first 3 deflections satisfied the same conditions as in (i). Classifications was verified manually to choose thresholds but no manual corrections were made. To plot example PSTHs (Fig.2-10), we computed confidence bounds using a state space method (Smith et al., 2010). These analyses were used for visualization purposes only. The random position of deviant stimuli resulted in more baseline than deviant deflections, and more baseline stimuli early in the train (stimuli after the deviants were not analyzed). For analyses that are susceptible to biases of unequal N and adaptation effects, such as change coefficients, a histogram matching procedure was used to match the number and position in the stimulus train across baseline and deviant stimuli (Fig.2-17). The effect of stimulus history on firing rates was analyzed using subsets of trials in which the stimulus amplitudes were matched but were preceded by higher or lower amplitude stimuli by matching the stimulus amplitude distributions (Extended Data Figure 3). All statistics of change coefficients were computed as 95% confidence intervals (CI) of the median using a 1,000 or 10,000-fold bootstrap. Spread of distributions of change coefficients was quantified as the difference between the observed and a surrogate distribution (computed from position matched, randomly re-sampled baseline stimuli) via the interquartile range (75th -25th percentile) and Shannon entropy (in bits):

$$H(\text{observed}) - H(\text{surrogate}) \text{ with } H(h) = - \sum_{i \in \text{bins}} (P(h_i) * \log_2 P(h_i))$$

95% confidence bounds and significance levels for these statistics were determined via bootstrap. Entropy was quantified via the difference between pairs of binned distributions, so choice of bin size had no significant effect. Where paired samples per cell were available, as in the effect of the optogenetic manipulation, a bootstrap on the median of the absolute value minus the population median was used: $\text{median}(|\text{coeff}_{\text{cell}} - \text{coeff}_{\text{population}}|)$. See Extended Data Figure 21 for change coefficients for RS and FS neurons across all layers 2/3, 4 and 5 in the control and laser condition. We quantified encoding in individual neurons with a GLM. We analyzed parameters for spike count as a function of stimulus deviation, mirroring the direct computation of change-coefficients (Fig.2-10, Fig.2-12). The features used in the model were stimulus deviation (-1:decreases, +1: increases), baseline amplitude, and spiking history (for 7 precedent deflections). The adaptation profile was modeled with a separate feature per deflection, linked with a quadratic penalty term on the pairwise difference (weight 10). A separate quadratic regularization term with (weight 1) penalized large parameters (other than constant) to avoid over-fitting. The regularizing matrix (q) was:

$$q = 1 * I + 10 * D, \quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 2 & \ddots & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & 2 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

(I: Identity matrix, D: difference operator).

Model parameters (w) were estimated from spike counts (Y) via

$$\min_w (-\log(p(Y|w)) + 0.5 * w'q'w)$$

and 95% confidence bounds were obtained using a 100-fold bootstrap. False positive rates were calculated by fitting to surrogate data (as described). Control, laser, and deviant conditions (increases/decreases) were fitted independently. See Fig.3-12 for results for all neuron types and conditions.

2.3 Laminar specific encoding of stimulus changes

Investigation of the neocortical encoding of stimulus changes requires a stimulus design that engages cortical mechanisms for change encoding, and specifically one that results in an observable transformation between the afferent thalamic input to the cortical column, and the resulting change encoding.

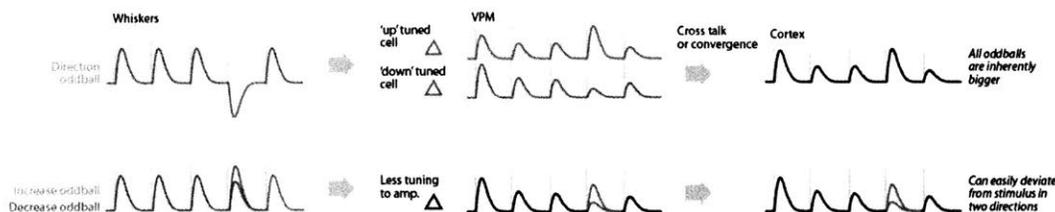


Figure 2-7: **Feature changes can elicit pre-cortical SSA.** **Top:** Directional vibrissa deflections drive tuned thalamic populations that adapt over multiple deflections, and subsequent changes in direction recruit new, less adapted neurons. Cross talk and/or convergence in the thalamocortical afferents then result in higher cortical afferent drive for such deviants. **Bottom:** By instead deviating only in stimulus amplitude, this pre-cortical process could be avoided, giving us a more direct means of controlling the stimulus driven afferent cortical signal.

Typical stimulus deviant paradigms employ variations in stimulus features, such as orientations of gratings or other visual features (Woods and Frost, 1977; Sobotka and Ringo, 1994; Müller et al., 1999), deviant tones (Ulanovsky et al., 2003; Taaseh et al., 2011; Chen et al., 2015), or vibrissa identity (Katz et al., 2006) or deflection direction (Andermann and Moore, 2006) which lead to pre-cortical adaptation and subsequent higher afferent cortical drive for all deviant stimuli. Stimulus-tuned neurons along the afferent pathway and at thalamocortical synapses adapt to repeated stimulation (Chung et al., 2002; Katz et al., 2006; Khatri and Simons, 2007), and subsequent deviant stimuli activate new pools of less adapted neurons at higher rates, leading to increased neocortical drive, even though SSA also involves cortical representations, specifically SOM positive interneurons (Natan et al., 2015). This process is termed stimulus-specific adaptation or 'SSA' (Movshon and Lennie, 1979). In theory, this effect can occur whenever there are tuned populations of pre-cortical neurons (Fig.2-7).

Here, we sought to avoid SSA because it presents a confound for studying layer-wise differences in change encoding: If all novel stimuli are also of higher amplitude, that is encoded with increased cortical afferent drive, this first order effect could mask more specific laminar computations. It is also not clear whether local cortical mechanisms of change encoding would be engaged by stimulus changes that are accompanied by such an increased afferent drive. Also, most of the computations that motivate this thesis, that is the com-

putation of a mismatch signal between a stimulus expectation and the bottom-up input in a cortical circuit, are not required for stimuli that are sufficiently different from previous stimuli to be marked as salient by SSA. Finally, it seems clear from the effects of ‘descending’ cortico-cortical input on local stimulus representations (Knierim and van Essen, 1992; Chelazzi et al., 1993; Reynolds et al., 2000; Maunsell and Treue, 2006) that there should be mechanisms for change encoding beyond pre-cortical SSA.

To disambiguate local neocortical change encoding from pre-cortical SSA, and to make it possible to observe laminar differences of change encoding, we used stimuli free of feature changes, deviating instead by changes of their amplitude. In this stimulus design, novel stimuli differed from the preceding pattern by either increased or decreased deflection amplitude, but otherwise identical. This results in stimuli that are novel and result in stronger afferent drive (as in SSA) but also stimuli that are novel but smaller. Importantly, these deviants should not drive differentially tuned pools of neurons along the afferent pathway, which makes them more suited to studying differences in their encoding across layers. Specifically, we can test whether this approach is successful in engaging cortical mechanisms of change encoding if the stimuli result in a clear difference in their encoding between L4, which would be predicted to monotonically increase firing rates if current stimuli are bigger, or decrease if they are smaller, and layers 2/3, that could encode a less directly driven change signal. Conversely, if we engage pre-cortical SSA and no cortical transformations, then we would expect an overall increase in firing rates for deviants that would be preserved across layers.

We delivered trains of deflections (7 deflections, 10 Hz, Fig.2-17a,c,d, Fig.2-10a) to a subset of vibrissae (B and C rows 1-3, matching recorded somatotopic positions) with a piezoelectric stimulator (See Fig.4-2). The waveform of the deflections was chosen to maximize instantaneous onset velocity while minimizing ringing and or hysteresis (See Fig.4-2), and the stimulus amplitude was varied randomly, either between 0 and ≈ 0 -30mm/sec for behavioural testing, or between ≈ 20 -30mm/sec for recording electrophysiological data. Deviants were single deflections of a randomly chosen amplitude between 80% – 120% of the baseline amplitude. This relatively small range was chosen for two main reasons: To keep the deviations small enough to engage cortical and no pre-cortical mechanisms (as discussed above), and to mimic the range of stimulus variation that a mouse would naturally experience when palpating its environment with its whiskers. Mice actively modulate the amount by which they protract their vibrissae with the effect of keeping the amplitude of vibrissa deflections relatively small (Towal and Hartmann, 2006; Mitchinson et al., 2007; Voigts et al., 2008), and stable across contacts (Voigts et al., 2015) regardless of what object they explore. The small range of deviants employed here is therefore likely a close approximation of naturally occurring stimulus patterns.

This stimulus design of varying baseline and deviant amplitudes yielded two primary analyses of baseline encoding, and two of deviant encoding: Effect of baseline stimulus amplitude on firing rates in the first deflection, on subsequent deflections, and encoding of current deviants, and of deviant history. For current deviants (Fig.2-17d and Fig.2-12c,d) we typically quantify the spike count difference per deflection between increases and decreases. For the deviant history, trials are analyzed that are of identical current (deviant) stimulus amplitude, but are preceded by higher or lower baseline amplitude (Fig.2-17e and Fig.2-12e).

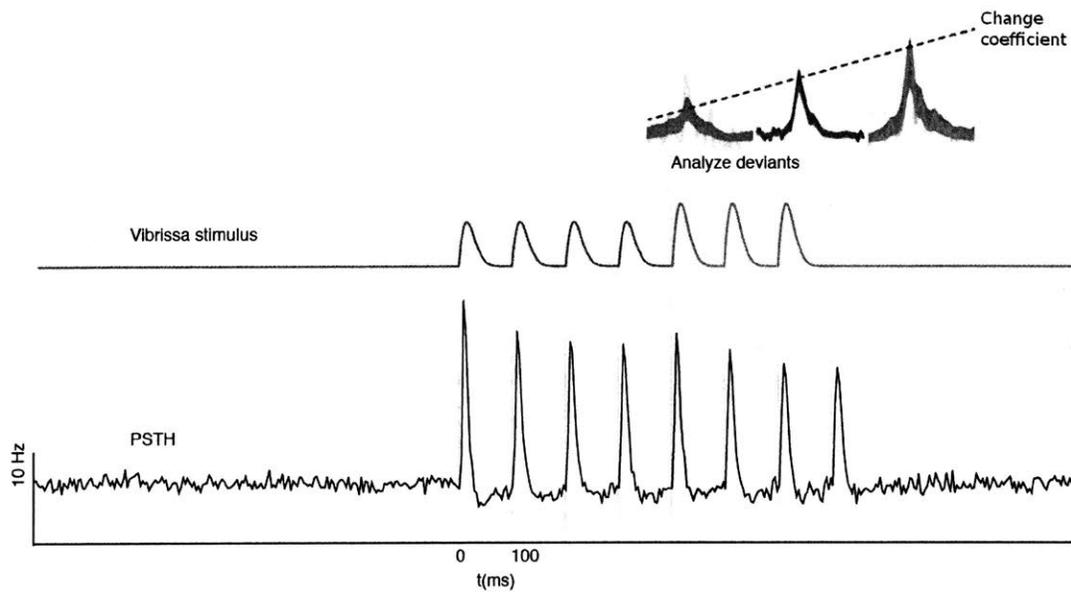


Figure 2-8: **Method for measuring change encoding in individual neurons.** Trains of vibrissa stimuli with variable baseline amplitude (grey) and small amplitude deviations (in this example an increase, red) are delivered with a piezo stimulator. The resulting evoked firing rates are computed for baseline stimuli, increases and decreases. The change coefficient is defined as the difference in spikes/stimulus between the decreases and increases. Similarly, overall firing rate increases for deviants can be computed by comparing baseline (matched for position in train) to deviant firing rates, or by comparing different baseline amplitudes.

Any stimuli after the deviant were excluded from further analysis, which results in a higher proportion of baseline stimuli early in the stimulus trains (Fig.2-17), while deviant stimuli are uniformly distributed in all positions > 2 . To compare these responses to baseline stimuli we therefore created a sub-sampled dataset from baseline stimuli that was matched to the deviant stimuli by only sampling from pre-deviant stimuli until an equal number of equally distributed baseline deflections.

For a subset of experiments we employed directional deviants to study stimulus deviations that increase the behavioural salience in a detection task (Fig.4-3). We separately examined the neural representations of these stimuli and found that stimulus direction is encoded in L6 CT cells and leads to conventional SSA (Chapter 3, Fig.3-13).

By recording with flexDrive implants, yielding ~ 25 neurons per session, we recorded 1242 neurons over 75 sessions in 5 mice. Of these cells, 395 were phasically stimulus-driven, classified via a GLM method that classifies phasic vibrissa driven responses (See Methods). To classify neurons by cortical layers I first tracked electrode depth on the drive implants and corrected this drive depth for cortical curvature, and drive angle. As a verification of recording depth I also analyzed the stimulus-evoked LFP (Castro-Alamancos and Connors, 1996) to classify neurons by cortical layer and categorized them as regular spiking

('RS'; putative excitatory pyramidal neurons) and fast-spiking ('FS'; putative inhibitory interneurons) by spike waveform (Bortone et al., 2014) (~30% were FS, Fig.2-18).

2.3.1 Layer 4 neurons encode stimulus deviations faithfully

Repetition of sensory stimuli increases the specificity of neural representations, enhancing psychophysical discrimination while responses to isolated stimuli are less specific (von Bekešy, 1967; Moore et al., 1999; Andermann et al., 2004; Tannan et al., 2006; Ollerenshaw et al., 2014).

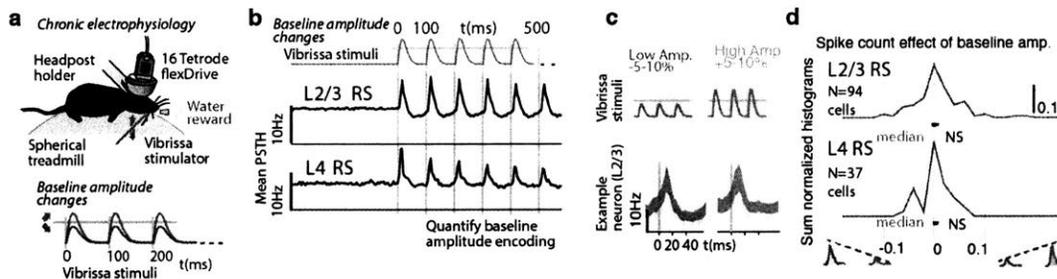


Figure 2-9: **The small amplitude of baseline stimuli employed in this study is not encoded in the firing rates of adapted cortical neurons.** **a**, Experimental preparation. Stimulus amplitudes were varied by about $\pm 15\%$ around a mean amplitude. **b**, firing rates in response to non-changing stimulus trains, excluding the first whisker deflection were analyzed. **c,d**, Neither L2/3 nor L4 cells significantly encoded the amplitude of the baseline stimulus. This is in contrast to changes in amplitude of the same magnitude, which are robustly encoded in L2/3 and L4 cells (Fig.2-12).

To study cortical mechanisms of change encoding that are likely not engaged equivalently for large stimulus deviations, we selected a narrow range of stimulus amplitudes (Fig.2-12a blue/red, baseline $\simeq 25\text{mm/sec}$, deviants = $0.8\text{-}1.2 \times \text{baseline}$, $N=72$ sessions). We chose this range of variation to be small enough not to be encoded in differences in the firing rates of layer 4 (L4) RS neurons when presented as single vibrissa deflections ($N=37$ RS; spike count difference between larger and smaller stimuli, signed rank $P=0.382$, 95% confidence interval (CI): [0.000, 0.003, 0.010]). This range of deviations is also well matched to the range of deflection amplitudes on vibrissa in free sensory behaviour (Voigts et al., 2008).

In contrast, the same range of amplitude differences was reflected in small but clearly distinguishable firing rate differences when stimuli were deviants (varying in amplitude from the preceding pattern of baseline stimuli). We quantified this encoding of deviations as the spike count difference between stimulus increases and decreases, per cell, per deflection. This *change coefficient* was positive for neurons that fired more for stimulus increases and/or less for stimulus decreases (Fig.2-12b (i, iv)). Coefficients for deviants in L4 were positive, faithfully representing stimulus amplitude changes (Fig.2-12c, $P=0.005$, CI:[0.012, 0.016, 0.035]; Fig.3-12 for generalized linear model, 'GLM' of rate changes separately for increases and decreases).

These spike count differences for deviations were also larger than the differences induced

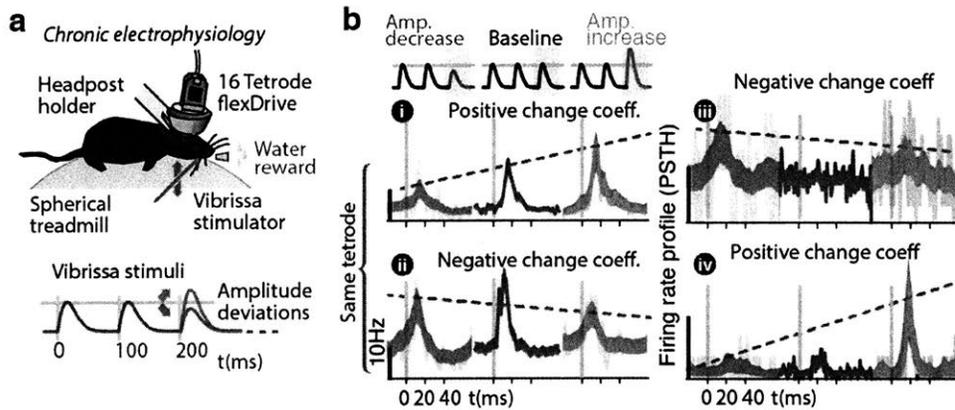


Figure 2-10: **Setup and example cell responses to small changes in stimulus amplitude** **a**, Experimental preparation. Stimulus amplitude deviations were presented after repetitions at baseline amplitude (Fig.2-17). **b**, Example sensory-driven PSTHs (95% confidence intervals, 'CI') for baseline (grey) and deviant stimuli (amplitude increases red, decreases blue). We calculated *change coefficients* (purple) as spike count differences across deviants. Neurons had positive (higher firing probability for stimulus increases - i, iv) and negative (higher probability for stimulus decreases - ii, iii) coefficients. The example in panel b (i,ii) are two neurons recorded on the same tetrode.

in the middle of the stimulus train by equivalent differences in baseline amplitude ($P=0.016$, signed rank). In sum, L4 was more sensitive to amplitude deviants (presented after repetitions of another baseline amplitude) than to equivalent differences in baseline stimulus amplitude. These differential responses, while small (<0.1 spikes/deflection) when considered for single neurons and for single vibrissal deflections, are substantial when considered across populations of neurons and for vibrissal motions during natural exploration (Fig.5-1).

To summarize, L4 RS neurons do not encode stimulus amplitudes in an adapted stimulus train, but do robustly encode small stimulus amplitude deviations (of comparable amplitude ranges as encountered by freely behaving mice (Voigts et al., 2008)). This encoding was 'faithful', that is increases in stimulus amplitude result in increases in firing probability, and vice-versa.

2.3.2 Layer 2/3 neurons encode stimulus deviations heterogeneously

In contrast to the faithful encoding in L4, L2/3 RS ($N=94$ vibrissa-driven out of 363) had no significant positive change coefficients. This finding suggests that the stimulus design was successful in engaging laminar cortical circuits for change encoding, as stimuli that elicit pre-cortical SSA would elicit increased firing rates for deviant stimuli in both L2/3 and L4 (Szymanski et al., 2009). Even though they were zero-centered, the firing rates for amplitude deviants in L2/3 had a wide distribution of change coefficients, including neurons with negative change coefficients, i.e. significant rate decreases for increased stimulus amplitudes, or increases for stimulus decreases (Fig. 2-12b (ii,iii)). To test the significance of this heterogeneity, whether L2/3 cells fired randomly (which could also result in a zero

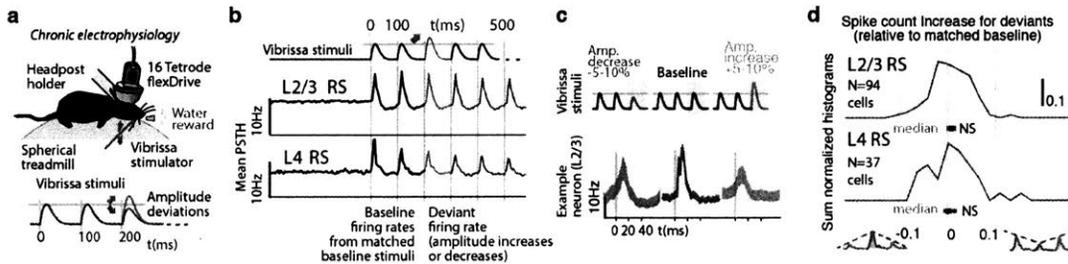


Figure 2-11: Amplitude deviants do not increase overall response amplitude. **a**, Stimulus amplitude deviations were presented after repetitions at baseline amplitude. **b,c**, Responses to deviants (same number of small amplitude increases or decreases) are measured by the spike count per deflection, and compared to responses to position-matched (Fig.2-17) baseline stimuli. The resulting overall excess or deficient spikes for deviants relative to baseline measure whether any cell fired more or less in response to the deviants, regardless of deviant identity. **d**, Histograms of this firing rate change across all stimulus-driven L2/3 and L4 RS cells. No significant firing rate effects were measured. This is in contrast to deviants defined by changes in the direction (up vs. down) of vibrissa deflections, which engages pre-cortical SSA (Fig.2-7) and results in higher firing rates (Fig.3-13).

centered distribution) or carried stimulus information, we first computed a surrogate distribution from shuffled baseline stimuli by resampling true deviant stimuli with position matched baseline stimuli of an equivalent amplitude range (Fig. 2-12c) and then compared the spread of the distributions via their interquartile range as well as their Shannon entropy (see Methods). The observed distribution of change coefficients was significantly broader than the surrogate distribution (interquartile range 'IQR': $P=0.006$, Shannon entropy: $P=0.029$, via bootstrapping, see Methods) showing that L2/3 firing was not random but carried information about the stimulus amplitude deviants.

The observed L2/3 change encoding could explicitly represent changes relative to previous stimuli (in this case ~ 100 ms before, as we used 10Hz trains of deflections), but our observations are also consistent with individual cells tuned to static stimulus amplitude ranges (Garion et al., 2014). In this case, cells with negative change coefficients could be tuned to lower stimulus amplitudes below our baseline amplitude range.

To disambiguate these possibilities, we analyzed trials where stimulus amplitudes were matched across current deviant stimuli, but were preceded by higher or lower amplitude baseline stimuli (55 sessions, Fig.2-17). In this setting, the current, analyzed stimulus amplitudes are constant, and if cells were indeed tuned to static amplitude ranges rather than explicitly encoding changes, they should not encode this preceding stimulus amplitude. We found that L2/3 RS neurons significantly represented this stimulus history (Fig.2-12e, Fig.2-13, IQR $P=0.004$, entropy $P=0.014$, $N=87$), showing that L2/3 encodes a change signal relative to preceding stimuli.

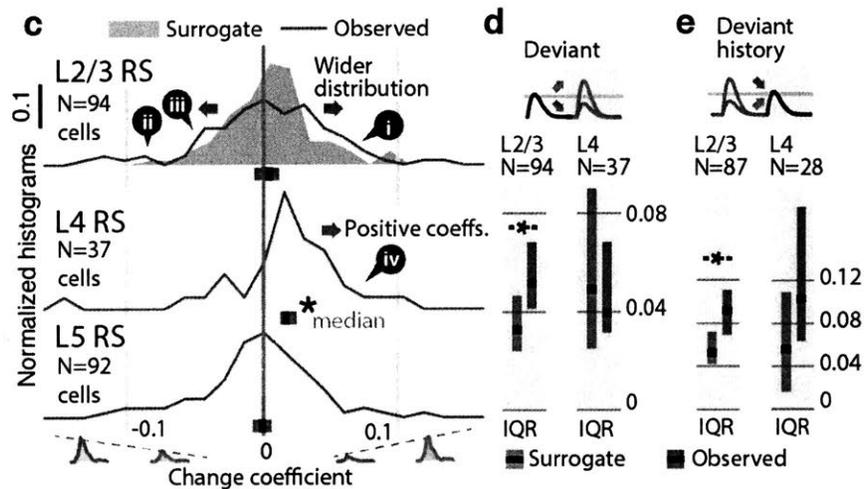


Figure 2-12: **Layer-specific encoding of small stimulus deviations.** c, Change coefficients (purple) for all stimulus driven regular spiking (RS) neurons. L4 faithfully represented stimulus changes with positive coefficients, while L2/3 coefficients were heterogeneous and broader than a null/surrogate distribution (gray). Bars are 95% CIs for median. d, Heterogeneous L2/3 encoding, quantified via interquartile range (IQR, 95% CIs vs. surrogate, see text). e, Similarly, L2/3 but not L4 firing probability encodes the amplitude of the preceding deflection (matched deviants, preceding stimuli are higher/lower amplitude).

2.3.3 Ideal observer decoding of L2/3 and L4 responses

In addition to quantifying the difference in spiking probability between upwards and downwards deviants (Fig.2-12), we examined the stimulus encoding in L2/3 and L4 by directly testing whether a spike count threshold per cell could decode the stimulus deviant identity with an above 50% correct rate. Optimal spike thresholds were chosen as the value that resulted in the best classification rate (maximally different from 50%). For cells that decrease their firing rate for increased stimulus amplitudes or vice versa (i.e., negative change coefficient neurons), the correct rate would therefore be < 50%, and for cells that carry no significant stimulus information, the rate would be at 50%.

