Canonical Neural Computation: A Summary and a Roadmap
A Workshop at Villa La Pietra, Florence, 17-19 April 2009

Overview

A variety of anatomical, physiological, and behavioral evidence suggests that the brain performs computations using modules that are repeated across species, brain areas, and modalities. What are these canonical modules, and how can we elucidate their underlying circuitry and mechanisms?

This workshop brought together a group of experimental and theoretical neuroscientists to outline a roadmap for research leading to one or more canonical models of neural computation. We aimed to develop a broad set of themes that span different levels of analysis (cellular, networks, systems, cognition, and behavior), species (mammals and invertebrates), brain areas (cortex, hippocampus, cerebellum), and modalities (vision, audition, vestibular, etc.). The participants were primarily researchers in basic science, but some also investigate developmental disabilities and disease.

Our intent was to use this meeting to catalyze new research efforts, to define a program of research that may be shared and collaborative between the various participants or subgroups of participants. To this end, we asked the speakers to give informal and discussion-oriented talks, rather than a more standard seminar. We wanted to hear mostly about what they are planning to do, and what they think the interesting questions are, rather than what they have already done. To maximize interactions, there was a strict 30 min limit for the talks, followed by 10 minutes of discussion without slides.

To plan the road ahead and identify the main goals of future investigations, and to outline the methods and collaborations needed to achieve these goals, there were two additional discussion sessions during which the participants were divided in four subgroups: (1) **physiology-systems**; (2) **physiology-mechanisms**; (3) **theory/computation**; (4) **human/ clinical**. The work of each subgroup is summarized below.

In addition to the researchers, the attendees included journal editors and program directors from funding bodies: Andrew Rossi (National Institute of Mental Health), Sarah Caddick (Gatsby Charitable Foundation), Charles Yokoyama (Neuron), Hannah Bayer (Nature Neuroscience), and Claudia Wiedemann (Nature Reviews Neuroscience). They provided advice (during a panel discussion) on how to design, implement, and communicate a transformative program of research involving multiple researchers aimed towards a common goal.

This document has multiple purposes. It is a report/summary of the workshop. But we also intend for it to be a roadmap for future research, both within the group of speakers and in a broader context. For example, it can be used as a source to outline the scope of collaborative grant applications, review papers, and position papers.

Organizers

David Heeger, NYU
Eero Simoncelli, HHMI and NYU
John Reynolds, The Salk Institute
Matteo Carandini, UCL

Generous funding was provided by The Swartz Foundation.

Schedule

Friday, 17 April 2009

7:45 AM Breakfast

9:00 AM Matteo Carandini, University College London
Normalization and competition in visual cortex
9:30 AM Discussion

9:45 AM Eero Simoncelli, New York University
Why normalization - some statistical justifications
10:15 AM Discussion

10:30 AM Coffee break

11:00 AM Dora Angelaki, Washington University
A divisive normalization model of multi-sensory integration
11:30 AM Discussion

11:45 AM David Heeger, New York University
Normalization model of attention
12:15 PM Discussion

12:30 PM Lunch

2:00 PM **Discussion session: future directions**
Subgroups: theory/computation, physiology-systems, physiology-mechanisms, human/clinical

3:00 PM Odelia Schwartz, Albert Einstein College of Medicine
Gain control in neural populations and natural scene statistics
3:30 PM Discussion

3:45 PM Dario Ringach, University of California Los Angeles
Implications of normalization for population coding
4:15 PM Discussion

4:30 PM Coffee break

5:00 PM Shihab Shamma, University of Maryland
Task difficulty and performance induce adaptive patterns in auditory cortex
5:30 PM Discussion

5:45 PM David McAlpine, University College London
Accumulating gain adaptation in the ascending auditory pathways
6:15 PM Discussion

6:30 PM Dinner

Saturday, 18 April 2009

7:45 AM Breakfast

9:00 AM Nicholas Priebe, UT Austin
Variability and invariance in cortical responses
9:30 AM Discussion

9:45 AM Frances Chance, UC Irvine
Gain modulation by subtractive and divisive mechanisms of inhibition
10:15 AM Discussion

10:30 AM Coffee break

11:00 AM John Reynolds, Salk Institute
Mapping the microcircuitry of attention
11:30 AM Discussion

11:45 AM Adrienne Fairhall, University of Washington
Intrinsic contributions to adaptive coding
12:15 PM Discussion

12:30 PM Lunch

2:00 PM **Discussion session: future directions**
Subgroups: theory/computation, physiology-systems, physiology-mechanisms, human/clinical

3:00 PM Nicole Rust, University of Pennsylvania
Commonalities of computation across the motion and object recognition pathways
3:30 PM Discussion