Ideal observer decoding of deviant amplitude direction (increased vs. decreases) showed the same difference between faithful L4, and heterogeneous L2/3 deviant encoding (Fig. 2-14). Even though the range of stimulus deviations was small (0.8-1.2* baseline amplitude), this analysis also showed that the firing of L2/3 cells was not random but carried stimulus information.

The observed encoding in L2/3 corresponds to a median difference of spike counts of at least ~0.03 spikes per deflection between the stimulus categories. For a 700ms stimulus at 10Hz (as used here) this corresponds to ~70 spikes per vibrissa, or 250–1000 spikes extrapolated across a typical bout of whisking in sensory decision making ((Celikel and Sakmann, 2007; Voigts et al., 2008), Fig.5-1). In sum, unlike L4, L2/3 RS neurons encoded stimulus deviations with a heterogeneous code, without following the direction of the deviation.

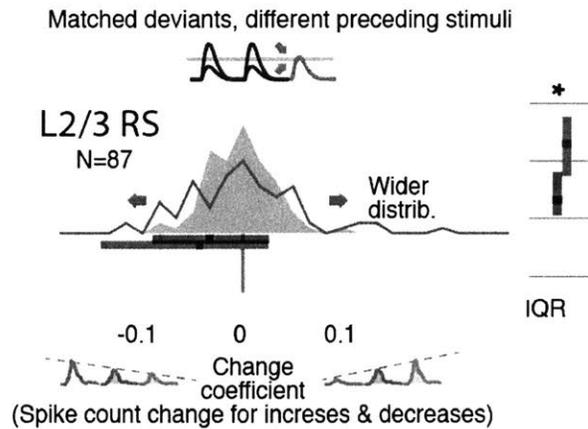


Figure 2-13: **Layer 2/3 RS cells encode stimulus history.** Stimuli were analyzed in groups that were matched for the stimulus amplitude of the analyzed stimulus, but differed by the amplitude of the preceding stimulus (Fig.2-17).

To summarize, L4 RS cells encode stimulus deviations faithfully (positive change coefficients / sign of spike rate change follows sign of stimulus amplitude changes), while L2/3 cells encode stimulus changes heterogeneously (positive and negative change coefficients), and encode stimulus history.

2.4 Stimulus encoding in Layer 6

As discussed above, L6 CT cells are well positioned to play a key role in the neocortical representation of prediction. The sparse responses in this layer in visual neocortex (Vélez-Fort et al., 2014) further suggest that highly specific patterns of activity in these neurons could play a role. We first investigated L6 responses using chronic neurophysiological recordings. Consistent with prior studies (Swadlow and Hicks, 1996; Lee et al., 2008; Thomson, 2010), L6 neurons were sparsely sensory responsive, with only 18/139 physically sensory driven (~ 13%), and low firing rates (Fig. 2c) making it impossible to reliably assess their stimulus encoding.

Further, these recordings did not allow us to distinguish between the less specifically tuned corticocortical (CC), and the very sharply tuned corticothalamic (CT) pyramidal cells that, in mouse V1, receive long-range input from higher cortical areas (Vélez-Fort et al., 2014) and are therefore well positioned to contribute to change encoding. Due to the higher stimulus selectivity of CT cells, electrophysiological recordings would likely be biased towards CC rather than CT cells (Fig.3-3).

Therefore, to obtain a sufficient sample size of genetically identified CT cells to assess receptive field properties, we used awake 2-photon calcium imaging (Denk et al., 1990) in NTSR1-Cre mice (Gong et al., 2007) expressing GCaMP6s (Chen et al., 2013) (N=3408 cell

bodies, N=2685 with amplitude deviants, 6 mice).

2.4.1 2-Photon Imaging setup

For 2-photon imaging, titanium headposts were used, and the skull around SI was thinned and flattened. A 3mm craniotomy was made, virus was injected, and a cranial window (Anderson et al., 2011; Goldey et al., 2014) 'plug' was made by stacking two 3mm coverslips (Deckgläser, #0 thickness (~ 0.1 mm), Warner, CS-3R) under a 5mm coverslip (Warner, CS-5R), using optical adhesive (Norland Optical #71). The plug was inserted into the craniotomy and the edges of the larger glass were sealed with vetbond (3M) and cemented in place (Extended Data Figures 8a, 9a). Dura was left intact. Animals were given >3d to recover. We performed intrinsic imaging (Polley et al., 2004) to localize the barrel field.

A two-photon microscope (Bruker/Prairie Technologies) using an 8 kHz resonant galvanometer (CRS) for fast x-axis scanning, and a non-resonant galvanometer (Cambridge 6215) for y-axis increments was used. In some sessions, non-resonant scanning in a smaller imaging window (variable region $\sim 100 \times 80$ px) was used. Frames were 512 x 512 pixels (resonant) or smaller (non resonant) and scanned at > 5Hz. Objectives (Nikon 25x 1.1 NA or Nikon 16x 0.8 NA) were rotated to the window plane. GCaMP6s was excited by a pre-chirped Ti-Sapphire laser (Spectra Physics; MaiTai) at 980 nm. All analysis routines were written in MATLAB. Motion artifacts, small movements in the x-y plane were corrected with rigid-body image alignment (Bonin et al., 2011) using a DFT based method (Guizar-Sicairos et al., 2008) or a similar affine deformation to register to templates averaged from 1000 low-motion frames. To manually identify ROIs, we calculated mean and standard deviation projections, and correlation coefficients for the entire image relative to a seed pixel, and areas of continuous or nearby highly correlated pixels were grouped into the ROI.

2.4.2 2-Photon Imaging in deep tissue

Even though 2-photon imaging has become a relatively routinely used method, imaging in deep cortical layers still presents a challenge. As outlined in Chapter 2, we were able to acquire good imaging data from layer 6 CT cells by a combination of several facilitating factors.

Total power at the objective entrance was 20-60 mW for superficial imaging ($< 450 \mu\text{m}$), 60-80 mW for deep imaging ($> 450 \mu\text{m}$), when scanning at ~ 5 -10Hz with an approximate pixel dwell time of 1-2 μs . Emitted photons were collected through the imaging path to a multi-alkali PMT (Hamamatsu; R3896, digitized with 14-bit resolution). A typical session lasted 2 hrs. We found no activity run-down, substantial bleaching or cellular damage over the session, consistent the low laser intensities used and with similar population imaging studies (O'Connor et al., 2010; Huber et al., 2012). In about half of implanted animals, we were able to image cell bodies of NTSR1+ layer 6 CT cells (3408 ROIs total) at depths between ~ 650 - $800 \mu\text{m}$. Good image quality at safe excitation laser powers (see above) at these depths was possible likely due to the sparse and relatively localized expression in L6 (approximate diameter of region with cell bodies ~ 300 - $400 \mu\text{m}$), which results in relatively little fluorophore above the imaging plane (Fig. 2c,d, Extended Data Figure 8), resulting in better signal to noise ratios at such depths than would be possible with denser labeling

(Theer et al., 2003). If the optical quality of the implanted window was non-optimal, due to dura re-growth, animal age or any surgical imperfections, L6 imaging became impossible.

At depths below L2/3, the window diameter can start to affect imaging quality. With large NA objectives (we almost exclusively used a 16x 0.8NA), deeper imaging planes, and imaging locations away from the center of the window, progressively more excitation light can get cut off by the edge of the window, resulting in power and effective NA loss.

Fig.2-21 shows a plot of the available 2-photon excitation power for a completely uniformly filled 0.8NA objective through a 1mm window, ignoring tissue scattering. Realistic beam profiles that deliver more power at lower angles will be affected less in terms of power, but will still lead to effective NA loss.

While a 2mm window should be big enough from this point of view when imaging in the window center, we used 3mm imaging windows, giving us plenty of room to search for sensory driven barrels to image without risking light cut off. Also, the edges of windows are rarely as clear as the center, so the extra safety margin is beneficial. This can mean not having to wait for an extra week for the window to clear sufficiently. Past 3mm, window size seems to offer little further advantages, at least for S1 imaging, and bigger windows are much harder to position flat on the cortex.

Large windows could also make it somewhat easier to collect the emitted (scattered) visible light. The rule of thumb for the surface area from which scattered photons are emitted is $\sim 1.5 \times$ imaging depth (Beaurepaire and Mertz, 2002), so a window that does not cut off excitation light should be near optimal for collection as well.

The choice of 980nm wavelength for exciting GCaMP6s gave us more ballistic photons than shorter wavelengths. Generally, longer wavelengths result in less scattering - this increase in mean free path length at longer wavelengths is a significant factor in deep imaging because scattered photons do not significantly contribute to the 2-photon excitation at the focal volume (Helmchen and Denk, 2005). We observe massively increased tissue autofluorescence at the dura for wavelengths of $> 1000\text{nm}$, so we settled on 980nm for all deep imaging.

Figure 2-22 shows a plot of the available power by depth via the Beer-Lambert law, squared to account for 2-photon excitation: $P_{\text{depth}} = (P_0 * e^{-\text{depth}/l_s})^2$ for a few wavelengths. The mean free path length estimates l_s are taken from Jacques (2013). The estimates depend on estimates of the scattering coefficients of alive neural tissue which vary substantially between tissue types, animal age and are likely affected by other experimental factors, especially the tissue condition after surgery (Beaurepaire and Mertz, 2002).

2.4.3 Layer 6 CT neurons encode stimulus amplitude

Even though L6 CT cells can be stimulus driven and can exhibit remarkable selectivity in their stimulus tuning (Vélez-Fort et al., 2014), it is not a priori clear that L6 could encode the stimulus features (amplitude ranges of $\pm 20\%$), across the entire stimulus train, and possibly for faster changes of amplitude within a stimulus train. It is especially important to explicitly test this because the sparse activity and rapid adaptation of L6 cells in S1

(Fig.3-3) make it non obvious whether stimulus history, current stimuli or deviations can be encoded by L6.

Due to the sparse labeling of L6 CT neurons by viral transfection and the very low density of fluorophore in superficial layers (Theer et al., 2003) we were able to image with good image quality (Fig.2b,e) throughout at least the upper $\sim 100\mu\text{m}$ of L6 (Fig.2c,d), and with good signal-to-noise-ratios in excess of 150% $\Delta\text{F}/\text{F}$ (Fig.2-23h, Fig.2-25).

Due to the slow timescale of the GCaMP6s indicator we did not attempt to distinguish encoding to individual vibrissa deflections (as in the electrophysiological data), and in animals where wider fields of view were possible, we chose low imaging rates as low as 5Hz in order to image larger number of cells. We did not attempt to quantify onset of change related activity within the stimulus train, but instead analyzed integrated activity by quantifying responses in a 0-2 second window post-stimulus offset. Consistent with electrophysiological data, only $\sim 13\%$ of cells were stimulus driven (469/3408 cells). When stimuli without deviants were presented, L6 CT cells encoded baseline amplitude variation faithfully, with significantly higher integrated calcium signals for larger stimuli ($P < 0.001$ signed rank, $N=346$, Fig.2f, Extended Data Figure 18) showing that L6 CT cells encode stimulus content at either a fast or $\sim 1\text{sec}$ timescale. To test explicitly whether this encoding persists at faster timescales, and after stimulus adaptation, and therefore could underlie the change representation in L2/3, we presented stimuli with amplitude deviations after 400ms, similar to those used for electrophysiology. As with baseline amplitude variations, L6 CT encoded these amplitude changes faithfully ($P < 0.001$, Fig. 1g, h). This encoding could represent either an explicit deviant encoding, as in L4, or a delayed stimulus or expectation encoding that emerges over timescales > 100 ms.

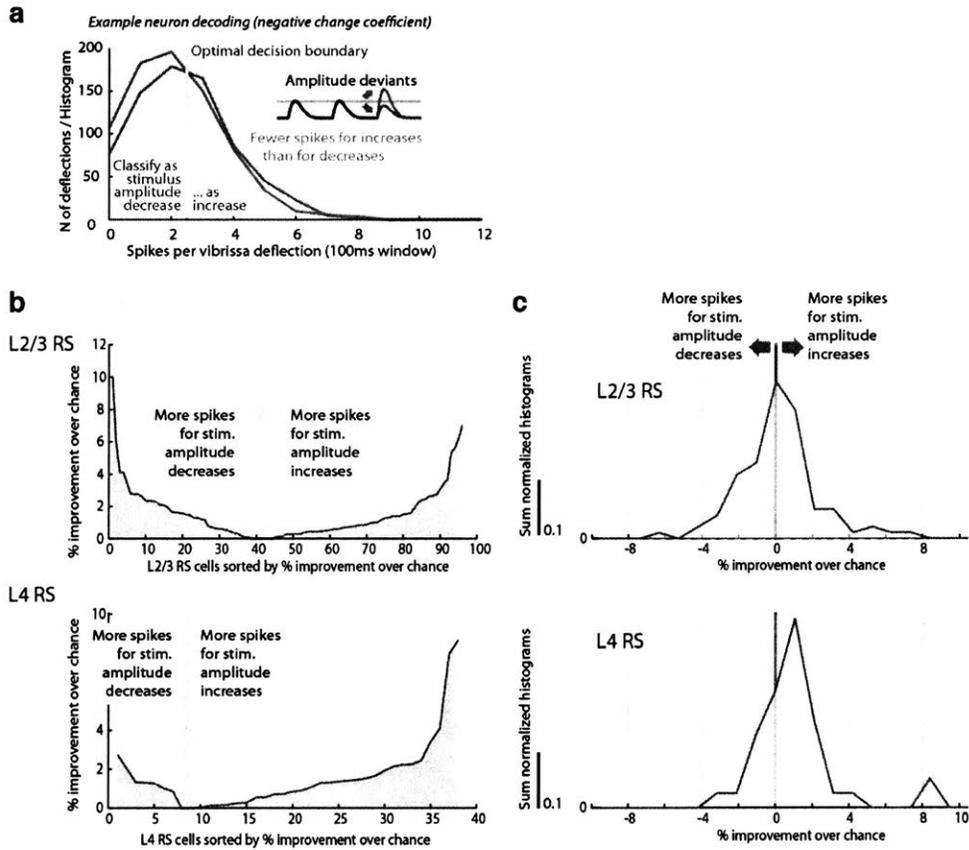


Figure 2-14: **L2/3 RS cells encode stimulus amplitude deviant identity heterogeneously, in contrast to L4.** L2/3 has an equal proportion of cells with increased or decrease firing probability for stimulus amplitude increases, in contrast to L4, which mirrors the same finding obtained by directly analyzing firing probability for increase/decrease deviants (Fig. 1). **a**, Example of ideal observer decoding for one cell. A optimal spike threshold is chosen as the value that results in the best classification rate (maximally different from 50%). Higher spike counts were always decoded as stimulus increases. For cells that decrease their firing rate for increased stimulus amplitudes or vice versa (i.e., negative change coefficient neurons), the correct rate would therefore be < 50%. A cell that fires on every trial with a significantly positive change coefficient could theoretically result in a 100% classification rate, while such a cell with negative change coefficient could yield an (equally informative) 0% hit rate. **b**, Rectified (always positive) correct classification rates relative to chance rate (which can slightly differ from 50% because of unequal proportions of stimulus increase versus decrease trials). In L4, most cells showed positive classification rates / spike count increases for stimulus increases, while in L2/3 the proportions are closer to even. **c**, Histogram of (non rectified) correct rates relative to chance level across cells the distribution of cells with positive or negative change encoding mirrors that of the change coefficients (Fig. 2-12).

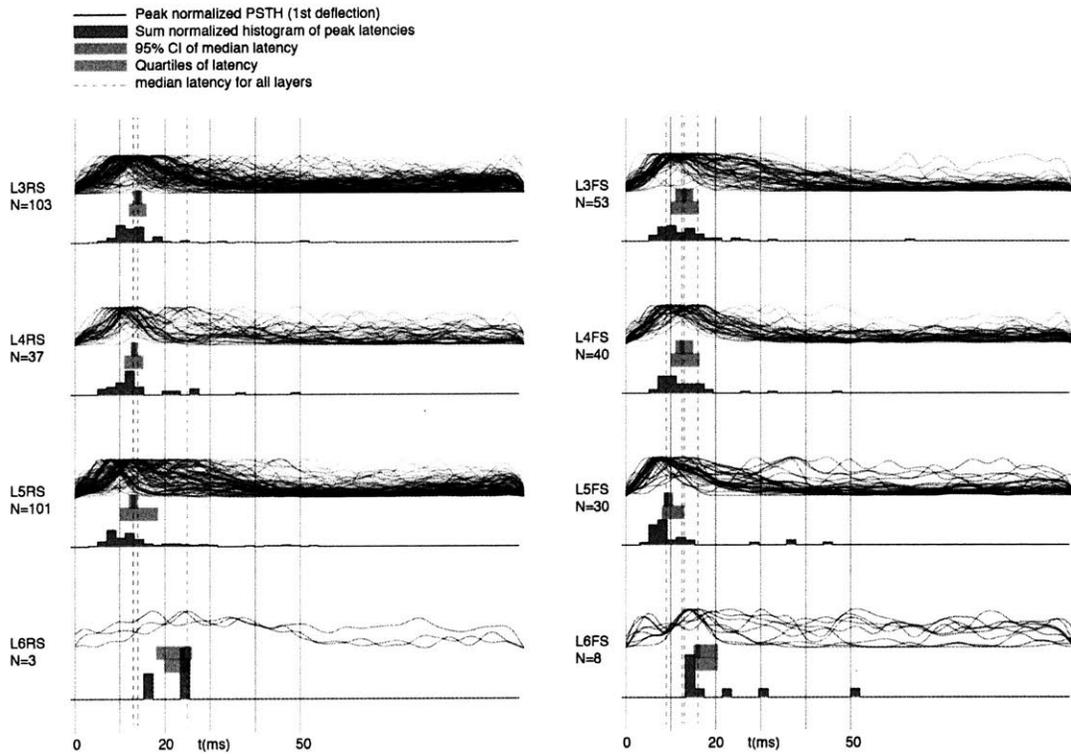


Figure 2-15: **Spike latency across layers for RS and FS cells.** PSTHs (0-100ms from stimulus onset) in the control condition (no laser) were smoothed with a gaussian filter ($\sigma=3\text{ms}$) and normalized to the peak amplitude. PSTHs with peaks after 60ms were excluded, and the peak times for the remaining neurons were analyzed. Bar graphs are either quartiles (green) or 95% CIs for the median. Red dotted lines show the median latency for each layer.

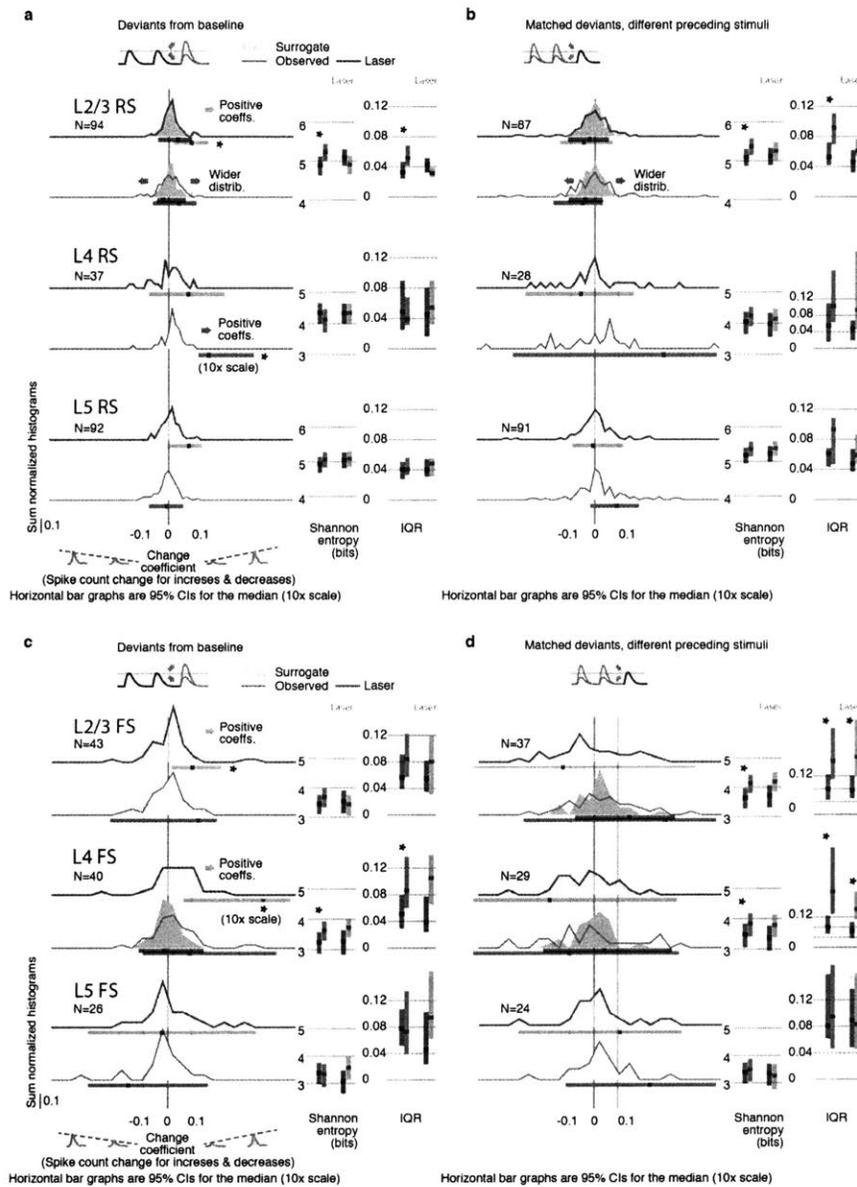


Figure 2-16: **Change coefficients for all layers, cell types and conditions.** **a**, CIs for RS cells. Layer 2/3 RS encoded deviations heterogeneously (median not different from zero, $P=0.395$, CI: [-0.007, 0.004, 0.011], see main text and Fig.1). In L4, stimulus increases cause RS to increase their firing rate relative to baseline stimuli ($P=0.025$ right tailed, rank-sum test, 95% CI of median firing probability change: [-0.003, 0.014, 0.028]). Stimulus decreases led to non-significant rate reductions ($P=0.155$ left tailed, CI: [-0.024, -0.012, 0.005]). Across both deviant conditions, L4 RS faithfully encoded deviants with positive change coefficients ($P=0.005$, CI: [0.012, 0.016, 0.035]). L5 RS had weakly positive but non-significant change-coefficients, and encoded neither stimulus deviants, nor stimulus history. **b**, Coefficients for history of stimulus deviations (matched amplitude in the deviant, with preceding stimuli of higher or lower amplitudes, Fig.2-17) for all regular spiking (RS) cells. **c**, Same as **a** but for fast spiking (FS) cells. Weak L6 drive made L4 FS change coefficients positive ($P=0.007$ individually, $P=0.006$ paired laser effect). **d**, Same as panel **b** but for FS cells.

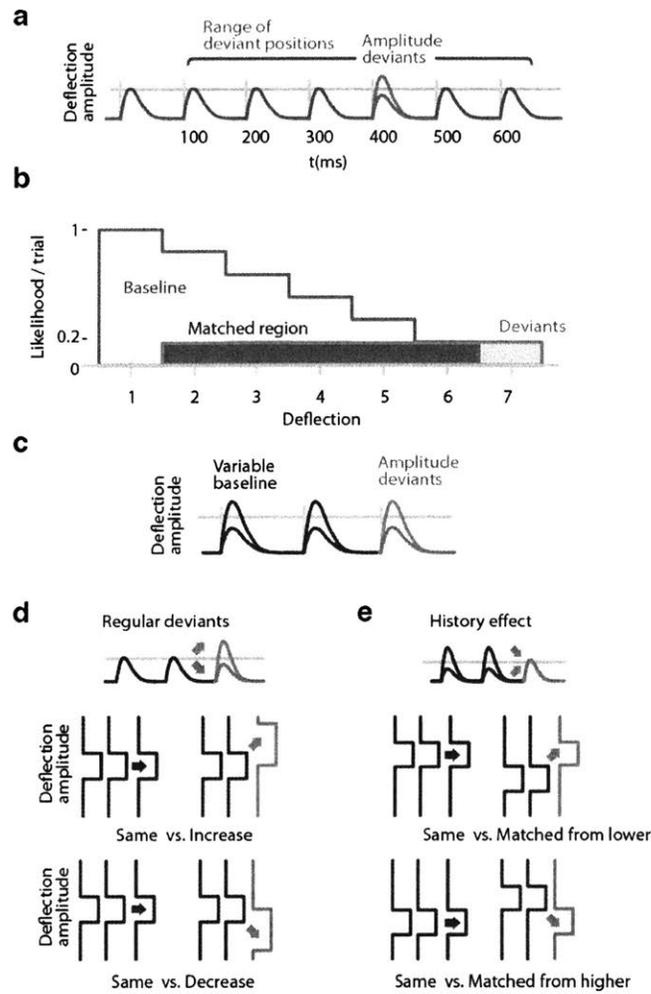


Figure 2-17: **Method for matching positions in stimulus trains and for analysis of encoding of stimulus history.** **a**, Possible positions of deviations in the stimulus train at deflections 2-7. **b**, Distributions of positions in trains of baseline (grey) and deviant (red) stimuli. Baseline stimuli are more likely to occur at the beginning of the train (post-deviant stimuli were excluded from analysis). Overlapping region (dark brown) shows the subset of trials that was sampled from to compute surrogate distributions to compare baseline and deviant stimuli while controlling for possible effects of different N and different positions in train. **c**, Stimulus design for sessions in which baseline amplitude was randomized across trials (N=55 sessions). **d**, Selection of trials for analysis of neural encoding of stimulus changes. Green: analyzed stimuli. Baseline stimuli are compared to increased or decreased stimuli. For analyses that are affected by adaptation or different sample sizes, positional matching (see a, b) was used. **e**, Selection of trials for analysis of stimulus history representation. Green: analyzed stimuli. Trials were sub-sampled to obtain conditions in which baseline amplitudes (grey) and deviant amplitudes (red/blue) were matched. The deviant condition was therefore defined by higher or lower preceding stimuli, allowing an analysis of the effects of stimulus history independent of the current stimulus amplitude. For analyses that are affected by adaptation or different sample sizes, positional matching (see a, b) was used in addition to the amplitude matching.

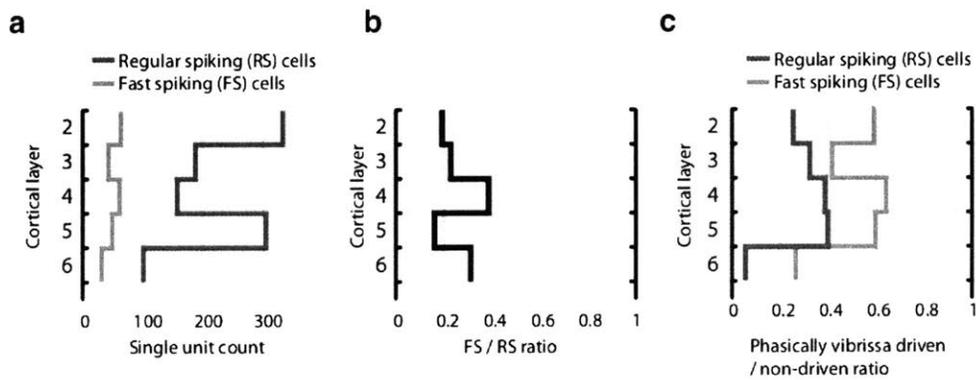


Figure 2-18: **Proportion of fast spiking (FS) and regular spiking (RS) neurons across cortical layers.** **a**, Total per-layer count of RS and FS cells. **b**, Proportion of FS to RS cells across all layers. **c**, Proportion of significantly vibrissa-stimulus driven (see Methods) RS and FS neurons across layers. The proportion of stimulus-driven cells is biased by our attempts to record from stimulus-driven cells: tetropdes were adjusted when recording conditions had stabilized but no stimulus-driven activity was recorded. Therefore, our data set likely significantly over-represents stimulus driven cells.

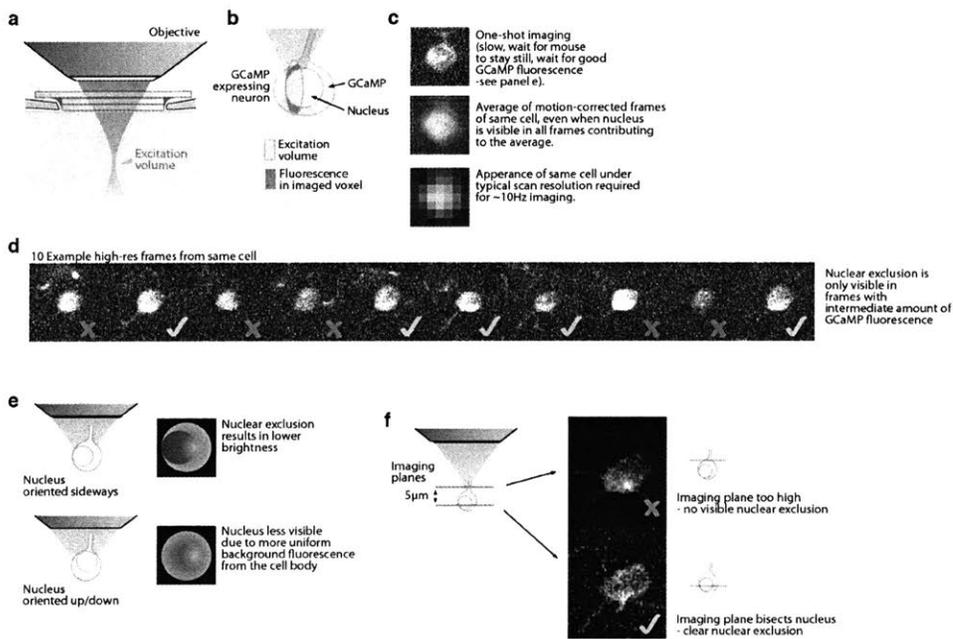


Figure 2-19: Visibility of cell nuclei in deep 2-photon imaging using GCaMP. **a**, Illustration of the imaging preparation using two stacked cover slips under a bigger top cover slip. **b**, Illustration of the 2-photon excitation volume in relation to a cell body. The low z-resolution at high imaging depths (Diaspro et al., 2006) reduces the apparent contrast of the cell nucleus. **c**, Demonstration of the effect of imperfect x/y motion correction and of the low scan resolution (typically $>2\mu\text{m}/\text{pixel}$, to achieve high frame rates of $>5\text{Hz}$) used in functional imaging on cell nucleus visibility. **d** Example frames from high-resolution scan of the same cell showing how z-motion and changes in fluorescent signal affect cell nucleus visibility. **e**, Illustration of how cell body shape can affect nucleus visibility. **f**, Illustration and example of how z-plane position relative to the cell can result in good calcium imaging data, but not necessarily show clear nuclear exclusion.

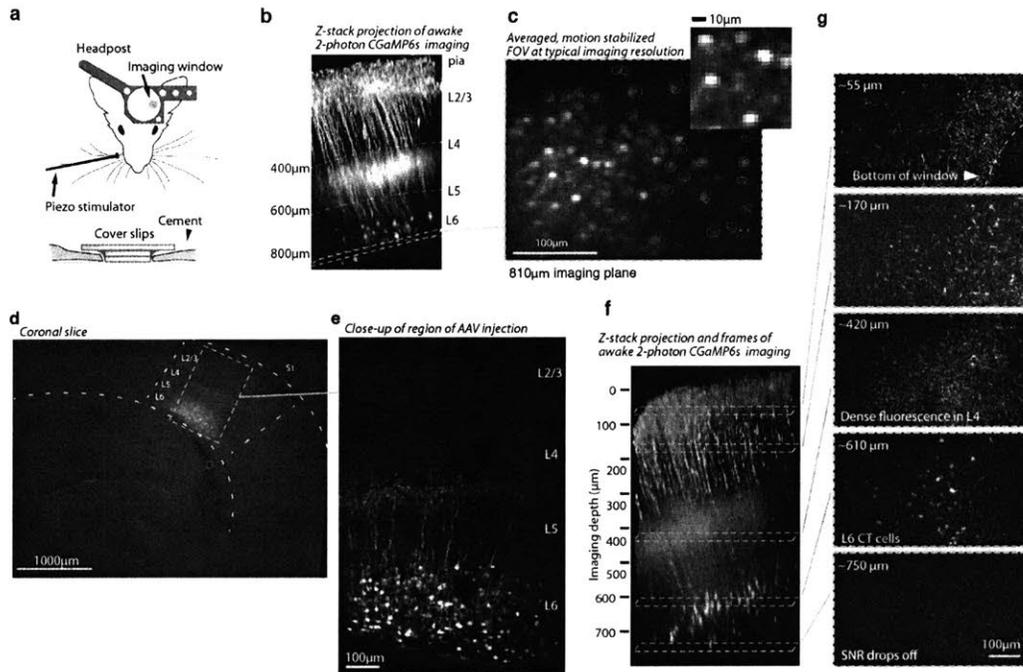


Figure 2-20: Histology and 2-photon imaging in L6 CT cells with GCaMP6s in the NTSR1-Cre line. **a**, Illustration of the imaging setup using two stacked cover slips under a bigger top cover slip. **b**, Z-stack projection of GCaMP6s image stack acquired in awake animal. **c**, Averaged, motion corrected field of view from typical L6 imaging session. **d**, Coronal section (150µm thickness) imaged at 525nm showing AAV mediated expression of GCaMP6s in NTSR-1 positive L6 CT neurons. **e**, Close-up of same slice (collage of 2 images), acquired with 2-photon scanning. **f**, Z-stack assembled from 170 images at 5µm steps, shown slightly oblique (~10 degrees) to the imaging Z-axis, maximum intensity projection. The stack was acquired in a resting, awake animal. Overall features of the image are identical to panel a and b, but superficial fluorescence appears brighter due to the strong image intensity falloff at greater depths in the 2-photon z-stacks. Beady appearance of processes in superficial layers is due to intermittent calcium driven fluorescence that makes individual dendrites/axons show up as very bright, large spots in individual frames (see panel e, 2nd example frame). Note that the relatively weak and sparse fluorescence above L4 is much more visible here than in the slice. This lack of superficial fluorescence is likely one of the features of the NTSR1 line that make this deep 2-photon imaging possible (Theer et al., 2003). **g**, Individual frames from Z-stack in panel f each image is remapped to equalize contrast and brightness and slightly smoothed to improve visibility (gaussian filter, $\sigma = 0.3\text{px}$).

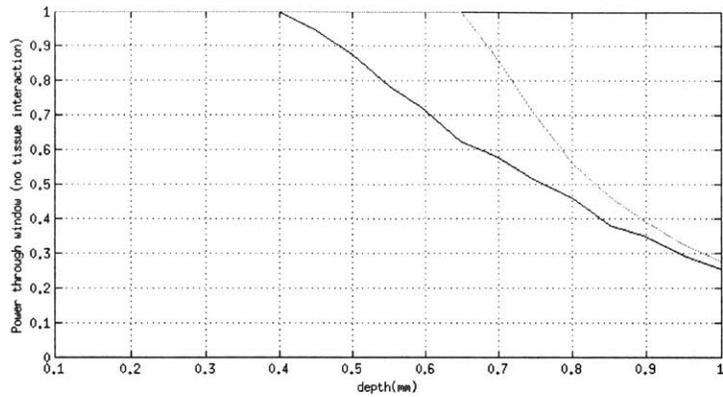


Figure 2-21: **Effect of window diameter and ROI position for deep imaging.** Available 2-photon excitation power (squared, not accounting for any scattering or absorption) for a completely uniformly filled 0.8NA objective through a 1mm window. The plot shows the squared fraction of photons that are not cut off by the window, for imaging in the center of the window (red), or 200 μm of center (black)

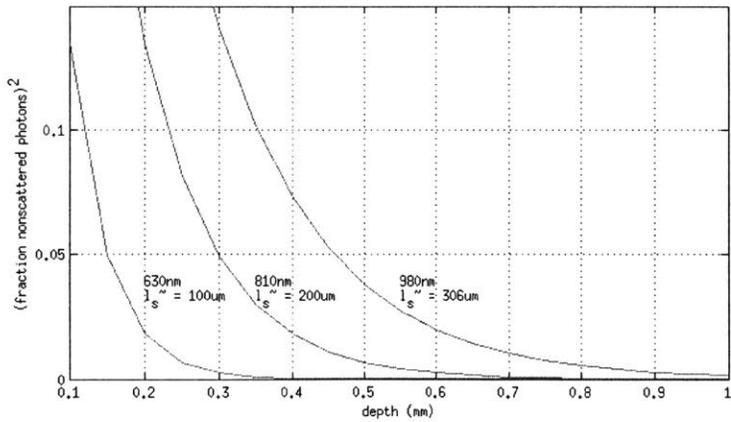


Figure 2-22: **Lambert-Beer exponential decay of non-scattered photons by depth for a few wavelengths.** All mean free path length estimates are approximations, the literature is not fully consistent on the numbers, so the values will not match specific setups.

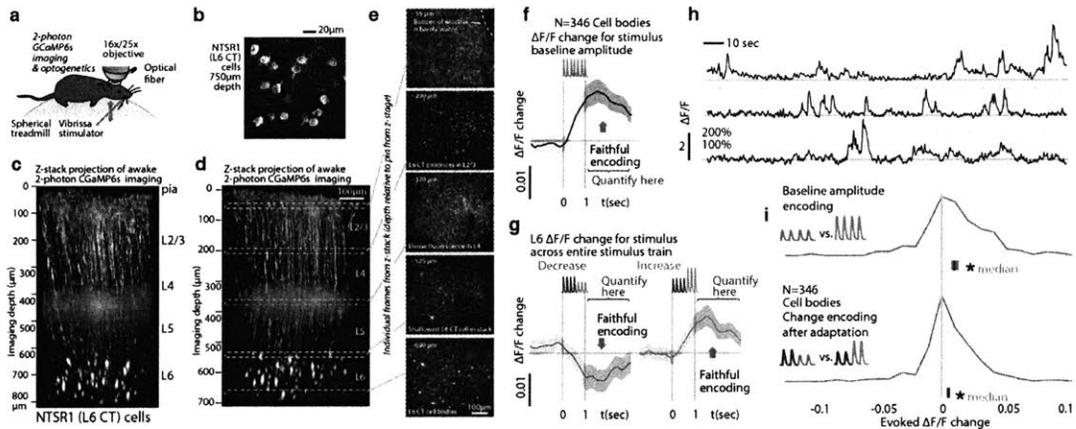


Figure 2-23: Layer 6 corticothalamic (CT) pyramidal cells faithfully encode stimulus amplitudes. **a**, Awake L6 2-photon calcium imaging with GCaMP6s in the L6 CT cell specific NTSR1-Cre line (See Methods for details). **b**, Example close-up image of L6 CT cells. To show visible cells, the image was combined from 12 individual frames in each of which a subset of the cells were spiking. **c,d**, Z-stack projections from awake mice. **e**, Sample frames from one of the z-stacks. **f**, L6 CT cells, though sparsely active, faithfully encode stimulus amplitude, **g**, even for amplitude deviations that occur later in the stimulus train (data were acquired at rates of down to ~ 5 Hz, so no attempt was made to quantify onset of any changes, and only the post-stimulus window was analyzed). **h**, Example traces for 3 neurons. **i**, Population data and summary statistics for L6 CT amplitude and change encoding.

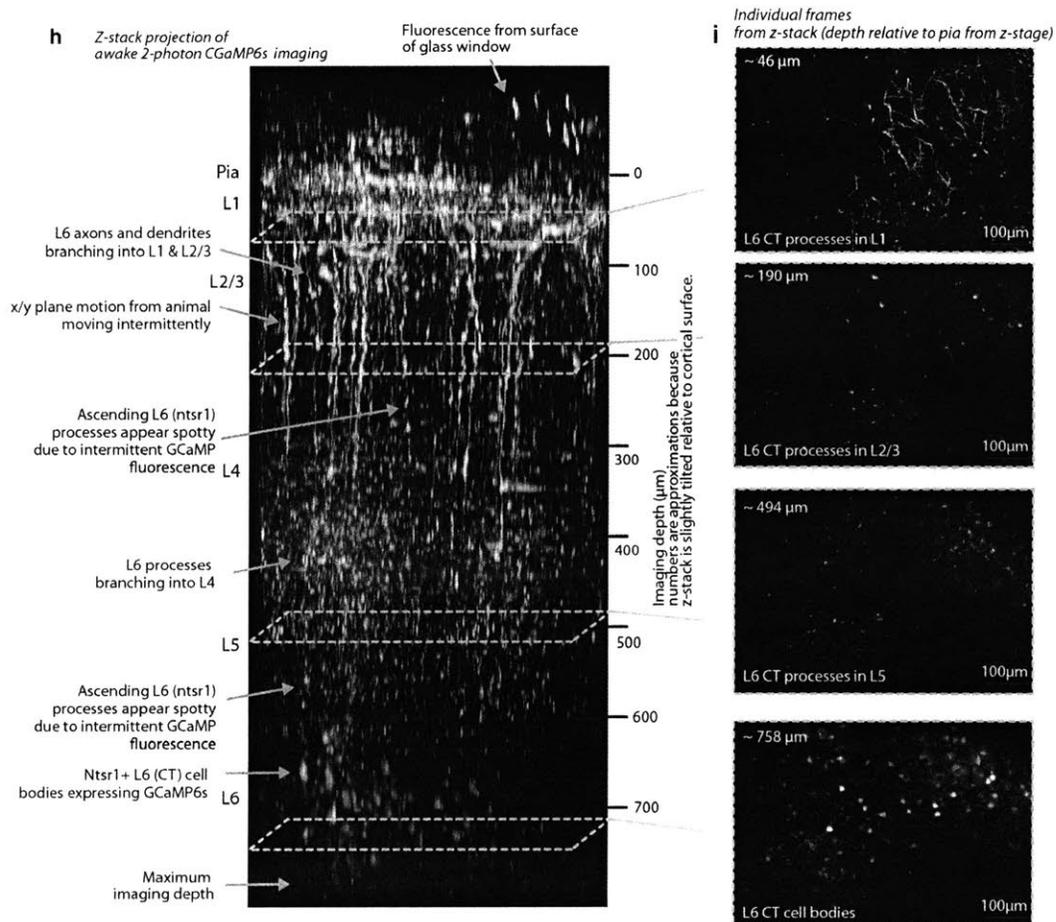


Figure 2-24: **Histology and 2-photon imaging in L6 CT cells with GCaMP6s in the NTSR1-Cre line.** **h**, Z-stack, same procedure as in **c**, assembled from 165 images at 5m steps. **i**, Individual images from the z-stack in panel **h**.

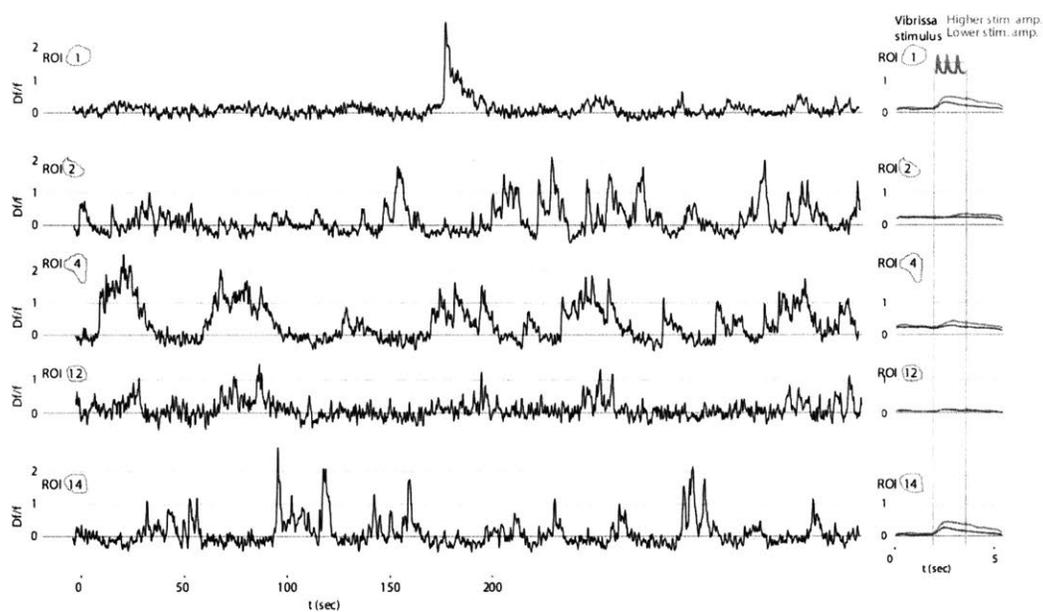


Figure 2-25: **Example fluorescence $\Delta F/F$ traces.** The traces were acquired at ~ 5 Hz, and smoothed with a gaussian filter ($\sigma = .15$ sec). **right**, Mean triggered average responses (95% CIs via bootstrap) of the cells to a 1 sec vibrissa stimulation. Red/Blue traces show responses to stimulation at $\sim 120\%$ or 80% of a mean stimulus amplitude. Same y-scale as in the left panel.

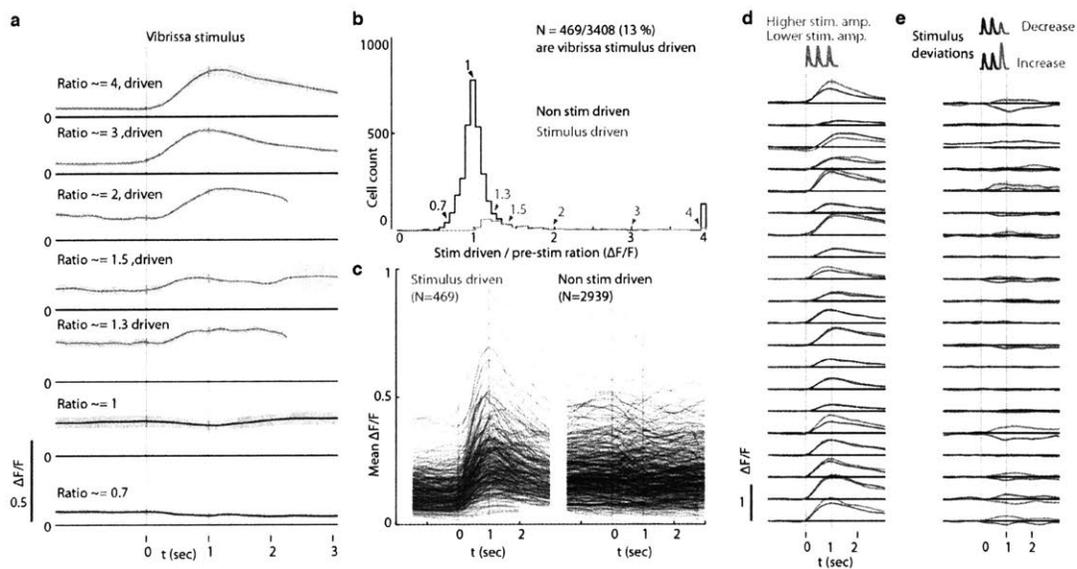


Figure 2-26: **Classification of L6 CT cells as vibrissa-stimulus driven.** **a**, Responses (stimulus triggered mean $\Delta F/F$) for 7 example ROIs showing approximate stimulus drive ratios of 0.7, 1, 1.3, 1.5, 2, 3, and 4. Stimulus drive amplitude was quantified as the ratio of the $\Delta F/F$ in a response window (0-2 seconds after stimulus offset) divided by the baseline $\Delta F/F$ (-2 to 0 sec). Confidence bounds are 95% for mean via bootstrap. **b**, Histogram of stimulus drive ratios for significantly vibrissa stimulus-driven (red) and non-driven (black) ROIs. ROIs were labeled as significantly stimulus driven if the 25th percentile of the mean $\Delta F/F$ in the response window was larger than the 95% percentile of the pre-stimulus period. **c**, Mean responses for all stimulus-driven, and the first 500 non-significant ROIs. **d**, Mean responses of highly stimulus-driven cells for high (red) and low (blue) ($\sim 120\%$ or 80% of a mean stimulus amplitude) non-deviant vibrissa stimuli. **e**, Mean difference between constant amplitude vibrissa stimuli and stimuli containing stimulus increase deviants (red) or decrease deviants (blue). Deviants were in the 80%-120% of baseline range.

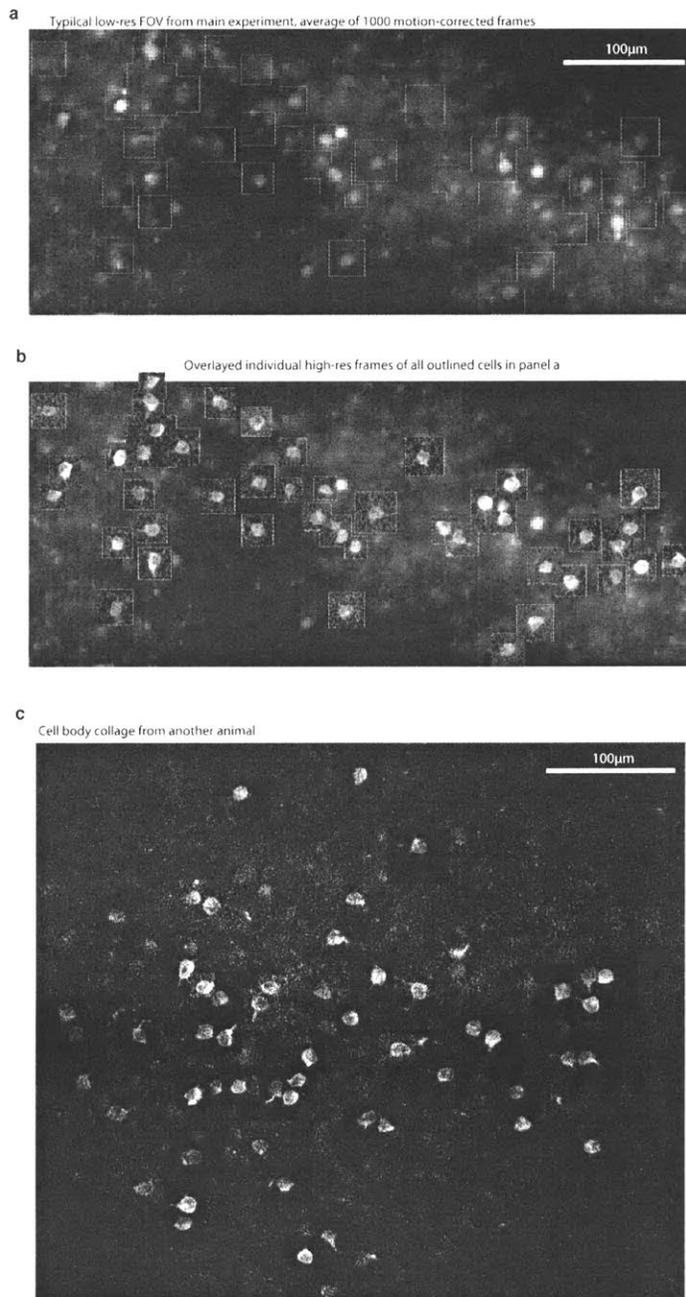


Figure 2-27: **Example high-resolution collages of L6 CT cells.** **a**, Typical Field of view (FOV) from a calcium imaging experiment using GcaMPs in NTSR1 positive L6 CT cells. This image shows the average of 1000 x/y motion corrected frames at a low ($\sim 2.5\mu\text{m}/\text{pixel}$) imaging resolution. Cell nuclei are not visible. **b**, Same FOV as in **a**, but overlaid with individual higher resolution scans from a subset of images. Each cell tile was selected to show an image when that cell was fluorescent. Cell bodies are visible in a large majority of cells. **c**, Same as panel **b**, from another animal.

Bibliography

- Andermann, M. L., Kerlin, A. M., Roumis, D. K., Glickfeld, L. L., and Reid, R. C. (2011). Functional specialization of mouse higher visual cortical areas. *Neuron*, 72(6):1025–1039.
- Andermann, M. L. and Moore, C. I. (2006). A somatotopic map of vibrissa motion direction within a barrel column. *Nature Neuroscience*, 9(4):543–551.
- Andermann, M. L., Ritt, J., Neimark, M. A., and Moore, C. I. (2004). Neural correlates of vibrissa resonance; band-pass and somatotopic representation of high-frequency stimuli. *Neuron*, 42(3):451–463.
- Anikeeva, P., Andalman, A. S., Witten, I., Warden, M., Goshen, I., Grosenick, L., Gunaydin, L. A., Frank, L. M., and Deisseroth, K. (2012). Optetrode: A multichannel readout for optogenetic control in freely moving mice. *Nature Neuroscience*, 15(1):163–170.
- Battaglia, F. P., Kalenscher, T., Cabral, H., Winkel, J., Bos, J., Manuputy, R., van Lieshout, T., Pinkse, F., Beukers, H., and Pennartz, C. (2009). The Lantern: An ultra-light micro-drive for multi-tetrode recordings in mice and other small animals. *Journal of Neuroscience Methods*, 178(2):291–300.
- Beaupreire, E. and Mertz, J. (2002). Epifluorescence collection in two-photon microscopy. *Applied Optics*, 41(25):5376.
- Biran, R., Martin, D. C., and Tresco, P. A. (2007). The brain tissue response to implanted silicon microelectrode arrays is increased when the device is tethered to the skull. *Journal of Biomedical Materials Research. Part A*, 82(1):169–178.
- Bonin, V., Histed, M. H., Yurgenson, S., and Reid, R. C. (2011). Local diversity and fine-scale organization of receptive fields in mouse visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 31(50):18506–18521.
- Bortone, D. S., Olsen, S. R., and Scanziani, M. (2014). Translaminar Inhibitory Cells Recruited by Layer 6 Corticothalamic Neurons Suppress Visual Cortex. *Neuron*, 82(2):474–485.
- Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G., and Deisseroth, K. (2005). Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neuroscience*, 8(9):1263–1268.
- Bragin, A., Hetke, J., Wilson, C. L., Anderson, D. J., Engel, J., and Buzsáki, G. (2000). Multiple site silicon-based probes for chronic recordings in freely moving rats: Implantation, recording and histological verification. *Journal of Neuroscience Methods*, 98(1):77–82.

- Buschman, T. J. and Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science (New York, N.Y.)*, 315(5820):1860–1862.
- Cardin, J. A., Carlén, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., Tsai, L.-H., and Moore, C. I. (2009). Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature*, 459(7247):663–667.
- Cardin, J. A., Carlén, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., Tsai, L.-H., and Moore, C. I. (2010). Targeted optogenetic stimulation and recording of neurons in vivo using cell-type-specific expression of Channelrhodopsin-2. *Nature Protocols*, 5(2):247–254.
- Castro-Alamancos, M. A. and Connors, B. W. (1996). Spatiotemporal properties of short-term plasticity sensorimotor thalamocortical pathways of the rat. *The Journal of Neuroscience*, 16(8):2767–2779.
- Celikel, T. and Sakmann, B. (2007). Sensory integration across space and in time for decision making in the somatosensory system of rodents. *Proceedings of the National Academy of Sciences*, 104(4):1395–1400.
- Chelazzi, L., Miller, E. K., Duncan, J., and Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, 363(6427):345–347.
- Chen, I.-W., Helmchen, F., and Lütcke, H. (2015). Specific Early and Late Oddball-Evoked Responses in Excitatory and Inhibitory Neurons of Mouse Auditory Cortex. *The Journal of Neuroscience*, 35(36):12560–12573.
- Chen, J. L., Carta, S., Soldado-Magraner, J., Schneider, B. L., and Helmchen, F. (2013). Behaviour-dependent recruitment of long-range projection neurons in somatosensory cortex. *Nature*, 499(7458):336–340.
- Chung, S., Li, X., and Nelson, S. B. (2002). Short-term depression at thalamocortical synapses contributes to rapid adaptation of cortical sensory responses in vivo. *Neuron*, 34(3):437–446.
- Cole, J. T., Yarnell, A., Kean, W. S., Gold, E., Lewis, B., Ren, M., McMullen, D. C., Jacobowitz, D. M., Pollard, H. B., O’Neill, J. T., Grunberg, N. E., Dalgard, C. L., Frank, J. A., and Watson, W. D. (2011). Craniotomy: True sham for traumatic brain injury, or a sham of a sham? *Journal of Neurotrauma*, 28(3):359–369.
- Crawley, J. N. (2007). *What’s Wrong With My Mouse?: Behavioral Phenotyping of Transgenic and Knockout Mice*. Wiley, 2nd edition.
- Denk, W., Strickler, J. H., and Webb, W. W. (1990). Two-photon laser scanning fluorescence microscopy. *Science*, 248(4951):73–76.
- Diaspro, A., Bianchini, P., Vicidomini, G., Faretta, M., Ramoino, P., and Usai, C. (2006). Multi-photon excitation microscopy. *BioMedical Engineering OnLine*, 5(1):36.
- Dickey, A. S., Suminski, A., Amit, Y., and Hatsopoulos, N. G. (2009). Single-unit stability using chronically implanted multielectrode arrays. *Journal of Neurophysiology*, 102(2):1331–1339.

- Dobbins, H. D., Marvit, P., Ji, Y., and Depireux, D. A. (2007). Chronically recording with a multi-electrode array device in the auditory cortex of an awake ferret. *Journal of Neuroscience Methods*, 161(1):101–111.
- 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(1):43–57.
- Fee, M. S. and Leonardo, A. (2001). Miniature motorized microdrive and commutator system for chronic neural recording in small animals. *Journal of Neuroscience Methods*, 112(2):83–94.
- Feingold, J., Desrochers, T. M., Fujii, N., Harlan, R., Tierney, P. L., Shimazu, H., Amemori, K.-I., and Graybiel, A. M. (2012). A system for recording neural activity chronically and simultaneously from multiple cortical and subcortical regions in nonhuman primates. *Journal of Neurophysiology*, 107(7):1979–1995.
- Freire, M. A. M., Morya, E., Faber, J., Santos, J. R., Guimaraes, J. S., Lemos, N. A. M., Sameshima, K., Pereira, A., Ribeiro, S., and Nicolelis, M. A. L. (2011). Comprehensive analysis of tissue preservation and recording quality from chronic multielectrode implants. *PloS One*, 6(11):e27554.
- Fujita, T., Yoshimine, T., Maruno, M., and Hayakawa, T. (1998). Cellular dynamics of macrophages and microglial cells in reaction to stab wounds in rat cerebral cortex. *Acta Neurochirurgica*, 140(3):275–279.
- Garion, L., Dubin, U., Rubin, Y., Khateb, M., Schiller, Y., Azouz, R., and Schiller, J. (2014). Texture coarseness responsive neurons and their mapping in layer 2-3 of the rat barrel cortex in vivo. *eLife*, 3:e03405.
- Goldey, G. J., Roumis, D. K., Glickfeld, L. L., Kerlin, A. M., Reid, R. C., Bonin, V., Schafer, D. P., and Andermann, M. L. (2014). Removable cranial windows for long-term imaging in awake mice. *Nature Protocols*, 9(11):2515–2538.
- 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. *The Journal of Neuroscience*, 27(37):9817–9823.
- Gray, C. M., Maldonado, P. E., Wilson, M., and McNaughton, B. (1995). Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. *Journal of Neuroscience Methods*, 63(1-2):43–54.
- Guizar-Sicairos, M., Thurman, S. T., and Fienup, J. R. (2008). Efficient subpixel image registration algorithms. *Optics Letters*, 33(2):156–158.
- Haiss, F., Butovas, S., and Schwarz, C. (2010). A miniaturized chronic microelectrode drive for awake behaving head restrained mice and rats. *Journal of Neuroscience Methods*, 187(1):67–72.
- Halassa, M. M., Siegle, J. H., Ritt, J. T., Ting, J. T., Feng, G., and Moore, C. I. (2011). Selective optical drive of thalamic reticular nucleus generates thalamic bursts and cortical spindles. *Nature Neuroscience*, 14(9):1118–1120.

- Helmchen, F. and Denk, W. (2005). Deep tissue two-photon microscopy. *Nature Methods*, 2(12):932–940.
- Huber, D., Gutnisky, D. A., Peron, S., O’Connor, D. H., Wiegert, J. S., Tian, L., Oertner, T. G., Looger, L. L., and Svoboda, K. (2012). Multiple dynamic representations in the motor cortex during sensorimotor learning. *Nature*, 484(7395):473–478.
- Jackson, N., Sridharan, A., Anand, S., Baker, M., Okandan, M., and Muthuswamy, J. (2010). Long-Term Neural Recordings Using MEMS Based Movable Microelectrodes in the Brain. *Frontiers in Neuroengineering*, 3:10.
- Jacques, S. L. (2013). Optical properties of biological tissues: A review. *Physics in Medicine and Biology*, 58(11):R37–61.
- Katz, Y., Heiss, J. E., and Lampl, I. (2006). Cross-whisker adaptation of neurons in the rat barrel cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 26(51):13363–13372.
- Khatri, V. and Simons, D. J. (2007). Angularly nonspecific response suppression in rat barrel cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, 17(3):599–609.
- Kloosterman, F., Davidson, T. J., Gomperts, S. N., Layton, S. P., Hale, G., Nguyen, D. P., and Wilson, M. A. (2009). Micro-drive array for chronic in vivo recording: Drive fabrication. *Journal of Visualized Experiments: JoVE*, (26).
- Knierim, J. J. and van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology*, 67(4):961–980.
- Korshunov, V. A. (2006). Miniature microdrive-headstage assembly for extracellular recording of neuronal activity with high-impedance electrodes in freely moving mice. *Journal of Neuroscience Methods*, 158(2):179–185.
- Kravitz, A. V. and Kreitzer, A. C. (2011). Optogenetic manipulation of neural circuitry in vivo. *Current Opinion in Neurobiology*, 21(3):433–439.
- Lansink, C. S., Bakker, M., Buster, W., Lankelma, J., van der Blom, R., Westdorp, R., Joosten, R. N. J. M. A., McNaughton, B. L., and Pennartz, C. M. A. (2007). A split microdrive for simultaneous multi-electrode recordings from two brain areas in awake small animals. *Journal of Neuroscience Methods*, 162(1-2):129–138.
- Lebedev, M. A. and Nicolelis, M. A. L. (2006). Brain-machine interfaces: Past, present and future. *Trends in Neurosciences*, 29(9):536–546.
- Lee, S., Carvell, G. E., and Simons, D. J. (2008). Motor modulation of afferent somatosensory circuits. *Nature Neuroscience*, 11(12):1430–1438.
- Maunsell, J. H. R. and Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, 29(6):317–322.
- McNaughton, B. L., O’Keefe, J., and Barnes, C. A. (1983). The stereotrode: A new technique for simultaneous isolation of several single units in the central nervous system from multiple unit records. *Journal of Neuroscience Methods*, 8(4):391–397.

- Miller, E. K. and Wilson, M. A. (2008). All my circuits: Using multiple electrodes to understand functioning neural networks. *Neuron*, 60(3):483–488.
- Mitchinson, B., Martin, C. J., Grant, R. A., and Prescott, T. J. (2007). Feedback control in active sensing: Rat exploratory whisking is modulated by environmental contact. *Proceedings. Biological Sciences / The Royal Society*, 274(1613):1035–1041.
- Moore, C. I., Nelson, S. B., and Sur, M. (1999). Dynamics of neuronal processing in rat somatosensory cortex. *Trends in Neurosciences*, 22(11):513–520.
- Movshon, J. A. and Lennie, P. (1979). Pattern-selective adaptation in visual cortical neurones. *Nature*, 278(5707):850–852.
- Müller, J. R., Metha, A. B., Krauskopf, J., and Lennie, P. (1999). Rapid adaptation in visual cortex to the structure of images. *Science (New York, N. Y.)*, 285(5432):1405–1408.
- Muthuswamy, J., Anand, S., and Sridharan, A. (2011). Adaptive movable neural interfaces for monitoring single neurons in the brain. *Frontiers in Neuroscience*, 5:94.
- Natan, R. G., Briguglio, J. J., Mwilambwe-Tshilobo, L., Jones, S. I., Aizenberg, M., Goldberg, E. M., and Geffen, M. N. (2015). Complementary control of sensory adaptation by two types of cortical interneurons. *eLife*, 4.
- Nguyen, D. P., Layton, S. P., Hale, G., Gomperts, S. N., Davidson, T. J., Kloosterman, F., and Wilson, M. A. (2009). Micro-drive array for chronic in vivo recording: Tetrode assembly. *Journal of Visualized Experiments: JoVE*, (26).
- Nicolelis, M. A., Lin, R. C., Woodward, D. J., and Chapin, J. K. (1993). Dynamic and distributed properties of many-neuron ensembles in the ventral posterior medial thalamus of awake rats. *Proceedings of the National Academy of Sciences of the United States of America*, 90(6):2212–2216.
- Nicolelis, M. A. L., Dimitrov, D., Carmena, J. M., Crist, R., Lehew, G., Kralik, J. D., and Wise, S. P. (2003). Chronic, multisite, multielectrode recordings in macaque monkeys. *Proceedings of the National Academy of Sciences of the United States of America*, 100(19):11041–11046.
- O’Connor, D. H., Clack, N. G., Huber, D., Komiyama, T., Myers, E. W., and Svoboda, K. (2010). Vibrissa-Based Object Localization in Head-Fixed Mice. *Journal of Neuroscience*, 30(5):1947–1967. WOS:000274246700037.
- O’Connor, D. H., Huber, D., and Svoboda, K. (2009). Reverse engineering the mouse brain. *Nature*, 461(7266):923–929.
- Ollerenshaw, D. R., Zheng, H. J. V., Millard, D. C., Wang, Q., and Stanley, G. B. (2014). The adaptive trade-off between detection and discrimination in cortical representations and behavior. *Neuron*, 81(5):1152–1164.
- Pastrana, E. (2011). Optogenetics: Controlling cell function with light. *Nature Methods*, 8(1):24–25.
- Pinault, D. (2004). The thalamic reticular nucleus: Structure, function and concept. *Brain Research. Brain Research Reviews*, 46(1):1–31.

- Polikov, V. S., Tresco, P. A., and Reichert, W. M. (2005). Response of brain tissue to chronically implanted neural electrodes. *Journal of Neuroscience Methods*, 148(1):1–18.
- Polley, D. B., Kvasnák, E., and Frostig, R. D. (2004). Naturalistic experience transforms sensory maps in the adult cortex of caged animals. *Nature*, 429(6987):67–71.
- Rennaker, R. L., Street, S., Ruyle, A. M., and Sloan, A. M. (2005). A comparison of chronic multi-channel cortical implantation techniques: Manual versus mechanical insertion. *Journal of Neuroscience Methods*, 142(2):169–176.
- Reynolds, J. H., Pasternak, T., and Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron*, 26(3):703–714.
- Ritt, J. T., Andermann, M. L., and Moore, C. I. (2008). Embodied information processing: Vibrissa mechanics and texture features shape micromotions in actively sensing rats. *Neuron*, 57(4):599–613.
- Scanziani, M. and Häusser, M. (2009). Electrophysiology in the age of light. *Nature*, 461(7266):930–939.
- Serruya, M. D., Hatsopoulos, N. G., Paninski, L., Fellows, M. R., and Donoghue, J. P. (2002). Instant neural control of a movement signal. *Nature*, 416(6877):141–142.
- Shein-Idelson, M., Ondracek, J. M., Liaw, H.-P., Reiter, S., and Laurent, G. (2016). Slow waves, sharp waves, ripples, and REM in sleeping dragons. *Science*, 352(6285):590–595.
- Siegle, J. H. and Wilson, M. A. (2014). Enhancement of encoding and retrieval functions through theta phase-specific manipulation of hippocampus. *eLife*, 3:e03061.
- Smith, A. C., Scalon, J. D., Wirth, S., Yanike, M., Suzuki, W. A., and Brown, E. N. (2010). State-space Algorithms for Estimating Spike Rate Functions. *Intell. Neuroscience*, 2010:3:1–3:14.
- Sobotka, S. and Ringo, J. L. (1994). Stimulus specific adaptation in excited but not in inhibited cells in inferotemporal cortex of macaque. *Brain Research*, 646(1):95–99.
- Swadlow, H. A. and Hicks, T. P. (1996). Somatosensory cortical efferent neurons of the awake rabbit: Latencies to activation via supra- and subthreshold receptive fields. *Journal of Neurophysiology*, 75(4):1753–1759.
- Szymanski, F. D., Garcia-Lazaro, J. A., and Schnupp, J. W. H. (2009). Current Source Density Profiles of Stimulus-Specific Adaptation in Rat Auditory Cortex. *Journal of Neurophysiology*, 102(3):1483–1490.
- Taaseh, N., Yaron, A., and Nelken, I. (2011). Stimulus-Specific Adaptation and Deviance Detection in the Rat Auditory Cortex. *PLoS ONE*, 6(8).
- Tannan, V., Whitsel, B. L., and Tommerdahl, M. A. (2006). Vibrotactile adaptation enhances spatial localization. *Brain Research*, 1102(1):109–116.
- Theer, P., Hasan, M. T., and Denk, W. (2003). Two-photon imaging to a depth of 1000 μm in living brains by use of a Ti:Al₂O₃ regenerative amplifier. *Optics Letters*, 28(12):1022.
- Thomson, A. M. (2010). Neocortical Layer 6, A Review. *Frontiers in Neuroanatomy*, 4.

- Towal, R. B. and Hartmann, M. J. (2006). Right-left asymmetries in the whisking behavior of rats anticipate head movements. *Journal of Neuroscience*, 26(34):8838–8846. WOS:000240006200022.
- Truccolo, W., Eden, U. T., Fellows, M. R., Donoghue, J. P., and Brown, E. N. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *Journal of Neurophysiology*, 93(2):1074–1089.
- Tseng, W.-T., Yen, C.-T., and Tsai, M.-L. (2011). A bundled microwire array for long-term chronic single-unit recording in deep brain regions of behaving rats. *Journal of Neuroscience Methods*, 201(2):368–376.
- Turner, J. N., Shain, W., Szarowski, D. H., Andersen, M., Martins, S., Isaacson, M., and Craighead, H. (1999). Cerebral astrocyte response to micromachined silicon implants. *Experimental Neurology*, 156(1):33–49.
- Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience*, 6(4):391–398.
- Vélez-Fort, M., Rousseau, C. V., Niedworok, C. J., Wickersham, I. R., Rancz, E. A., Brown, A. P. Y., Strom, M., and Margrie, T. W. (2014). The Stimulus Selectivity and Connectivity of Layer Six Principal Cells Reveals Cortical Microcircuits Underlying Visual Processing. *Neuron*, 83(6):1431–1443.
- Vetter, R. J., Williams, J. C., Hetke, J. F., Nunamaker, E. A., and Kipke, D. R. (2004). Chronic neural recording using silicon-substrate microelectrode arrays implanted in cerebral cortex. *IEEE transactions on bio-medical engineering*, 51(6):896–904.
- Voigts, J., Herman, D. H., and Celikel, T. (2015). Tactile object localization by anticipatory whisker motion. *Journal of neurophysiology*, 113(2):620–632.
- Voigts, J., Sakmann, B., and Celikel, T. (2008). Unsupervised whisker tracking in unrestrained behaving animals. *Journal of neurophysiology*, 100(1):504–515.
- Voigts, J., Siegle, J. H., Pritchett, D. L., and Moore, C. I. (2013). The flexDrive: An ultra-light implant for optical control and highly parallel chronic recording of neuronal ensembles in freely moving mice. *Frontiers in Systems Neuroscience*, 7.
- von Bekešy, G. (1967). *Sensory Inhibition*. Princeton University Press, Princeton.
- Wang, X. and Carlén, M. (2012). Optogenetic dissection of cortical information processing—shining light on schizophrenia. *Brain Research*, 1476:31–37.
- Williams, J. C., Rennaker, R. L., and Kipke, D. R. (1999). Long-term neural recording characteristics of wire microelectrode arrays implanted in cerebral cortex. *Brain Research. Brain Research Protocols*, 4(3):303–313.
- Wilson, M. A. and McNaughton, B. L. (1993). Dynamics of the hippocampal ensemble code for space. *Science (New York, N.Y.)*, 261(5124):1055–1058.
- Woods, E. J. and Frost, B. J. (1977). Adaptation and habituation characteristics of tectal neurons in the pigeon. *Experimental Brain Research*, 27(3-4):347–354.

- Yamamoto, J. and Wilson, M. A. (2008). Large-scale chronically implantable precision motorized microdrive array for freely behaving animals. *Journal of Neurophysiology*, 100(4):2430–2440.
- Zhang, F., Wang, L.-P., Brauner, M., Liewald, J. F., Kay, K., Watzke, N., Wood, P. G., Bamberg, E., Nagel, G., Gottschalk, A., and Deisseroth, K. (2007). Multimodal fast optical interrogation of neural circuitry. *Nature*, 446(7136):633–639.

Chapter 3

Optogenetic manipulation of stimulus representation in L6

This chapter describes findings on how a subtle optogenetic manipulation of layer 6 cells can lead to a disruption of the stimulus encoding in that layer, without affecting overall firing rates. I will then outline a number of experiments that employ this method to quantify the role of stimulus encoding in layer 6 in the encoding of stimulus changes in other cortical layers.

3.1 Methods

The majority of animals used for optogenetic experiments were the same as were used for the experiments described in Chapter 2, with the exception of mice used for acute recordings. Surgical procedures, and AAV injection for expression of ChR2 in L6 CT cells were performed as described before (Chapter 2, Methods).

3.1.1 Optogenetic stimulation

In all experiments, light was delivered through a jacketed fiber-optic cable 200 μ m in core diameter and 2.5 m long with a numerical aperture of 0.22 (Doric Lenses) connected to a 450nm diode laser (powertechnology.com) using a collimator (Thorlabs PAF-X-15-PC-A). The fiber was connected to the animal's head via mating metal ferrules in a zirconia sleeve. For head-fixed behaviour, ferrules were shielded with black plastic tape and the head of the mouse was illuminated with a blue masking LED that did not illuminate the stimulator or vibrissae. Light loss in the implanted fiber stub was measured for each implant. The amplitude of the light stimulus was calibrated regularly with an optical power meter (Thorlabs PM100D with S120C sensor) to up to 1mW at the surface of the skull, resulting in \sim 0.1–0.5mW in neocortex (measured through the skull and metabond after perfusion). In a subset of sessions, higher laser power (\sim 2–5mW, see Fig.3-2, or up to \sim 12mW for gap-crossing, Chapter 4, Fig.4-10) was used, or short pulses (4 or 40ms, \sim 12mW at fiber, 100 pulses/session) were delivered after the end of a recording session (Fig.3-1b). We note that the chronic implantation of an optical fiber results in a somewhat decreased power delivery to the brain due to inevitable regrowth of dura under the implantation site. Direct 1:1

comparisons of the light powers of the chronic experiments to the acute experiments where the fiber was placed directly on the brain (Fig.3-2) are therefore not possible, and it should be assumed that a somewhat higher light power is required in the chronic case to achieve the same extent of optogenetic activation as in the acute case. We estimate this light loss in the acute case to be around 50% but we did not implant enough mice or track the amount of light delivered to the brain with sufficient precision to quantitatively describe this effect.

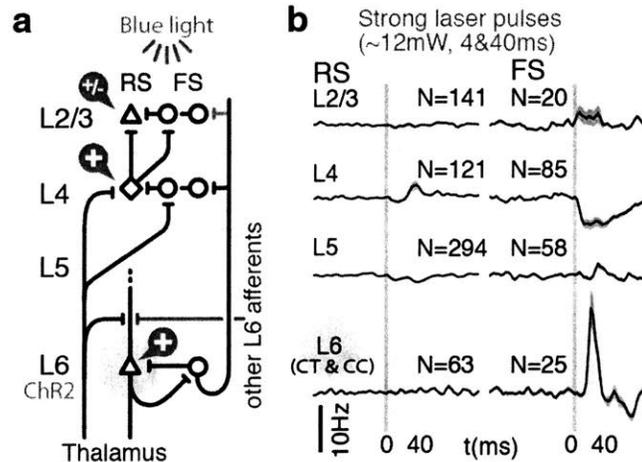


Figure 3-1: **Pathways by which L6 CT can affect other layers.** **a**, Simplified circuit diagram adapted from (Olsen et al., 2012; Helmstaedter et al., 2008). L6 CT cells represent stimulus amplitudes sparsely but faithfully ('+' Fig. 2-12g,h). **b**, Strong, short pulses of L6 CT activation robustly drive L6 FS (Bortone et al., 2014), and elicit a variety of mixed effects in superficial layers. Notably, we did not observe a direct activation of the (ChR2 expressing) L6 CT cells - possibly due to a mis-alignment of recording sites and sites of strongest ChR2 expression, or sparsity of activation.

3.2 Gain modulation by L6 CT cells

To test the functional role of the sparse stimulus encoding in L6 CT, we expressed Channelrhodopsin (ChR2) in the NTSR1-Cre line, and delivered blue light from a chronically implanted fiber on the surface of the brain, as described above.

We first replicated the effects of Olsen et al. (2012), who found that driving L6 CT cells in the same NTSR1 cre-lines results to increased firing rates results in a decrease of visually evoked responses in superficial layers. Similarly, strong suppression of the L6 cells resulted in an increase of visually evoked responses, showing the L6 CT cells can implement multiplicative/divisive gain modulation.

We replicated the L6 activation result from Olsen et al. (2012) in mouse S1 using acute laminar recordings (see methods). The experimental setup, stimulus design and delivery were identical to our awake experiments. Laser powers of >1mW were applied to the dura (corresponding to at least >2mW in the chronic setting where dura regrowth is expected after 2 weeks). As in visual cortex, strong L6 CT activation (>5mW power) resulted in increased L6 firing rates (both CT and others - we did not attempt to distinguish the

contribution of L6 CT from CC, and of L6 FS cells in this experiment). This activation recruited translaminar inhibition and reduced sensory gain in superficial layers (Fig.3-2a). We also observed that, possibly in addition to L6 FS driven translaminar mechanisms, FS cells throughout cortex could be activated by strong L6 CT drive. In a later experiment, we further validated this finding by showing that this gain modulation, even at moderate amplitudes, reduces the ability of mice to perform a freely behaving whisker-mediated sensory task (Fig.4-10).

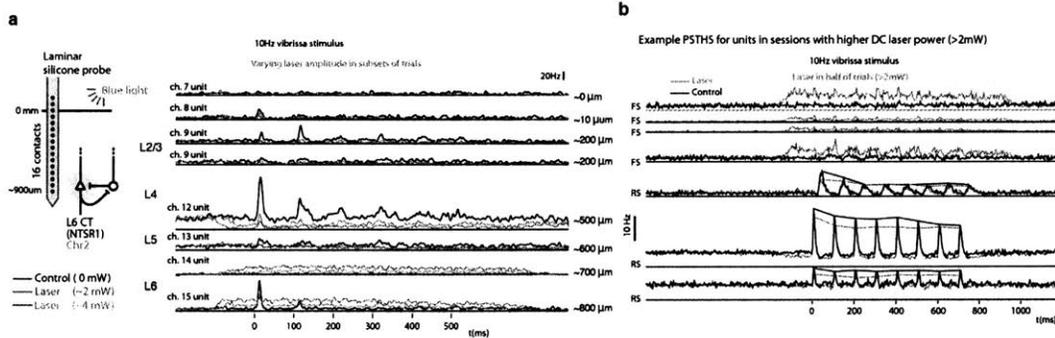


Figure 3-2: Strong optogenetic drive of Chr2 in L6 CT cells (NTSR1) reduces sensory gain across layers. **a**, Laminar silicon probe recording in anesthetized (isoflurane) mouse primary somatosensory mouse cortex. NTSR1+ L6 CT cells were transfected with Chr2 via AAV injection (see Methods). Blue light (~ 2 and ~ 4 mW of total power) was applied to the exposed cortex over the recording site and vibrissae were stimulated with short trains at 10 Hz, as in the main experiment. Many L6 cells were robustly driven by this laser power while neurons in layers 5 and above showed laser power dependent gain modulation (suppression, as in (Olsen et al., 2012)). **b**, Example PSTHs of FS and RS cells in L4 and L2/3 for large DC laser amplitudes (>2 mW) in awake mice. FS cells were driven, RS cells showed reduced gain.

3.2.1 Simultaneous 2-Photon Imaging and Optogenetic stimulation

In order to investigate the effect of optogenetic drive on L6 CT cells, especially in the case of the low amplitude drive employed for the most part of this thesis, a large number of reliably measured L6 CT cells was required. While the results in Chapter 2 showed that relatively good quality GCaMP imaging in L6 CT cells is possible, the integration of optogenetic drive and 2-photon imaging is challenging because the blue (450-470nm) light can easily mask any photons emitted by the GCaMP. In the basic case of just applying optigenetic stimulation to an unmodified 2-photon experiment, the light intensity is high enough to severely desensitize or even damage the photomultipliers (PMTs).

Here, to minimize light artifacts and PMT damage, we used a blocking filter (Semrock OD 6, custom NIR block, notches to block 460-470nm & 560-570nm). Light from blue (470 nm) or yellow (560 nm) LEDs was directed at the entire imaging area from a 200 μ m fiber at a ~ 40 degree angle. The LEDs were driven with a high-speed, high-precision LED driver (cyclops, designed by Jon Newman, www.open-ephys.org/cyclops), and was pulsed for

75 μ s after each 4th or 8th x-scan line. This stimulation scheme ensured that no visible light was directed at the brain while imaging data was acquired. Overall pulse rates were > 200 Hz, functionally equivalent to constant light (Lin et al., 2009). Light levels were adjusted manually to integrated powers of $\sim 0,1$ mW (for ChR2). Here, we did not pulse the PMT voltage to stop the PMTs from picking up the visible light pulses, but simply exploited the fact that the light induced artifact would decay sufficiently during the x-axis flyback. As a result, X-scan lines following LED pulses were slightly brighter due to the light stimulation. We removed this artifact by replacing these lines with interpolated data from preceding and following x-scan lines, whether the LED was on or off. Remaining slight image brightening was corrected off-line (see below).

3.2.2 2-Photon data analysis

Unless indicated, Wilcoxon rank sum or signed rank tests, and bootstrapping for testing IQRs were used, as described for electrophysiological data. Fluorescent values F were extracted from ROIs, the baseline fluorescence F_0 was computed as the 30th percentile in a 200 sec sliding window and $\Delta F/F$ was computed as $(F-F_0)/F_0$. Annulus-shaped ROIs were computed to estimate neuropil contamination (Kerlin et al., 2010; Bonin et al., 2011; Chen et al., 2013) by eroding out 20 pixels in the x-direction from each somatic ROI (this ensures that if there is any specific artifact from the pulsed optogenetic drive, it affects the cell body and neuropil ROIs equally) and excluding other cell bodies from this neuropil ROI (Kerlin et al., 2010; Bonin et al., 2011). For all analyses of firing rates, residual image brightening due to light artifacts was corrected by subtracting an average image brightening profile averaged from all neuropil ROIs over the entire session. All other analyses are computed as differences in evoked fluorescence between stimulus conditions within the same cell and laser condition, and were therefore not affected by the light artifact correction. Stimulus driven ROIs were identified by comparing the 90% quantiles of the $\Delta F/F$ for each ROI in the pre-stimulus period for all stimulus conditions (-1500 – 0 ms) with the 10% quantile in the stimulus period (500 – 1500ms), and ROIs with non-overlapping quantiles were analyzed further. Cells were classified either in control trials or optogenetic drive trials, or both (Fig.3-6, Fig.3-8). Change coefficients were defined as in the spiking data, but owing to the slow timescale of GCAMP6s, we analyzed the difference of the stimulus evoked fluorescence between baseline and deviant stimuli over the entire post-deviant stimulus time (0-2ec after stimulus offset) instead of analyzing individual deflections. (Fig.2-4, Extended Data Figs.11,16-18). Control levels for stimulus encoding were computed by shuffling of stimulus labels as described before.

3.3 Weak drive of L6 CT cells

To gain basic insight into the effect of L6 drive in our experiment, we first tested the responses to short, high-powered L6 CT stimulation. Strong pulses (4ms and 40ms duration, ~ 12 mW) activated L6 FS, directly via CT cells (Bortone et al., 2014), or via other pathways such as L6 CC cells (Zhang and Deschênes, 1997). However, we did not find robust activation of L6 RS using electrophysiology (Fig. 3-1, ~ 60 – 65% are expected to be NTSR-1+ cells (Vélez-Fort et al., 2014; Olsen et al., 2012; Zhang and Deschênes, 1998)). This could be due to the sparsity of the evoked L6 CT response, which nevertheless robustly

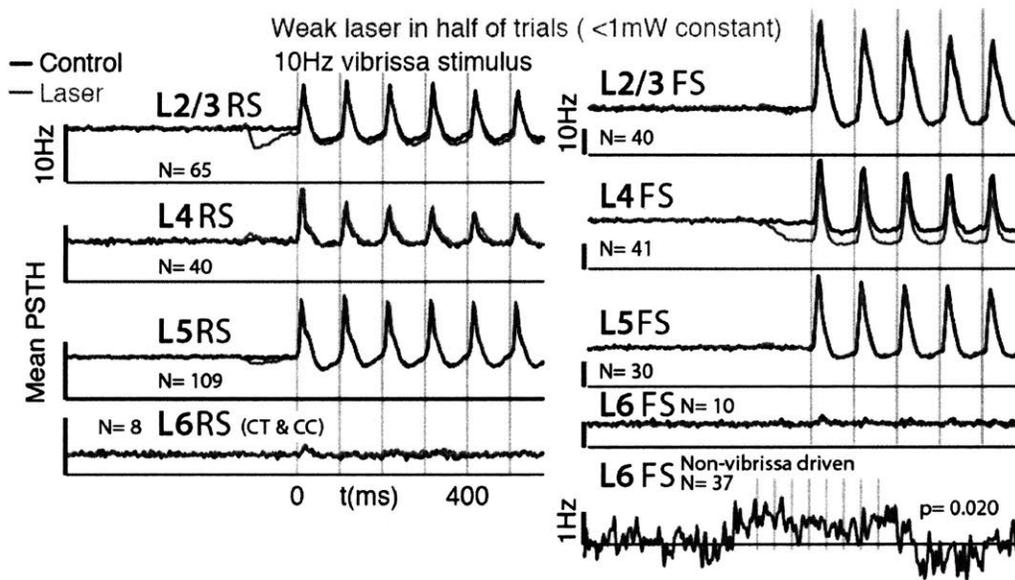


Figure 3-3: **Weak drive of L6 CT maintains overall firing rates in other layers a**, Weak drive suppresses only L4 FS, and weakly but significantly drives non vibrissa-stimulus driven L6 FS cells (green: laser effect, paired $P=0.02$ ranksum)

drives FS cells. L6 is typically very sparsely active, so it seems possible that a small number of highly synchronized L6 spikes could activate a much larger population of L6 FS cells. Another possibility is that the recording sites were not perfectly aligned with the site of L6 stimulation, allowing us to see the effects on nearby L6 FS cells, and in more superficial layers, while obscuring the L6 activation itself.

As discussed, strong, sustained L6 drive that increases or reduces L6 firing rates results in sensory gain modulation (Olsen et al., 2012; Bortone et al., 2014) (Fig.3-2). Here, I hypothesized that the sensory information, or expectation, or history thereof could be encoded by the L6 CT population. The results in Chapter 2 show that L6 CT cells encode stimulus direction (Fig.3-13), baseline amplitude, as well as small amplitude deviations (Fig.2-23), even though the sparse L6 responses and the slow timescale of the calcium imaging made it impossible to disambiguate these possibilities without extensive experiments using higher order stimulus manipulations. In order to test whether the L6 encoding has an effect on the stimulus change and history encoding in L2/3 (Fig.2-12), I wanted to avoid inducing gain modulation. A L6 manipulation that is strong enough to induce increased or decreased sensory evoked firing in L2/3 is by definition changing the responses to both repeated and novel stimuli, posing a confound for studying the circuitry underlying change detection. Further, attempting to validate any change-specific effects in behavioural experiments (Chapter 4) is impossible if even non-changing stimuli are suppressed or dis-inhibited by the manipulation.

To selectively probe the role of the observed sparse stimulus encoding in L6 CT cells on change encoding without causing overall gain modulation, we therefore employed significantly lower light levels ($\sim 0.1-0.5$ mW in tissue). This optogenetic stimulus modestly suppressed L4 FS (31.7 vs. 24.6 Hz median peak rates, $P=0.001$, Fig. 3c) but not L4 RS

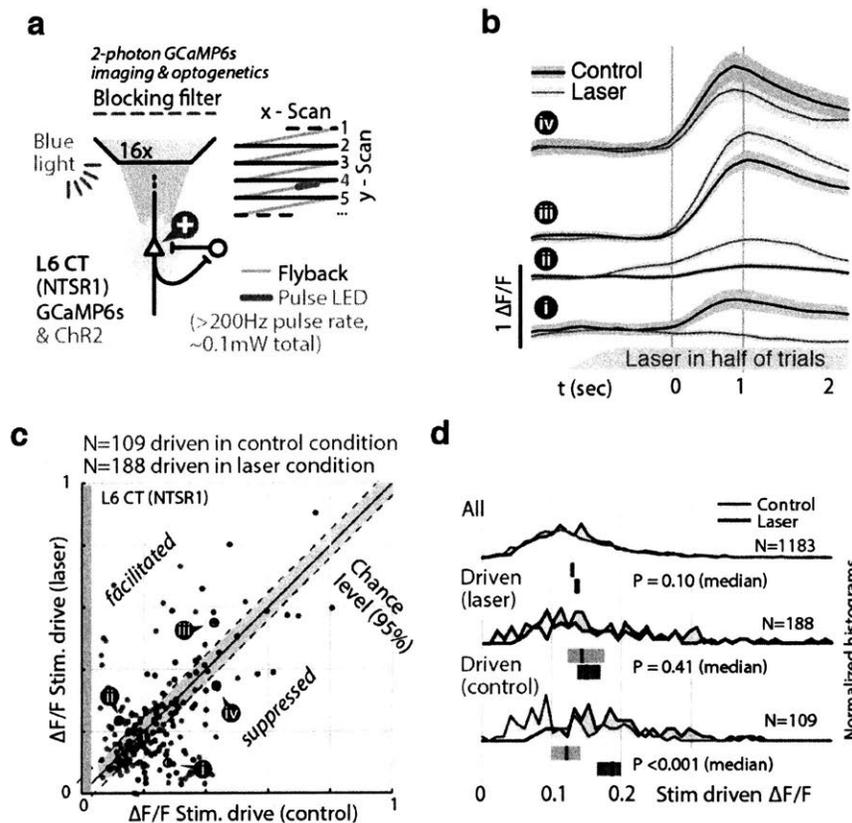


Figure 3-4: Weak drive of L6 CT maintains overall firing rates but changes stimulus driven populations in L6. **a**, Simultaneous calcium imaging and L6 optogenetic drive. **b,c**, Weak L6 CT drive can significantly facilitate or suppress L6 stimulus responses, changing which L6 cells are stimulus driven. ($P < 0.001$ IQR vs. shuffled control). **d**, Optical drive did not change overall L6 CT firing rates during sensory input, but did suppress the fraction of neurons selected to be stimulus driven in the control condition.

(15.5 vs. 16.9 Hz, $P = 0.20$) or L2/3 RS (14 Hz vs 13.9 Hz $P = 0.38$, all signed rank). Similarly to the strong pulsed drive (Fig.3-1), the weak sustained drive increased the firing rates of L6 FS interneurons (Bortone et al., 2014; Zhang and Deschênes, 1997) Fig.3-3, $P = 0.02$ sign-rank, laser - control firing rates, $N = 37$ non-vibrissa driven), this increased inhibition presumably mediates L6 CT suppression and offsets the direct optogenetic L6 CT drive, keeping the L6 RS firing rates at a baseline level ($P = 0.18$, $N = 89$, signed rank). This electrophysiological finding therefore shows that there is a regime of weak L6 CT drive in the NTSR1 line that activates L6 CT cells enough to in turn activate L6 FS cells, but keeps the overall L6 CT firing rates stable.

To better understand the effect of this weak manipulation on the population activity of L6 CT cells, we used simultaneous optogenetic drive and 2-photon calcium imaging (see methods above, $N = 1183$ cells, 4 mice, Fig. 3d). The electrophysiological findings of slight L6 FS drive but maintained L6 CT rates strongly suggest that subsets of L6 T cells (possibly the cells with higher AAV mediated Chr2 expression, or those receiving less FS inhibition)

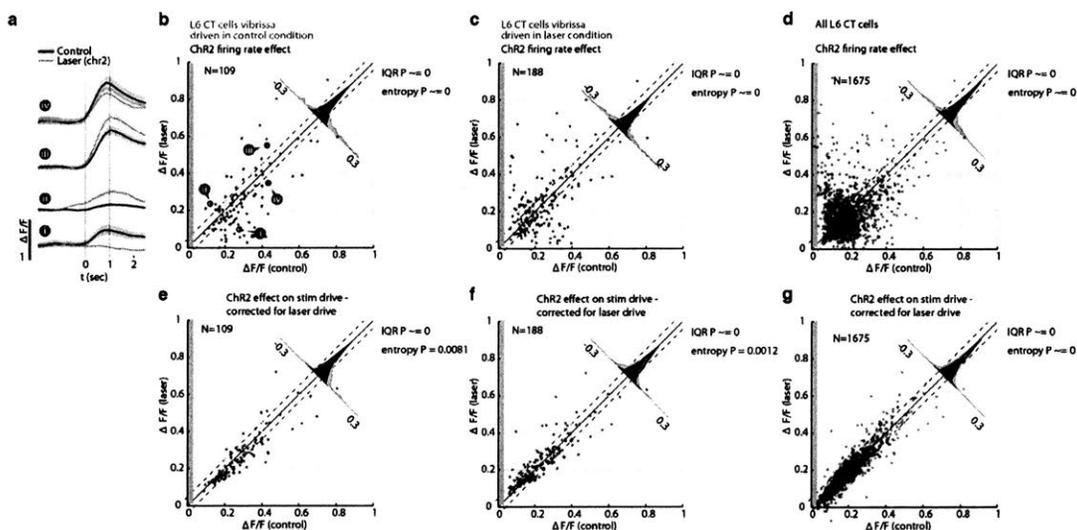


Figure 3-5: Weak optogenetic L6 CT drive suppresses and drives subsets of L6 CT cells. Effect of weak L6 CT ChR2 drive ($\sim 0.1\text{-}0.2\text{mW}$ through imaging window, Fig. 3-4 and Methods) on L6 CT cell firing rates during vibrissa stimulation, measured via calcium imaging via GCaMP6s in stimulus driven cells. **a**, Example responses of L6 CT to vibrissa stimulation in the control and laser conditions. Individual cells can show direct activation to light presentation, suppression of vibrissa evoked responses, or facilitation of vibrissa evoked responses. **b**, Per-cell $\Delta F/F$ for vibrissa stimuli in the control and laser conditions for cells that were classified as vibrissa stimulus driven in the control condition (Extended Data Figure 11). Diagonal insert: sum-normalized histogram of the paired change per cell (blue). The control distribution (black) and 95% significance bounds/chance level (black dotted lines) were calculated by shuffling the laser activation labels across all trials for each cell. Same as panel a but cells were classified as vibrissa stimulus driven in the laser condition. **d**, Same as panels a, b but for all L6 CT cells regardless of vibrissa stimulus drive. **e-g**, Same as panels b, c, d but for each cell the mean effect of the optogenetic drive in absence of vibrissa stimulation (difference between laser and control activation in catch / 0-amplitude trials) was subtracted prior to analysis, resulting as an (additively corrected) estimate of the optogenetically induced change in purely vibrissal-driven firing rate.

should be driven to slightly increased rates, while others (with less Chr2 drive, and/or more FS inhibition) should be suppressed.

We found that weak optogenetic input indeed drove or suppressed the sensory responses of individual L6 CT cells (Fig. 3-4b), presumably by generating inhibition through the activation of L6 FS by CT Chr2 drive. We quantified this effect by analyzing the changes in sensory evoked firing rates across L6 CT cells between the laser and control conditions (Fig. 3-4c), and comparing the differences to the differences obtained from randomly assigning laser or control trial labels (Grey dotted region: 95% confidence bound).

As in the electrophysiological data, mean CT calcium signals (relative firing rates) remained unchanged during laser drive given the balance of these effects across neurons ($P=0.10$, signed rank (Fig. 3-4d)), as in the electrophysiological data. In order to exam-

ine whether the effect was limited to a suppression of cells that were classified as stimulus driven in the control condition and suppressed in the laser condition or vice-versa, we analyzed the firing rate and ‘shuffling’ effect separately for all cells, for cells that are classified as stimulus-driven in control trials, or as stimulus driven in laser trials (Fig. 3-4d)). As expected by a random ‘re-shuffling’ of the population, cells that were stimulus driven in the control trials were less driven in laser trials. However, cells that were classified as driven in the laser condition were not less sensory driven in the control condition, indicating that these cells might be less specifically stimulus driven. We further repeated the scatter-plots and analysis of the changes of sensory evoked firing in the laser condition minus control condition versus a shuffled null distribution in data that were corrected for the effect of optogenetic stimulation in absence of vibrissa stimuli by subtracting out the laser evoked response in catch trials (Fig.3-5).

3.4 Effect on stimulus encoding in L6

Even though mean L6 CT population rates were not significantly changed, the weak L6 drive changed which cells were activated by vibrissa stimulation (Fig.3-4, Fig.3-5). To understand the impact of this shuffling on the downstream read-out of L6 information, we investigated if, and how the encoding of stimulus features was affected by this manipulation, using the same stimulus design as used for testing basic L6 encoding (Fig.2-23).

We found that the encoding of stimulus baseline amplitudes ($P < 0.05$, ranksum, evoked $\Delta F/F$ change between stimuli, control vs. laser) and of stimulus amplitude changes in an ongoing train was significantly disrupted ($P < 0.05$, Fig. 3-6).

This effect could be explained by the shuffling of stimulus driven L6 populations (Fig. 3-4c,d): Sensory responses for cells that are selected to be stimulus driven in the control condition would be expected to be decreased in the laser condition due to any, possibly random changes in population activity. Indeed, this population carried less stimulus information ($P < 0.05$). However, we also found that encoding was affected in cells that were stimulus driven in the control condition and remained so the laser condition ($P=0.044$ rank sum across cells, $P<0.0001$ across trials, Fig.3-7), showing that even for cells that retained significant stimulus-driven firing, the information content of the firing rate was reduced. This leaves two non-exclusive main mechanism hypotheses on how this change in encoding could affect the downstream read-out: The downstream area could receive information from a new population of L6 cells that are active in the laser condition which could affect interpretation of the neural code because the synaptic weights of these cells onto their targets were not subject to learning and plasticity, and/or the L6 population that is active carries less information, so regardless of the interpretation of the downstream neural population, there is less information available.

All analyses of the effect of weak L6 drive on stimulus encoding in L6 relied on the assignment of cells to stimulus driven or non stimulus driven categories (Fig.2-26). Even though we found consistent effects for cells that were stimulus driven in the laser condition as well as in the control condition, or both, this assignment procedure still classifies cells on the same data that is then used to quantify the stimulus encoding. To avoid any possible effects of the classification on the analysis, we quantified the stimulus encoding in a separate analysis that uses a leave-one-out (LOO) procedure (Fig.3-8). For each cell, and for each

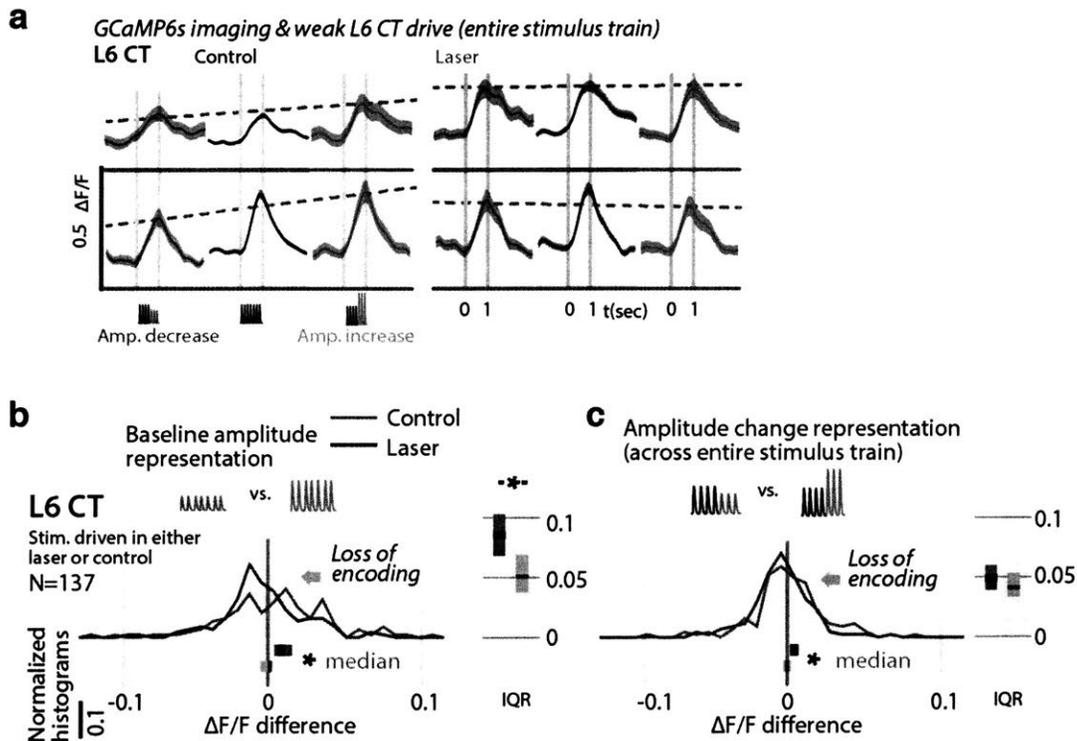


Figure 3-6: **Weak L6 drive disrupts stimulus encoding in L6** a, Example L6 CT stimulus responses. Weak L6 CT drive disrupts representation of stimulus amplitude for b, constant stimulus amplitude trains and, c, trains with stimulus amplitude changes.

trial, the cell was only counted as stimulus driven, if it was quantified as stimulus driven in all but the analyzed trial. Firing rates were then compared across stimulus conditions (large vs. small amplitude, amplitude changes etc.) across these trials. As before, classification of stimulus drive was restricted to control or laser condition trials.

Overall, the LOO procedure replicated the findings of the static class assignment analysis: L6 CT cells significantly encode stimulus amplitude (Fig.3-8a), and stimulus amplitude changes later within the stimulus train (Fig.3-8b,c). The same result applied for classification of stimulus drive in laser trials (Fig.3-8d-f). In all cases, the stimulus information was reduced in the laser condition, with the exception of encoding of stimulus changes for cells/trials classified in the laser condition (Fig.3-8f).

To summarize, weak L6 CT drive recruited L6 FS cells, kept L6 CT firing rates and all stimulus evoked firing rates unchanged, but suppressed some L6 cells and facilitated others. The information content of the L6 CT cells was significantly reduced by the manipulation.

3.5 Effect on stimulus encoding in other layers

We next examined whether and how this weak optogenetic drive of L6 CT impacted stimulus representation in other layers (Fig.2-12). The findings presented in Chapter 2 show

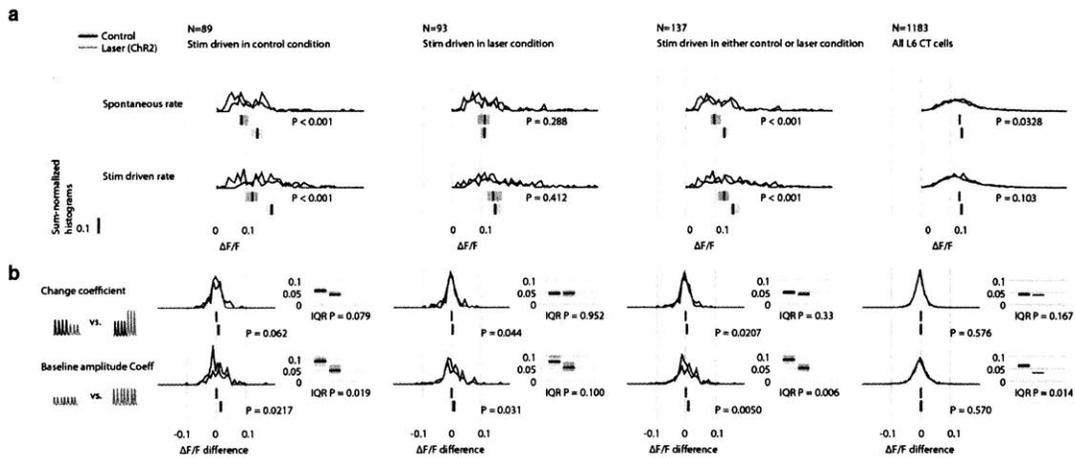


Figure 3-7: Weak optogenetic L6 CT drive with Chr2 disrupts encoding of small stimulus amplitude changes and deviants, quantified via raw $\Delta F/F$. Effect of weak L6 CT Chr2 drive (~ 0.1 - 0.2 mW through imaging window, see Fig. 2d) on L6 CT cell firing rates during vibrissa stimulation, measured by calcium imaging via GCaMP6s in stimulus driven cells. All vibrissa stimulus responses were quantified in a 0-2 sec window relative to stimulus offset (or equivalent time in catch trials). **a**, Histogram of spontaneous (catch trials) or vibrissa stimulus evoked mean $\Delta F/F$. Bar graphs show 95% CIs of the median per group via bootstrapping (1000 fold). P values are from two-sided Mann-Whitney rank sum tests between the control and laser conditions. Data is analyzed as all stimulus driven cells classified in control condition (left), classified in laser trials (middle left), in either condition (middle right), and for all L6 CT cells (far right). **b**, Histograms of change coefficient, computed as the difference in evoked $\Delta F/F$ between constant and increasing amplitude and constant and decreasing amplitude trials, and baseline amplitude encoding (mean difference in $\Delta F/F$ between low and high-amplitude constant stimuli). Analysis is split into groups as in panel a. P values are either from Mann-Whitney rank sum tests (unpaired) or Wilcoxon sign rank test (paired by cell). Spread of the distributions is quantified via IQR, CIs are computed via 1000 fold bootstrap, P values via rank sum test.

that in contrast to L4 pyramidal cells that encode stimulus amplitude deviations faithfully, L2/3 cells encode heterogeneous change signals that depend both on the current (Fig.2-12) as well as on the preceding (Fig.2-13) vibrissa deflections. If this transformation of RFs between L4 and L2/3 RS cells depends on the information that is encoded by L6, for instance by subtraction of the fast, faithful L4 encoding, and the slower L6 encoding, then the weak L6 drive that reduces the information L6 carries about the baseline stimulus, would be expected to have a specific impact on the L2/3 RFs.

In agreement with this prediction, we found that the weak L6 drive disrupted the heterogeneous change representation in L2/3 (Fig.3-9b, $P=0.010$ entropy reduction, $P=0.020$ IQR reduction, $P=0.008$ paired left tailed IQR, $N=94$). Instead of encoding heterogeneous change signals, during optogenetic drive of L6 CT, L2/3 neurons came to represent current stimulus amplitudes with net positive change coefficients (CI: [0.002, 0.009, 0.016], $P=0.003$

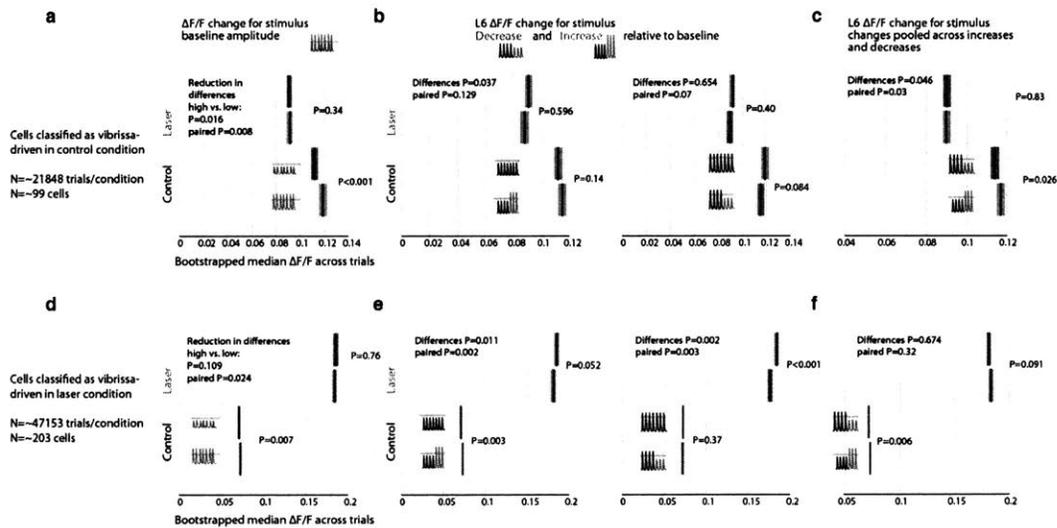


Figure 3-8: Weak optogenetic L6 CT drive with ChR2 disrupts encoding of small stimulus amplitude changes and deviants, quantified via leave-one-out cross-validation for stimulus driven vs. non-driven cells. Effect of weak L6 CT ChR2 drive on L6 CT cell firing rates during vibrissa stimulation, measured by calcium imaging in stimulus driven cells. **a**, To test whether the encoding of stimulus features in L6 CT cells was affected by weak L6 drive independently of a possible selection bias in choosing cells that are stimulus driven in the control condition and then possibly lose significance in the laser condition purely because of the variance in their responses, we used a leave-one-out (LOO) method to quantify L6 CT cell responses (also used for quantifying the directional tuning on L6 CT cells, see Fig.3-13). Instead of assigning cells as stimulus or not stimulus driven, we classified a cell as stimulus driven per trial, using all but one trial, and then analyzed this trial only if the other trials showed significant stimulus drive. This procedure avoids selecting data on the same dataset as is used for the analysis. In this analysis, positive change coefficients were observed (small stimuli led to smaller, bigger stimuli to bigger $\Delta F/F$). P values for comparisons within control or laser conditions are from rank sum tests. This difference is significantly reduced in the laser condition. P values for the decrease of this difference, in control versus laser trials are computed by bootstrapping across trials and testing the median of the difference (laser vs. control) of differences (large in control-small in control) (large in laser-small in laser) versus zero for each sample. For paired P values by cells, each bootstrap sample computed differences of the difference of mean responses as before, but within cells, for a bootstrapped sample of cells and then tested the median across cells versus zero. **b**, Comparison of $\Delta F/F$ for constant and changing stimuli for increasing (red) and decreasing (blue) trials. **c**, Same data as in panel d but pooled across increasing and decreasing deviant conditions. Deviants were significantly encoded and this encoding was disrupted in the laser condition. Because of the timescale of GCaMP6s, multiple interpretations exist for this encoding. This encoding may be determined by a delayed response to repeated presentation of the deviant amplitude and does not necessarily represent a true deviant encoding. **d,e,f**, Same as panels c, d, e but cells were classified as stimulus driven (via LOO procedure) in the laser condition. Baseline stimuli are significantly encoded in these cells, and this encoding is significantly disrupted in the laser condition. Stimulus deviants are encoded in the control condition but the encoding is not reduced as clearly in the laser condition.

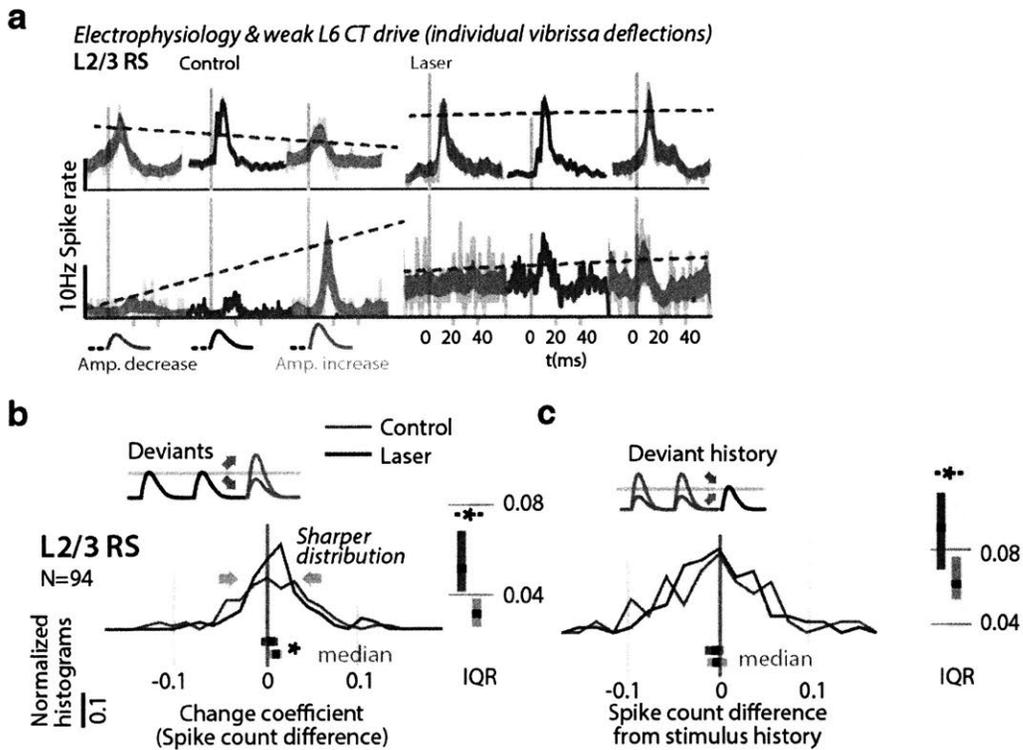


Figure 3-9: **Weak L6 drive disrupts change as well as history encoding in L2/3** a, Example L2/3 RS responses (as in Fig.1b). Weak L6 drive made L2/3 change coefficients positive and, b, reduces their heterogeneity as well as, c, stimulus history representation.

signed rank), similar in principle to the code in L4 RS. History-dependence was similarly reduced (Fig.3-9c, $P=0.014$ entropy reduction, $P=0.004$ IQR reduction, $P=0.043$ paired left tailed IQR), indicating that sparse stimulus encoding in L6 contributes to heterogeneous change representations in superficial layers.

In addition to examining the effect of weak L6 drive on the highly laminar-specific encoding of stimulus amplitude changes, we also examined more classical stimulus-specific adaptation (SSA) effects. In a subset of experiments, either in the electrophysiology experiments ($N=92$ units and MUAs) or 2-photon L6 CT imaging experiments ($N=84$ cells), stimuli were delivered as in the other experiments, but instead of amplitude deviations, the direction of the whisker stimulus was changed at some point in the stimulus train (see also Fig. 2-7).

Vibrissa deflection direction was chosen randomly, and for each cell the preferred direction was computed from all but one trial, in order to compute the encoding in this hold-out trial. L6 CT cells significantly encoded the direction of vibrissa stimuli, both for changes in direction over the entire train, as well as for changes within the train Fig.3-13bc, mirroring our finding about the encoding of stimulus amplitude, and amplitude deviants in L6 CT cells (Fig.2-23).

Similarly to our findings for amplitude changes, encoding of stimulus direction (upwards vs. downwards vibrissa deflections) in L6 CT (GCaMP6s imaging, N=84 cells), was disrupted by weak optogenetic drive ($P < 0.05$, baseline direction and direction changes, Fig.3-13). Similarly, L2/3 and 4 neurons (N=92 units and MUA recordings), which in the control condition showed both an increased firing for preferred directions and an overall increase for deviants (SSA), became more tuned to static preferred directions ($P = 0.038$ left tailed, change in spike count for preferred over non-preferred) during optogenetic activation, rather than increasing firing rates for deviants ($P = 0.003$ right tailed decrease). These findings are in line with our results using amplitude changes, and suggest that even though directional deviants could possibly arise from non-cortical mechanisms, they are still affected by the specific pattern of activity in L6.

To summarize, weak L6 CT drive maintained the overall firing rates of L6 CT cells, but changed which cells were stimulus responsive, and reduced the information content of the L6 CT population. The manipulation caused L2/3 RS cells to lose their heterogeneous stimulus change encoding, making them into weakly but significantly faithful encoders (increased firing rates for vibrissa stimulus amplitude increases), and reduced the effect of stimulus history on L2/3 cells.

3.6 Role of fast spiking interneurons

The negative change coefficients of L2/3 neurons suggest that these cells receive more inhibition for amplitude increases. We therefore examined change encoding by fast spiking (FS) interneurons. FS interneurons are the largest group of inhibitory interneurons in cortex (Markram et al., 2004; Gonchar et al., 2007) and play a fundamental role in shaping basic receptive field properties (Sillito, 1975; Priebe and Ferster, 2008; Ozeki et al., 2009; Liu et al., 2010). Specifically, inter-laminar transformations of receptive fields, as we observed here between L4 and L2/3, often depend on inhibitory interneurons (Foeller et al., 2005; Wehr and Zador, 2003; Ramirez et al., 2014; Bortone et al., 2014).

In our recordings, FS interneurons (N=43 in L2/3, 40 in L4) were classified by the spike waveform (Fig.2-5). Overall, the reliability of the stimulus encoding in FS cells was lower than in RS cells, which, together with the lower number of recorded cells made quantification of the FS encoding problematic. Nevertheless, the results obtained in this study hint at some potential mechanisms for how FS cells could mediate the observed effects, through an yet to be determined pathway from L6 cells, and could be of interest in hypotheses about FS function in the neural encoding of time-varying stimuli in neocortex.

Unlike L4 RS cells, L4 FS cells did not represent amplitude changes faithfully (CI: [-0.010, 0.009, 0.044] $P = 0.166$), but appeared to possibly encode deviations heterogeneously, ($P = 0.071$ entropy, $P = 0.028$ IQR).

FS cells in L2/3 encoded stimulus history (L2/3: $P = 0.01$ entropy, $P = 0.005$ IQR), like L2/3 RS. The encoding in L4 FS was less clear (L4: $P = 0.332$ entropy, $P = 0.029$ IQR). FS firing in superficial layers was therefore not determined by thalamic (Pouille et al., 2009; Swadlow, 2003) or local (L4) feed-forward drive. Instead, change coefficients in L2/3 and

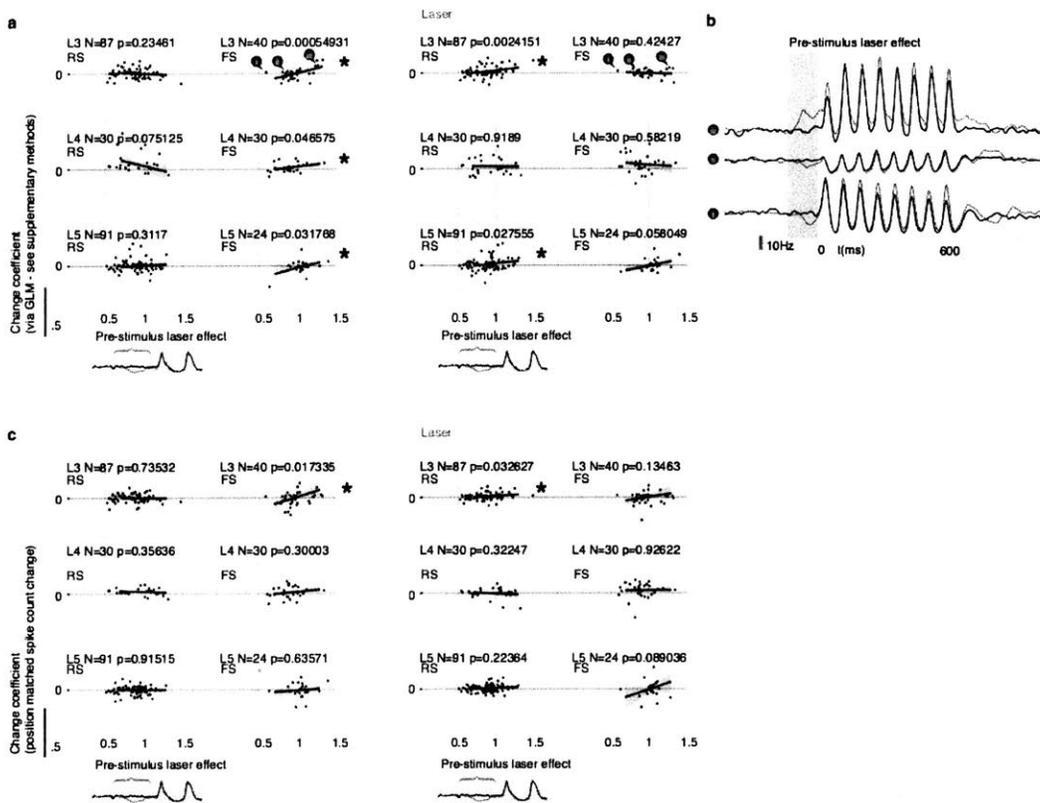


Figure 3-10: **L6 CT drive/suppression of FS cells predicts responses to stimulus deviations (change coefficients)**. **a**, Correlation between pre-stimulus laser effect (spike count in 300-50ms pre stimulus onset in laser trials / spike count for control trials) and the cell's change coefficient estimate via GLM (same as for other GLM analysis). Data analyzed individually for control (left) and laser trials (right, shaded blue). **b**, Example PSTHs for control (black) and laser (blue) conditions for 3 example FS cells, as indicated in (a). **c**, Same as (a) but using the estimate for the change-coefficient from the difference in spike counts in position matched baseline and deviant stimuli (same as in Fig.3-12) instead of the GLM estimate.

L4 FS were correlated with their response to transient L6 CT activation (Fig.3-10). We quantified the transient response of all FS cells to the onset of the weak light stimulus, and correlated the amount of inhibition or excitation that each FS cell received from this L6 drive with its change coefficient. Stronger inhibition from L6 predicted lower coefficients (L2/3: $P=0.0005$, $N=40$; L4: $P=0.046$, $N=30$, regression on cells with laser drive $< 1.5x$ baseline - the requirement of excluding 2 large outliers in this analysis shows that the FS population is more heterogeneous than the RS population), indicating that inhibition from L6 CT could contribute to change encoding by superficial FS, possibly via L6 FS (Bortone et al., 2014). Stimulus encoding in L6 could therefore act on L2/3 through a balance of drive onto RS and FS cells.

One important caveat of these findings, in addition to the weak statistical significance,

low N, and presence of outliers, is that it seems unlikely that FS cells, which are very widely connected (Bock et al., 2011) and therefore expected to receive and encode less specific information than their RS counterparts (Hofer et al., 2011), play a specific role in encoding the content of stimulus deviations. Instead the presented findings suggest that FS cells implement a gain modulation mechanism, offsetting a separate, likely more specific and information-rich stream of excitatory input. One interpretation of this idea is that the weak effects that we observed in Fig.3-10 could be due to the choice of stimulus amplitude as the deviant-defining feature, and that other, more higher-order stimulus deviations would not engage the FS population in the same manner.

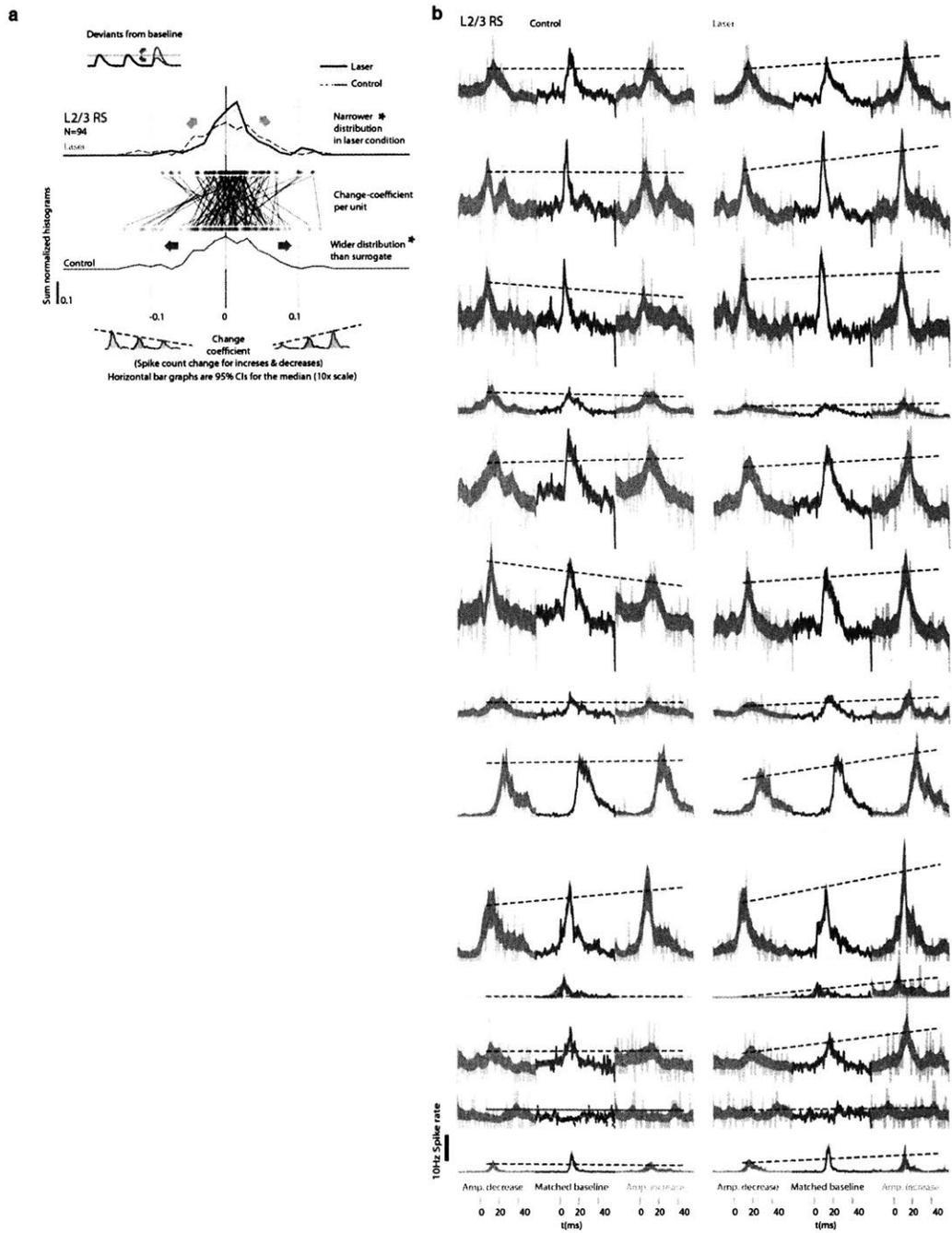


Figure 3-11: **Weak L6 drive reduces the diversity of change encoding in L2/3 responses and shifts them toward positive change coefficients.** **a**, Change coefficients (change in spike count per deflection for stimulus increases versus for stimulus decreases, for position matched stimuli) changes for control (bottom) versus weak L6 drive (top) conditions. **b**, Examples of L2/3 cells showing change representations, same analysis as in Fig.2-12

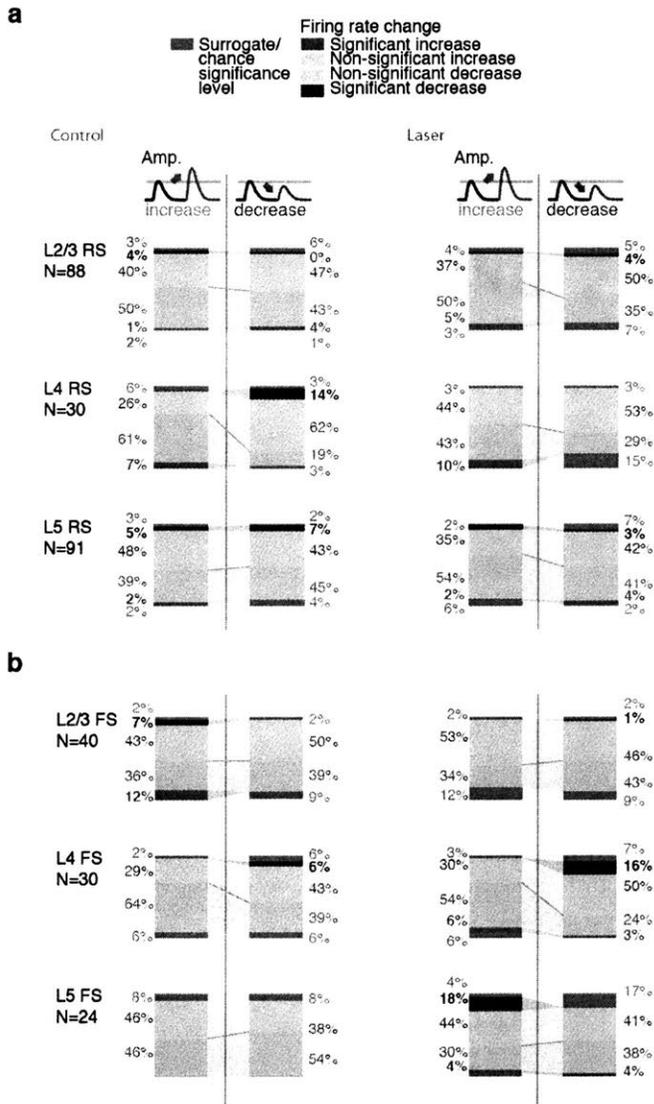


Figure 3-12: GLM coefficient statistics for all layers, cell types and conditions. **a**, Percentages of neurons with significantly increased or decreased firing rates, split by stimulus increases or decreases (each group rounded to nearest integer, as such totals do not necessarily add to 100%). Same analysis was run for control and laser conditions. **b**, Same as A but for fast spiking (FS) cells. The shift away from heterogeneous deviant encoding in L2/3 RS in the laser conditions is reflected in the higher proportion of positive coefficients. Significance of individual neurons was assessed using a GLM (see Supplemental Methods), using a 95% significance level at ~500 trials. The (relatively small) numbers of individually significant neurons using this criterion do therefore not reflect the information content of their spike output, but should be interpreted only in their difference across layers, cell types and control versus laser conditions. In L2/3, population-level analysis of deviant encoding shows robust encoding of deviants via the variance of their change coefficients (Fig. 2-12), and via ideal observer analysis (Fig. 2-14). As shown in Fig.5-1, relatively small changes in firing rate per cell correspond to robust stimulus encoding across the population of sensory responsive neurons.

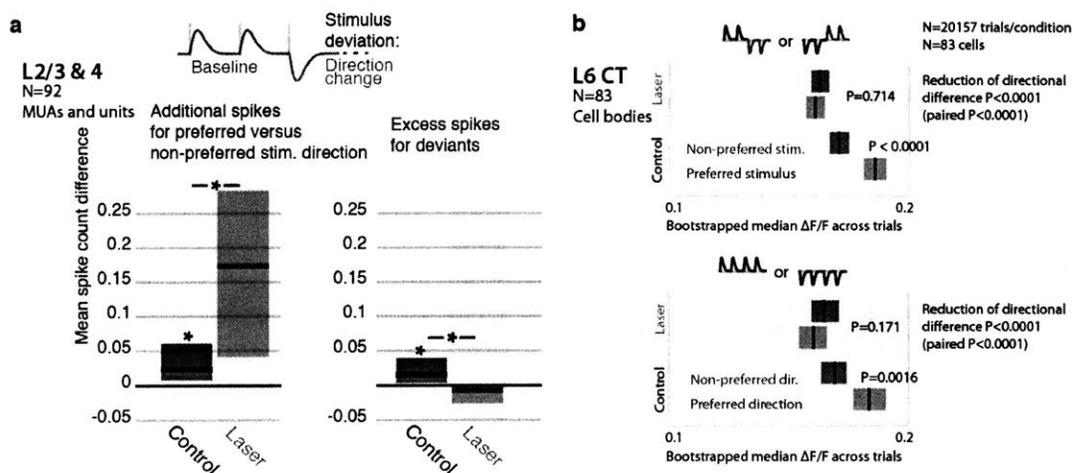


Figure 3-13: L6 CT cells encode stimulus direction, and weak L6 drive disrupts stimulus-specific-adaptation effects for vibrissa deflection direction changes a, Stimulus design for electrophysiological study of directional deviants, same as for head-fixed behavioral stimulus detection task (9 sessions). Baseline direction was randomly chosen (see Methods). Spike rate increase to preferred deviant direction (N=92 units and MUs) relative to baseline. Preferred direction was determined by the tuning to baseline stimuli. L6 activation increased the relative response to the preferred versus non-preferred direction (rank sum, $p=0.038$ left tailed increase in spike count increase for preferred). Spike rates were increased for directional deviants (relative to baseline, directions are balanced). L6 drive removes this change encoding ($p=0.003$ right tailed decrease in additional spikes). **b,** Direction encoding in L6 CT cells (83 cell bodies, GCaMP6s 2-photon, see main text and methods). Top: L6 CT cells preferentially responded to either trains of upwards, followed by downwards deflections or vice-versa ($P < 0.0001$ rank-sum, 8 deflections, 10Hz, see Methods). For each trial, all other trials of that L6 CT cell body in the same session were analyzed to determine the preferred stimulus and the held out trial was counted either as preferred or non-preferred for the analysis (see Figure 3-8 for details). Bottom: Same analysis, but whole trains of vibrissa deflections were made up of upwards or downwards deflections. L6 CT significantly encoded direction ($P=0.0016$). Weak L6 drive disrupted the encoding of deflection direction in L6 CT cells, in both cases (reduction of differences across preferred and non-preferred between control and laser conditions $P < 0.0001$ via non-paired bootstrap across trials in both cases, $P < 0.0001$ paired by cell via bootstrap, see Fig.3-8).

Bibliography

- Bock, D. D., Lee, W.-C. A., Kerlin, A. M., Andermann, M. L., Hood, G., Wetzell, A. W., Yurgenson, S., Soucy, E. R., Kim, H. S., and Reid, R. C. (2011). Network anatomy and in vivo physiology of visual cortical neurons. *Nature*, 471(7337):177–182.
- Bonin, V., Histed, M. H., Yurgenson, S., and Reid, R. C. (2011). Local diversity and fine-scale organization of receptive fields in mouse visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 31(50):18506–18521.
- Bortone, D. S., Olsen, S. R., and Scanziani, M. (2014). Translaminar Inhibitory Cells Recruited by Layer 6 Corticothalamic Neurons Suppress Visual Cortex. *Neuron*, 82(2):474–485.
- Chen, J. L., Carta, S., Soldado-Magraner, J., Schneider, B. L., and Helmchen, F. (2013). Behaviour-dependent recruitment of long-range projection neurons in somatosensory cortex. *Nature*, 499(7458):336–340.
- Foeller, E., Celikel, T., and Feldman, D. E. (2005). Inhibitory sharpening of receptive fields contributes to whisker map plasticity in rat somatosensory cortex. *Journal of Neurophysiology*, 94(6):4387–4400. WOS:000233317100066.
- Gonchar, Y., Wang, Q., and Burkhalter, A. (2007). Multiple distinct subtypes of GABAergic neurons in mouse visual cortex identified by triple immunostaining. *Frontiers in Neuroanatomy*, 1:3.
- Helmstaedter, M., Staiger, J. F., Sakmann, B., and Feldmeyer, D. (2008). Efficient Recruitment of Layer 2/3 Interneurons by Layer 4 Input in Single Columns of Rat Somatosensory Cortex. *The Journal of Neuroscience*, 28(33):8273–8284.
- 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. *Nature Neuroscience*, 14(8):1045–1052.
- 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(5):858–871.
- Lin, J. Y., Lin, M. Z., Steinbach, P., and Tsien, R. Y. (2009). Characterization of Engineered Channelrhodopsin Variants with Improved Properties and Kinetics. *Biophysical Journal*, 96(5):1803–1814.

- Liu, B.-h., Li, P., Sun, Y. J., Li, Y.-t., Zhang, L. I., and Tao, H. W. (2010). Intervening inhibition underlies simple-cell receptive field structure in visual cortex. *Nature Neuroscience*, 13(1):89–96.
- Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., and Wu, C. (2004). Interneurons of the neocortical inhibitory system. *Nature Reviews. Neuroscience*, 5(10):793–807.
- Olsen, S. R., Bortone, D. S., Adesnik, H., and Scanziani, M. (2012). Gain control by layer six in cortical circuits of vision. *Nature*, 483(7387):47–52.
- Ozeki, H., Finn, I. M., Schaffer, E. S., Miller, K. D., and Ferster, D. (2009). Inhibitory Stabilization of the Cortical Network Underlies Visual Surround Suppression. *Neuron*, 62(4):578–592.
- Pouille, F., Marin-Burgin, A., Adesnik, H., Atallah, B. V., and Scanziani, M. (2009). Input normalization by global feedforward inhibition expands cortical dynamic range. *Nature Neuroscience*, 12(12):1577–1585.
- Priebe, N. J. and Ferster, D. (2008). Inhibition, spike threshold, and stimulus selectivity in primary visual cortex. *Neuron*, 57(4):482–497.
- Ramirez, A., Pnevmatikakis, E. A., Merel, J., Paninski, L., Miller, K. D., and Bruno, R. M. (2014). Spatiotemporal receptive fields of barrel cortex revealed by reverse correlation of synaptic input. *Nature Neuroscience*, 17(6):866–875.
- Sillito, A. M. (1975). The contribution of inhibitory mechanisms to the receptive field properties of neurones in the striate cortex of the cat. *The Journal of Physiology*, 250(2):305–329.
- Swadlow, H. A. (2003). Fast-spike Interneurons and Feedforward Inhibition in Awake Sensory Neocortex. *Cerebral Cortex*, 13(1):25–32.
- Vélez-Fort, M., Rousseau, C. V., Niedworok, C. J., Wickersham, I. R., Rancz, E. A., Brown, A. P. Y., Strom, M., and Margrie, T. W. (2014). The Stimulus Selectivity and Connectivity of Layer Six Principal Cells Reveals Cortical Microcircuits Underlying Visual Processing. *Neuron*, 83(6):1431–1443.
- Wehr, M. and Zador, A. M. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature*, 426(6965):442–446.
- Zhang, Z.-w. and Deschênes, M. (1997). Intracortical axonal projections of lamina VI cells of the primary somatosensory cortex in the rat: A single-cell labeling study. *J Neurosci*, pages 6365–6379.
- Zhang, Z. W. and Deschênes, M. (1998). Projections to layer VI of the posteromedial barrel field in the rat: A reappraisal of the role of corticothalamic pathways. *Cerebral Cortex (New York, N.Y.: 1991)*, 8(5):428–436.

Chapter 4

Role of L6 in the perception of stimulus changes

The findings described in Chapter 3 show that stimulus encoding in L6 CT cells plays a role in the encoding of stimulus deviations in L2/3, and suggest that disrupting this change encoding should also selectively affect the perception of stimulus changes. We tested this prediction in two separate behavioral paradigms, one head-fixed threshold stimulus detection task and one freely behaving gap-crossing task.

4.1 Saliency of deviant stimuli

As outlined in Chapter 2 and Fig.2-7, stimulus deviations are expected to increase stimulus saliency. In behavioural experiments, stimulus deviations markedly improve (decrease) detection thresholds (Goble and Hollins, 1993; Musall et al., 2014), suggesting that this effect could be used to examine the role of L6 encoding. Even though, as outlined before, we generally sought to avoid stimulus paradigms in which deviants result in increased stimulus salience because many of the underlying mechanisms are likely different from the cortical circuit that underlies explicit comparison between (potentially top-down) expectations and current input, we attempted to test the impact of disrupting the L6 encoding on behaviour.

Even though human psychophysics results suggest that amplitude changes increase the detectability of weak stimuli (Krauskopf, 1980). Even contrast decreases in a visual detection task that are conceptually very similar to the amplitude in-, and decreases employed here, can result in decreased detection thresholds (Zhaoping and Jingling, 2008). Nevertheless, here we first tested the impact of weak L6 drive on this effect using deflection direction deviations (Khatri and Simons, 2007) that we found could reliably decrease detection thresholds in mice.

4.2 Head-fixed stimulus detection

4.2.1 Animal subjects

For head-fixed behaviour, one NTSR1-Cre mouse using viral injection, and 3 reporter line crosses (NTSR1/ChR2 +/-) were used. For gap-crossing, 6 NTSR1 mice (2 ChR2 viral injection, 4 reporter line crosses) were used. Surgical methods were as described before, but no drive implants were used, and instead only fiberoptic fibers and head-posts were implanted.

4.2.2 Behavioral training and stimulus design

Training began > 10 d after postoperative recovery and at least > 7 d after onset of water restriction (1 ml/d). Mice were secured to the head-post apparatus, and rested on a platform. Initial training procedure was as described before (Siegle et al., 2014). White noise (~65dB) was used to mask auditory cues. If mice licked up to 800 ms after the onset of the vibrissae stimulus, water was delivered. There was an additional time-out period of 2 sec for false alarms, and a pre-stimulus delay period (1–4 sec) was gradually introduced, during which licking resulted in a reset of the delay timer. This design was chosen in order to make it harder for the mice to guess the onset time of the stimulus and therefore obtain reward independent of perceiving the stimulus. See Fig.4-1 for an overview of the timing and structure of the experiments.

Vibrissae were stimulated with a custom stimulator based on piezoelectric wafers (Noliac CMBP09, Extended Data Figure 6). Stimulations consisted of deflections with a fast onset velocity and a slower ~ 80 ms return to baseline with a small (10% of peak amplitude) negative deceleration period to reduce a 2nd deceleration peak and to reduce the impact of piezo hysteresis. See Extended Data Figures 3,6 for stimulus design. Several vibrissae, centered around the C2 vibrissa, were gripped ~ 5 mm from the mystacial pad. Amplitudes were calibrated using videography. Piezo elements were replaced if ringing exceeded 10% of the peak amplitude, or if the stimulus amplitude deviated by > 5% (Fig.4-2), or if any hysteresis was measured.

Water delivery was controlled by a solenoid valve (Lee Co.), calibrated to give an ~8µl per opening (30–60 ms). Licking was detected via infrared detectors. After reaching criterion, optogenetic stimulation was added on half of trials. Stimulation started at a variable offset of 0.2 – 1.5 seconds preceding tactile stimulation and persisted for the duration of the stimulus. Laser power was ramped up with a gaussian onset profile lasting ~ 200ms (Fig.3-3). Vibrissa stimulus amplitudes were drawn uniformly from a range (~=0-30mm/sec), adjusted manually to maintain performance while probing small stimulus amplitudes. In 10% of trials, maximum amplitude stimuli were delivered. Behavioral experiments were controlled using a custom state machine (www.github.com/open-ephys/behavioral_state_machine, developed by Tim Buschman) written in Matlab via PCI DIO boards (National Instruments). Mice were weighted daily, and animals that did not consume 1ml of water/session or lost weight were supplemented with water in their home cage several hours after the experiment finished.

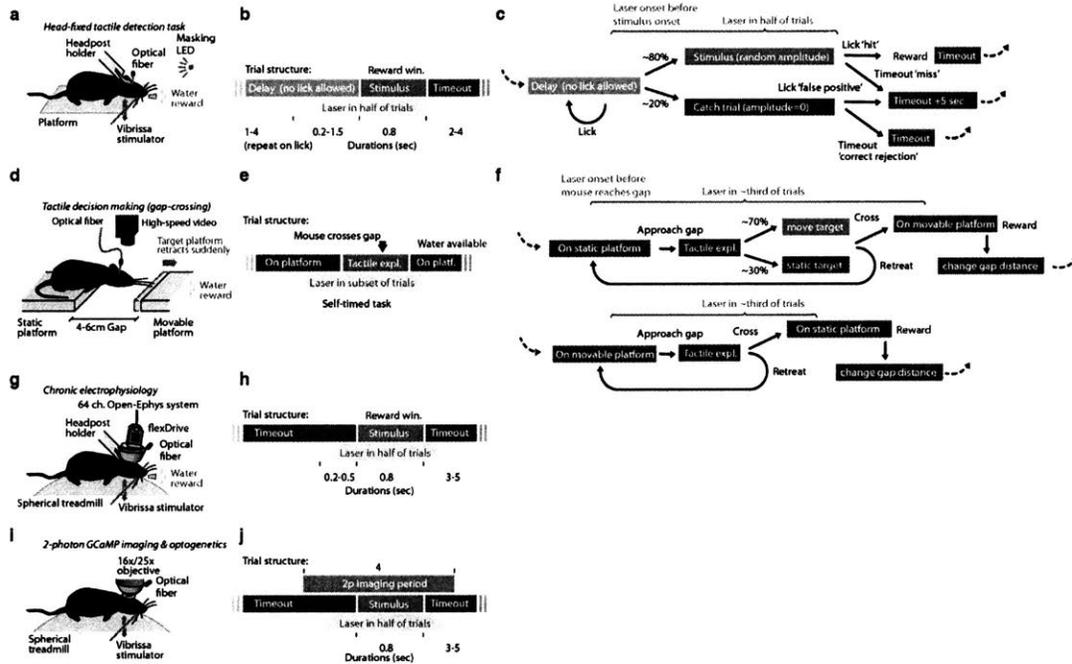


Figure 4-1: Overview of trial structure and timing for the experimental paradigms used in this study. **a**, Experimental preparation for head-fixed detection behavior (N=4 mice). Mice were head-posted and rested on a platform. Correct stimulus detection was rewarded with water. **b**, Trial timing for behavioral testing (Fig. 4a,b, Extended Data Figure 28). Early licking in a variable timeout period reset the timeout. **c**, State diagram for head-fixed detection task. **d**, Setup for unrestrained gap-crossing behavior (N=6 mice, Fig. 4c-f, Extended Data Figure 31). Mice freely crossed between two platforms for water reward. The platform distances were chosen randomly between 4-6cm, and on a subset of trials the target platform was pulled back by 2mm mid-exploration. **e**, Timing structure for gap-crossing. All events except for the timing of the laser stimulus and the gap-repositioning were chosen freely by the animals. **f**, State diagram for gap-crossing. **g**, Setup for chronic electrophysiology (N=5 mice). Correct stimulus detection was rewarded with water. Mice rested or walked/ran on a spherical treadmill to promote comfort and longer recording sessions. **h**, Trial timing for electrophysiology. No timeout periods or catch trials were used. **i**, Setup for 2-photon imaging (N=2 mice) and simultaneous 2-photon imaging and optogenetics (N=4 mice). Mice were not water restricted and rested or walked/ran on a spherical treadmill. **j**, Trial timing for imaging experiments. No reward was used, but stimulus timing was randomized as in other conditions.

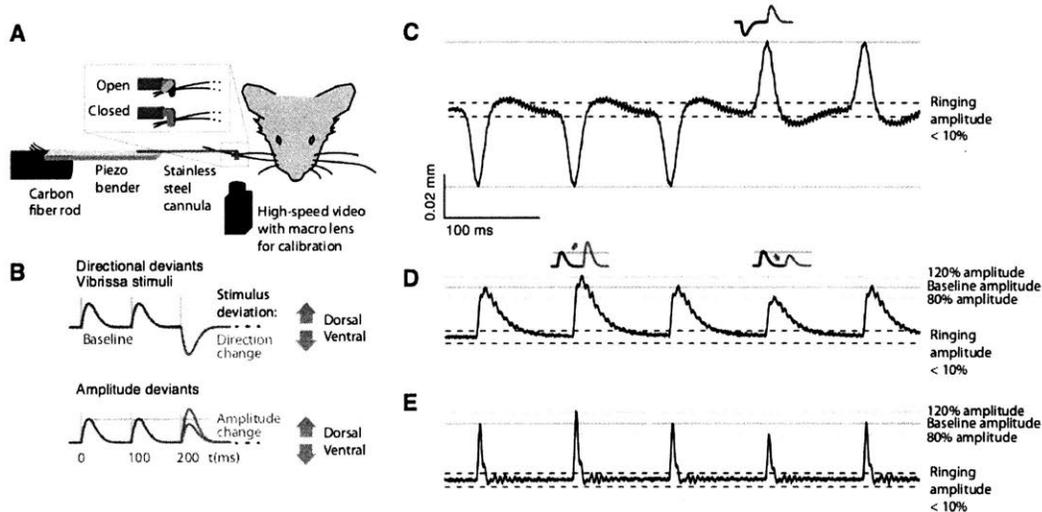


Figure 4-2: Vibrissa Stimulator design and calibration. **a**, Design of the vibrissa stimulator. A piezoelectric plate bender (Noliac CMBP09) is glued to a flattened carbon fiber rod with cyanoacrylate glue. The vibrissa grabbing arm is made from 18-gauge stainless-steel cannula to which a piece of steel is soldered at a 90 deg angle. A 22-gauge cannula with a 90 degree bend is free to rotate inside the main cannula, creating small jaws that can lock vibrissae into place reliably (see insert), ensuring consistent locking between vibrissa and stimulator with no additional ringing or slippage. Contact between the jaws is created by a thin coat of epoxy (5 min cure time) on one of the jaws that is left to cure for ~15 min before the clamp is closed, creating a closely molded fit between the jaws. **b**, Overview of the stimulus design for the directional and amplitude deviant conditions. Baseline directions in the direction deviant condition were randomized. **c**, Calibration test result from a vibrissa stimulator. The tip of each stimulator was filmed with a high-speed camera at 500 Hz with an exposure time of < 0.02 ms, about once per week. Repeated measurements at small time offsets were interleaved to obtain measurements at a 10 kHz imaging rate. The tip of the stimulator was located in each frame by calculating the center of the image in a thresholded region of interest, and the resulting motion trace was scaled to mm via measuring a calibration target of known size. The method was used to verify peak amplitude, absence of significant hysteresis and absence of significant post-stimulus ringing that could indicate damage to the steel cannula or fatigue of the piezoelectric. **d**, As in panel c but for higher onset velocity stimuli in which onset velocity and peak amplitude was varied. The Baseline (10%) and deviant amplitudes (80 and 120%) are followed precisely by the motion of the stimulator. **e**, Stimulators were also tested with very short deflections to verify that the offset ringing of the piezoelectric stayed within a 10% margin.

4.2.3 Head-fixed stimulus detection

Mice were trained to lick to indicate stimulus detection for water reward (Stüttgen and Schwarz, 2008; Sachidhanandam et al., 2013; Miyashita and Feldman, 2013; Siegle et al., 2014) (go/no-go task, Fig. 4-3). Mice that learned the task ($N = 3$) were run for 50-100 sessions. In order to analyze periods where active stimulus processing in neocortex contributed to stimulus detection (Siegle et al., 2014), only periods of high performance were analyzed via a threshold of the d' statistic, which is defined as the difference $d' = z(H) - z(F)$ in Z-scores between the hit-rate H and the false-positive rate F and measures task engagement ($d' > 1.2$).

In half of trials, direction deviants were present in stimulus trains at positions 2-4. For less salient, harder to perceive stimuli that require neocortex (Miyashita and Feldman, 2013) ($< 60\%$ of threshold amplitude, estimated per session, 1436 trials), deviations increased the hit rate (Fig. 5, 30.0% base vs. 41.5% deviant, $P = 0.001$ binomial). This finding recapitulates similar effects in rodents, nonhuman primates, and humans (Goble and Hollins, 1993; Musall et al., 2014; Ferrington et al., 1977). This result shows that even though the directional deviants employed here might be subject to SSA (Fig.2-7, Fig.3-13), they can be used to assess the basic behavioural effect of stimulus changes.

4.2.4 Head-fixed behavioral analysis

Data analysis was performed in Matlab (Mathworks) as described before (Siegle et al., 2014). Trials were selected based on a d' threshold of 1.2. We also excluded mice in which the false-positive rate was increased in the laser condition ($> 95\%$ binomial confidence bound). One mouse was excluded based on this criterion, which was established before the start of experimentation. To account for differences in the clamping distances from the follicle, angles and number of vibrissae across sessions, we estimated the threshold amplitude per session: (i) The d' was computed as described and trials with $d' > 0.8$ were analyzed. (ii) A cumulative gaussian was fit to the stimulus amplitude and hit rate, and the median point of the curve was defined as the threshold amplitude for that session. Subsequent analyses were performed on amplitudes normalized to this threshold. This normalization resulted in a hit rate of $\sim 80-100\%$ for stimuli of normalized amplitude 1 in the $d' > 1.2$ filtered data. Trials with responses within 50 ms of stimulus onset as were trials with stimulus deviations later than 400ms, to exclude licking that was not elicited by the stimulus, or trials in which the stimulus deviation was too late to contribute to the detection performance (reaction times for smallest stimulus amplitudes were $\sim 200-500$ ms 95% CI, Fig.4-3d,f). Binomial confidence bounds were computed with the Clopper-Pearson method at the 95% level. Statistical significance of comparisons between hit rates across conditions was calculated using a bootstrap (10000 samples) on binomial distributions. Mice were performing at chance level when the stimulator was detached from the vibrissa ((Siegle et al., 2014) and Fig.4-4).

4.2.5 Effect of weak L6 CT drive on change induced stimulus saliency

Consistent with the disruption of change encoding with simultaneous maintenance of sensory response amplitudes in S1, weak L6 drive did not affect detection of stimuli that did not include a deviant (30% vs. 31%, $P=0.33$). This finding is consistent with the hypothesis

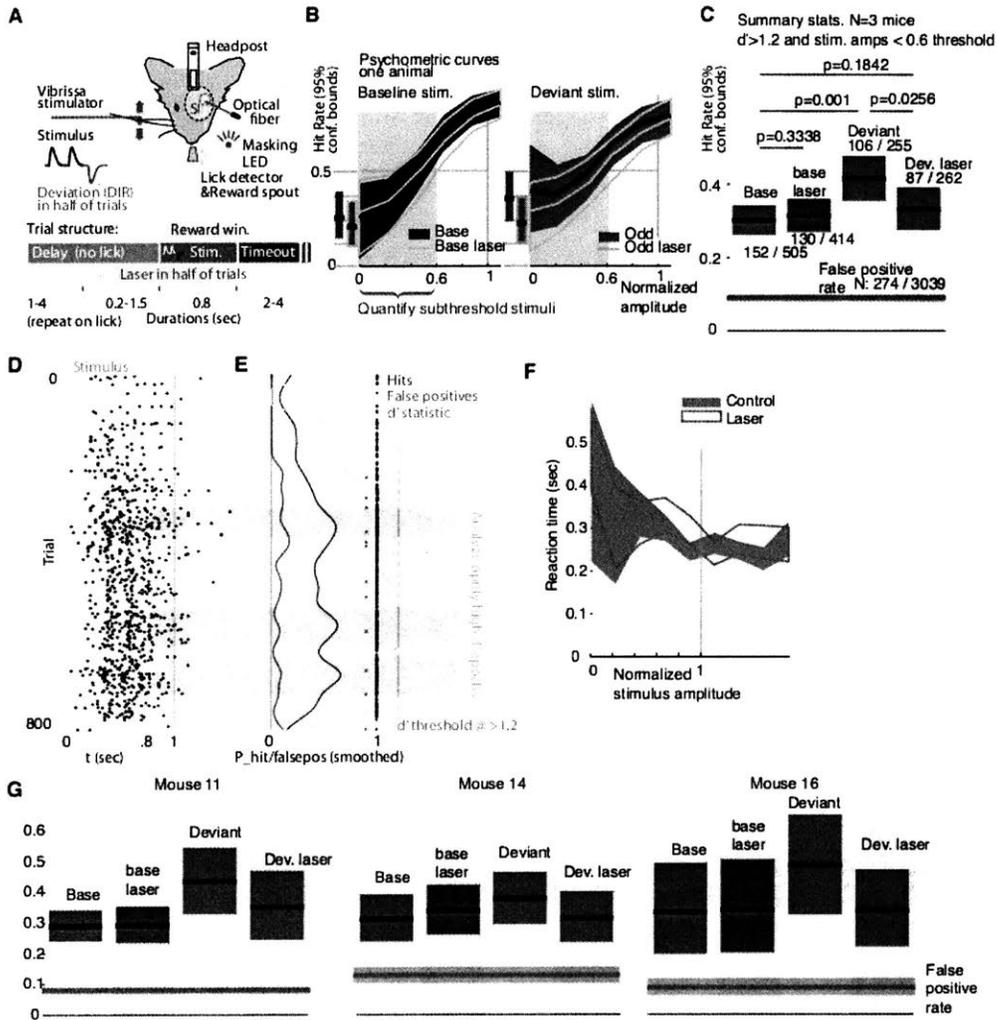


Figure 4-3: Data analysis and per-mouse results for head-fixed detection behavior analysis. **a**, Experimental conditions and task timing diagram (see Extended Data Figure 1). **b**, Psychometric curves (95% CIs, Clopper-Pearson method) for control (black, green) and laser (blue outline) conditions for baseline (left) and deviant (right) stimuli. For description of how normalized amplitudes were computed, see Methods (Head fixed behavioral analysis). **c**, Binomial confidence bounds and p values (bootstrapped binomial) across all 3 mice. (N=1436 trials with $d' > 1.2$ and normed stimulus amplitude < 0.6). **d**, Raster of lick times in one session. **e**, Hits (blue) and false positives (red) for same session as in **a**, with smoothed traces (gaussian window, $\sigma = 20$ trials) and resulting d' statistic (green). Only trials with $d' > 1.2$ were analyzed further. **f**, Reaction time as a function of stimulus amplitude for the baseline (gray) and laser (blue) condition. Shaded areas are 95% confidence bounds for mean. **g**, Results for individual mice (same analysis as in **c**).

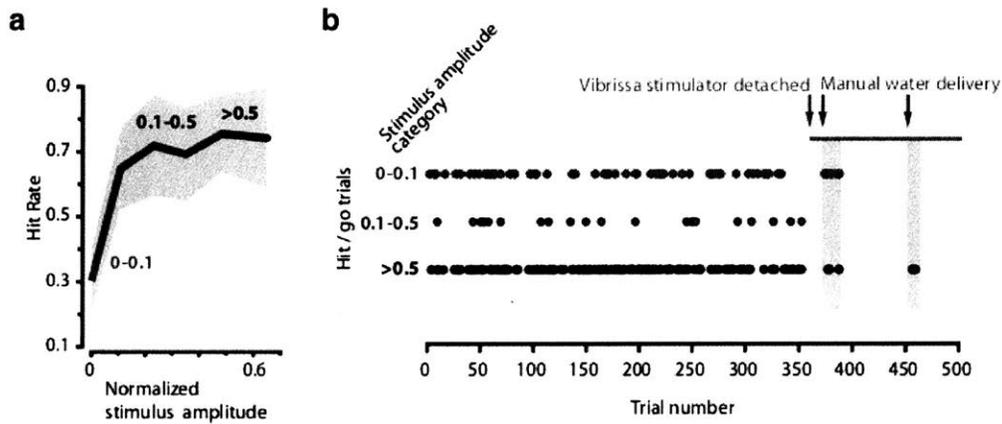


Figure 4-4: Head-fixed behavior depends on vibrissa stimulation and is independent of other (auditory or visual) cues. Mice were trained to detect vibrissa stimulation (in a separate experiment) using the same vibrissa stimulator, masking noise, absence of <650nm light, and behavioral protocol as describes in Fig.4-3. **a**, Psychometric curve for one session, stimulus amplitudes were split into three categories based on stimulus amplitude (< 0.1, 0.1 – 0.5 and >0.5 x maximum amplitude). **b**, Raster plot of successful detections (licks in go-trials) over the session (500 trials). Misses or catch trials are not plotted. Towards the end of the session, the stimulator was opened freeing the vibrissae, but remained in the same position, so that vibrissa contact to the stimulator was still possible and any vibration, visual or auditory cues were the same as in the beginning of the session. In two blocks, reward was given manually to verify that the animal was still attentive and could react to reward delivery (indicated by the click of a solenoid) with licking (indicated as hits in cases where water was given in non-catch/go trials).

that threshold detection tasks are predominately determined by the number of evoked spikes in primary sensory cortices.

However, this manipulation selectively reduced performance for deviant-containing stimuli (Fig.4-3b, $P=0.025$), making them about as hard to detect as baseline stimuli. This shows that even though the directional deviants could possibly obtain higher saliency through subcortical SSA, the L6 dependent mechanism outlined in this thesis still selectively affects their preferential detection.

In agreement with similar results in humans (Wickelgren, 1977; Niemi and Näätänen, 1981), we observed that the reaction times of the mice were somewhat reduced for larger stimulus amplitudes (Fig.4-3f). This suggests that even in simple detection tasks, mice might integrate information over multiple whisker deflections in order to overcome the noise in individual deflections, similar to the integration behaviour during free gap-crossing (Hutson and Masterton, 1986; Celikel and Sakmann, 2007). This integration did not appear to be affected by the L6 drive, consistent with the lack of effect on the hit rate for non-changing stimuli.

Because this behavioral stimulus design differed from the amplitude deviants used in the other experiments, we also examined the impact of weak L6 CT drive on neural representation of direction deviations in separate experiments (Chapter 3, Fig. 3-13), and found that, as expected from other SSA findings (Antunes and Malmierca, 2013; Müller et al., 1999; Ferrington et al., 1977; Hollins and Favorov, 1994), in the control condition, directional deviants result in an overall increase in firing rates, and the weak L6 drive reduces this increase, making individual cells more responsive to their preferred direction instead (Fig.2-7).

4.3 Untrained free behaviour

The head-fixed detection task has the feature of providing well-defined psychometric function, but is in many ways a non-natural condition. We therefore further examined the impact of small, sudden stimulus changes in a naturalistic and untrained sensory decision making task, gap-crossing (Hutson and Masterton, 1986) (Fig. 6a), where mice use their vibrissae to locate and cross between elevated platforms whose distance is changed after each trial ($\sim 4-6$ cm, 6 mice). The experiments were performed under near infrared illumination and with white noise to exclude visual or auditory contributions to task performance (Voigts et al., 2008, 2015).

4.3.1 Gap-crossing

Mice ($N=6$) were implanted with plastic head-posts and fiber stubs, and water restricted as described, vibrissae on the side ipsilateral to the fiber were trimmed. Two mice also had implants for electrophysiology. The gap-crossing apparatus consisted of two facing platforms (Hutson and Masterton, 1986) (58mm wide) over a custom LED backlight (650nm). Mice were habituated for 2 days prior to the experiments. On day 3, the optical fiber was attached and masking noise (~ 80 dB) was introduced. After mice crossed the gap in either direction, a water reward ($\sim 0.01 - 0.05$ ml) was delivered manually, and a new platform position was chosen between ~ 45 and 65mm. In half the trials the laser was on for > 1 sec prior till ~ 2

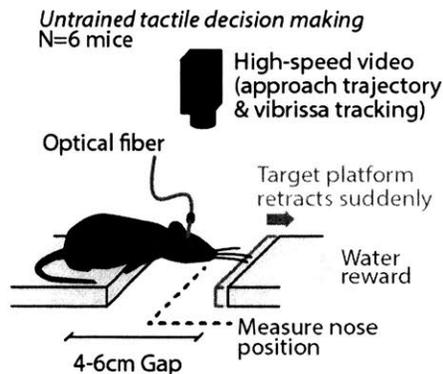


Figure 4-5: **Experimental setup for gap-crossing.** In an unrestrained gap-crossing task, the target platform can suddenly retract by ~ 2 mm during a bout of tactile sampling (Voigts et al., 2008). Observation and tracking of the vibrissae via high-speed video allows quantification of when and how vibrissae contacted the target platform (allowing, amongst other things to exclude trials when vibrissae contacted the platform during platform movement), and measurement of mouse posture over time allows quantification of how much time the mouse spends locating the target. This assay yields a simple read-out of whether mice perceived the small, sudden platform motion without requiring any task-specific training.

sec after the crossing. In a subset of trials, one platform was retracted by 2mm within 8ms via a voice coil actuator ((Voigts et al., 2015) and Fig.4-6) while the mouse was palpating it.

Mice were run every other or third day, sessions ended when mice either lost interest in crossing, fell from the platform, or tangled the optical tether. In control sessions, the optical cable was attached to a mock ferrule that directed the light to a position rostral of the actual fiber stub implant.

4.3.2 Gap-crossing analysis

The gap was filmed from above at 315 Hz (Pike 032B, Allied Vision Technologies). The mouse nose distance to the target platform was tracked using custom scripts in Matlab. The high frame rate (330Hz), low ($< 250\mu\text{s}$) exposure times and very small aperture (Navitar HR F1.4 16mm f stop adjusted down to a depth of field of a few cm). There should be no motion or out-of-focus blurring. This is vital for detailed tracking of the mices whiskers. The requirement for low exposure times and small aperture mean that a lot of light is needed.

The only light source was a very uniform and bright backlight, we use red ($> 650\text{nm}$), because this wavelength is invisible to mice (Jacobs et al., 1999). We made the backlight from 12 red 700mA LEDs glued to a thin ($\sim 2\text{mm}$) aluminum plate that is bolted to the optical breadboard, which then acts as a passive heat-sink. On this sits a box made from mirrored acrylic, and on top of that two sheets of frosted glass, acting as a diffuser (a few mm between the two sheets make the diffuser much more efficient than a single layer of diffusing glass). The LEDs were positioned to achieve good uniformity this made it possible to use simple thresholding for many operations, and is important for whisker tracking later. No room lights were used, and glare from computer screens etc. was minimized

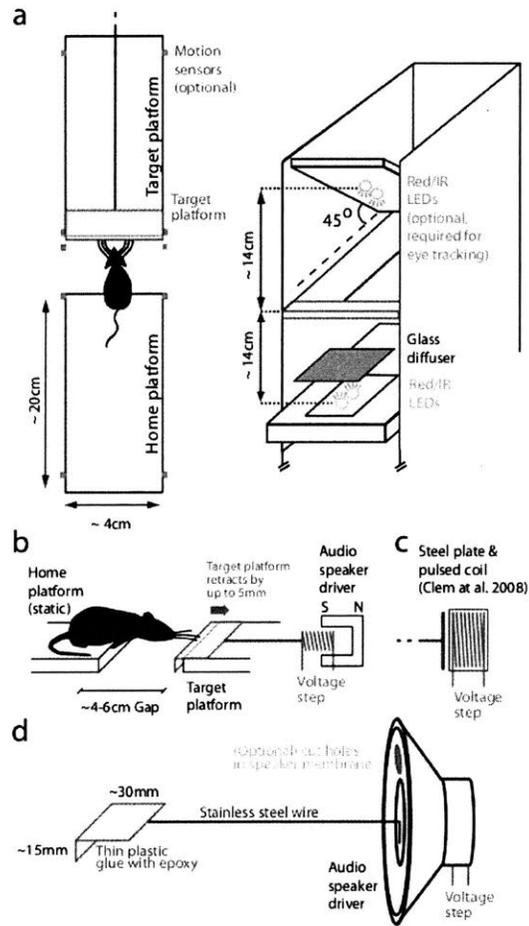


Figure 4-6: Method for movable target platform. **a**, Overview and measurements of the gap-crossing setup, adapted from Voigts et al. (2015). The platform is made from sheets of transparent acrylic. The moving target platform (red) is placed on top of the target platform and moved via a thin stainless steel wire (blue) that extends through the back of the target platform. The gap distance is adjusted by moving the entire home platform. An array of infrared LEDs and a diffuser glass is mounted at 14cm below the gap. **b**, Overview of the moving platform assembly. The moving target platform (red) is moved by a wire that is attached to an audio speaker driver and application of a voltage step will retract or protract the target platform by a few mm. The platform is calibrated so that it is quickly stopped by the acrylic platform at the retracted position. **c**, In Voigts et al. (2015) a steel plate was placed 5mm in front of a copper coil, so that application of a voltage step retracted the steel plate until it hit the coil enclosure. **d**, Measurements of the moving platform assembly.

with red filters. The platforms were at least somewhat transparent and left some of the backlight through, making it possible, if still hard to track the mouse once it intersects them.

The position of the mouse was tracked automatically using basic image operations and temporal filtering (5 frames moving median filters for excluding one-frame errors in position tracking etc.). Trials were identified as attempted (mouse reached over the gap) or completed crossings. For analysis of sensory disruption using high laser powers, the probability of crossing was computed from all trials with gap distance $< 6\text{cm}$. For other analyses, only trials in which the mice crossed within 5 sec were further analyzed. The nose position over time was aligned to the position at which the mouse had committed to a crossing attempt without touching the target yet, extending over the home platform by $\sim 7\text{ mm}$, corresponding to a position of -20mm in the imaging reference frame.

4.3.3 Unsupervised approximate whisker tracking

In order to assess whether and when vibrissae contact the platform before and/or after the platform motion, and to quantify whether our experiments induced any anomalies in the whisking pattern of the mice we use an unsupervised vibrissa tracking system.

Subsets of vibrissae were tracked using an automated tracker (Figs.4-11,4-7, www.github.com/jvoigts/whisker_tracking). This approach is different from precise all-whisker tracking, were typically all vibrissa are tracked (this is currently done almost exclusively in subsets of vibrissae) and where precise contact times (for electrophysiology etc.) and contact parameters such as whisker bending over time are desired to estimate torques. Usually for these cases, imaging frequencies of $\geq 1\text{kHz}$, clipped whiskers, where only a row (usually C row) or even only one whisker is left, and even go to head-fixed experiments (O'Connor et al., 2010). In these cases the convnet step detailed here should still be useful, but a more sophisticated method to track parametric whisker shapes would have to be used. The method I developed as an undergrad project (Voigts et al., 2008) is one example of this, but much better methods have been published since, like Clack et al. (2012) (well documented code available), Knutsen et al. (2005) (initial paper in 2005) and updated tracker code (github.com/pmknutsen/whiskertracker), or the BIOTACT Whisker Tracking Tool (Gyory et al. (2010), <http://bwtt.sourceforge.net/docs/>).

The method we used here starts by first locating whether there's a mouse in each frame (to avoid tracking empty images), the position of the mouse's head, and the position of the target platform. This was accomplished by a series of elementary image transformations and manually selected thresholds.

We then performed pixel-wise whisker labeling with a convolutional neural network. We next needed to label all pixels that represent whiskers, ideally independently of light conditions, background noise etc. If this labeling is sufficiently clean, a relatively simple method can be used later to get the location and orientation of individual whisker segments. Here, we used a very small convolutional neural network (Krizhevsky et al., 2012) to identify whisker pixels. This code uses ConvNet code by Sergey Demyanov (<https://github.com/sdemyanov>), but re-implementing this code in Theano, Torch, Tensorflow or any other machine learning library should be fairly trivial.

We first generated a training set of raw and binary label images in which all whiskers are manually annotated, using photoshop. In this step it is crucial to get all whiskers in these images with very high precision, and paint over or mark to ignore (and then exclude from training set) all non-labelled whiskers so that the training can run on a clean training set. We also included negative examples, including pixels from all possible occluders such as optical cables, recording tethers etc. Here, we used just 4 images and supplemented the training set with rotated and mirrored copies.

The network for this example was set up as:

An input radius of 5px, corresponding to input tiles of 11×11 pixels,

First layer has 8 outputs, second layer has 4 outputs, softmax function for output.

The input radius, size of the image tiles around the pixel that is to be identified was chosen to be as small as possible. Large radii mean more parameters to learn and slow down processing. here, we needed around 5 pixels to do a proper line/ridge detection, and maybe a few more in order to train the NN to avoid labeling whisker-like structures that are part of the target platform etc.

In order to avoid accidentally tracking pieces of fur that are too close to the head but locally look like whiskers, we need a fairly large input radius for the cnn so it could be trained to label every hair that is too close to the head as negative. Instead, because locating pixels that are part of the head is dead simple via smoothing and thresholding (the head is the only big very dark object in the images) we accepted that the cnn would give a few false positives here, and just ran a very fast cleanup pass with a much simpler convolution with larger kernel (20px diameter circle). This way the CNN ran on small easy to train 11×11 tiles and we still avoid fur labeling.

To make the training set we picked all positive examples, plus rotated copies, plus a large number of negative ones picked from random image locations. Further, to avoid over-training to the specific image brightness levels of the training set, we added random offsets to each training sample. Because of the small number of training images, we were not using a separate test set to track convergence. The Training is then run to convergence, for around ~ 4 hrs on a regular 4-core i7 system.

Once the per-pixel whisker labels looked sufficiently clearly differentiated, we ran an approximate whisker angle tracking with a Hough transform (see Gyory et al. (2010) for a similar method). Of course the labeled image would make a good input for a proper vibrissa tracking tool that can track a proper parametric whisker shape and even attempt to establish and maintain whisker identity over frames.

To get the approximate whisking pattern, a simple median or mean of the angles coming from the hough transform was used, and simply averaging (and thresholding) the CNN output (pixel wise label of whether there was a vibrissa or not) at the platform edge gives a good measure of whether vibrissae overlapped the target in any frame. This is of course no direct indicator of whether there was contact between the two, but for all analyses in this thesis it is a sufficient proxy, and at the very least gives a clear indication of whisking cycles where there was no contact.

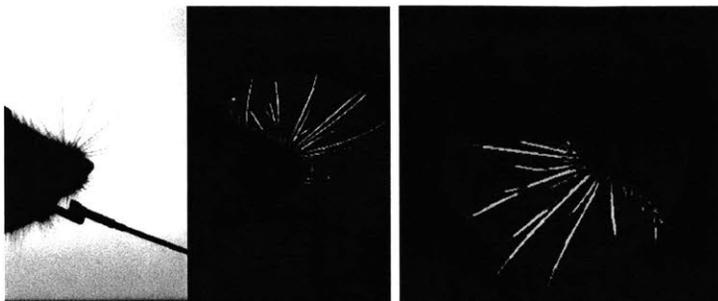


Figure 4-7: **Fast approximate whisker tracking.** **Left:** Raw input image showing unilateral whiskers and a sham-ferrule. **Middle:** Output of the CNN method, only a circular section around the mouse's face was processed. Thresholded pixel values from the CNN were used to mark intersections between vibrissae and the target platform. **Right:** Vibrissa segments tracked using a simple Hough transform. This method does not yield identified vibrissae, or curvature estimates, as do full vibrissa trackers (Knutsen et al., 2005; Voigts et al., 2008; Clack et al., 2012), but is sufficient for estimating the overall whisking pattern.

4.3.4 Assessing stimulus novelty perception in Gap crossing

To induce small stimulus changes during gap-crossing, in $\sim 70\%$ of trials where mice were crossing towards a move-able platform, the target was rapidly pulled back by $\sim 2\text{mm}$ mid-exploration using a modified bass-speaker as a voice-coil actuator (Voigts et al., 2015). This stimulus deviation was small enough not to startle mice, but mice consequently spent more time palpating, and precisely re-locating the target before crossing. We specifically verified this effect by comparing 'change' trials in which the mice contacted the retracting platform both before and after the retraction ($N=317$) to 'no-change' trials in which they only contacted it either only before ($N=49$) or only after ($N=134$) a platform motion. On these trials, mice would not perceive any change in position between the vibrissa-to-target contacts, but would experience any non-specific effects of platform motion, such as noise, air motion, table vibration, or visual cues. Only trials with crossings within 5 seconds were considered for this analysis. In the change condition, but in neither of the control conditions, mice significantly slowed their approach before crossing ($P < 0.005$ change vs. only before, and $P < 0.01$ change vs. only after, rank sum, Fig.4-8a). This shows that the platform retraction assay can be used to assess whether freely behaving, untrained mice can perceive small sudden stimulus changes.

4.3.5 Effect of weak L6 CT drive on change preception in gap-crossing

We next tested whether the same weak optogenetic manipulation used throughout would impair change detection in this naturalistic behavior. Weak L6 drive removed the extra sampling time that platform motion would otherwise have generated ($P < 0.001$ laser vs. control, $N=317$ trials, Fig.4-8b left). Optical drive did not, however, affect baseline rates of motion between platforms when targets were static ($P=0.9$, $N=751$ trials, Fig.4-8b right). Impacting L6 CT therefore selectively affected perception of sudden small stimulus changes, but had no effect for non-changing stimuli in naturalistic sensory decision making.