3:45 PM Sam Solomon, University of Sydney
One model or more? Gain control in the context of parallel visual pathways
4:15 PM Discussion

4:30 PM Coffee break

5:00 PM Larry Abbott, Columbia University
Nonlinear dimensional reduction for discrimination in the fly olfactory system
5:30 PM Discussion

5:45 PM Rachel Wilson, Harvard Medical School
Gain control in an olfactory circuit
6:15 PM Discussion

6:30 PM Banquet dinner (at La Pietra)

Sunday, 19 April 2009

7:45 AM Breakfast

9:00 AM Josh Solomon, City University London
Contextual influences on perceived orientation: gain control or not?
9:30 AM Discussion

9:45 AM Concetta Morrone, Istituto di Neurofisiologia CNR, Pisa
Cross-orientation inhibition measured with fMRI
10:15 AM Discussion

10:30 AM Coffee break

11:00 AM Anthony Norcia, Smith-Kettlewell Eye Research Institute
Normal and abnormal development of gain control and contextual interactions
11:30 AM Discussion

11:45 AM Steven Dakin, UCL
Abnormal gain control and sensory processing deficits in schizophrenia
12:15 PM Discussion

12:30 PM Lunch

2:00 PM **Panel discussion: Research opportunities**

Andrew Rossi, NIMH
Sarah Caddick, Gatsby Charitable Foundation
Charles Yokoyama, Neuron
Hannah Bayer, Nature Neuroscience
Claudia Wiedemann, Nature Reviews Neuroscience

3:00 PM Massimo Scanziani, University of California San Diego
Dynamic range and hippocampal inhibitory circuits
3:30 PM Discussion

3:45 PM Michael Hausser, University College London
Dendritic computation
4:15 PM Discussion

4:30 PM Coffee break

5:00 PM Kevan Martin, Swiss Federal Institute of Technology Zurich
Mapping the matrix: circuits, cells, and synapses in neocortex
5:30 PM Discussion

5:45 PM J Anthony Movshon, New York University
Encoding and decoding with feedforward computations in cortical networks
6:15 PM Discussion

6:30 PM Dinner

Physiology: Systems

Participants: Dora Angelaki , Sarah Caddick, Tony Movshon, John Reynolds (Coordinator), Nicole Rust, Shihab Shamma, Sam Solomon, Charles Yokoyama.

Introduction

Research in sensory systems gives strong indications that the brain may apply similar computations to different problems, and has thus identified a number of these canonical computations. They have proven capable of accounting for a wide variety of observed neurophysiological measurements.

A classical example of canonical computation is the linear receptive field. It has been found to be a powerful description of neuronal responses in the visual system including primary visual cortex (Movshon et al. 1978) and area MT (Rust et al. 2006), in somatosensory cortex (DiCarlo and Johnson 2002), and in auditory cortex (Depireux et al. 2001).

A second example of canonical computation is divisive gain control and normalization, which has been found to be a key computation in retina (Solomon et al. 2006), thalamus (Mante et al. 2008), primary visual cortex (Heeger 1992; Rust et. al., 2005), area MT (Simoncelli and Heeger 1998) and inferotemporal cortex (Zoccolan, Cox, & DiCarlo, 2005). Divisive gain control has also been implicated in mediating adaptation (Wainwright, Schwartz & Simoncelli, 2002), to adjust the sensitivity of neural systems to prevailing sensory statistics, thereby improving the quality of information encoded by a neural population code (Dean, Harper & McAlpine, 2005). Divisive gain control has been also shown to account for a wide variety of modulatory effects of attention on responses of visual cortex (Reynolds and Chelazzi, 2004; Reynolds and Heeger 2009) and may play a key role in multisensory integration (Morgan et al., 2008).

A third example of canonical computation is soft-thresholding of noisy signals. The conversion of input currents into output firing rates introduces a thresholding stage. This threshold sets the operating point of visual cortex (Ringach and Malone 2007) and allows neurons to have

invariant tuning curves (Priebe and Ferster 2008), and to effectively amplify the variability (Carandini 2004) of their responses.

These canonical computations, in turn, are typically combined in a canonical way (e.g. Rust et. al., 2005; Kouh and Poggio, 2008). Models of both areas V1 and MT, for example, posit that the greater selectivity and invariance exhibited by neurons in comparison to their inputs derives from an appropriate summation of incoming inputs, coupled with normalization and spike threshold. Models of ventral stream processing posit an analogous hierarchy of computations such that neurons in successive stages of processing exhibit selectivity for increasingly more complex combinations of certain visual features while also exhibiting increased invariance to other stimulus attributes (Cadieu et. al., 2007; Rust and DiCarlo, 2008).
Related theories have been proposed to explain the physiology of other neural systems (cortical and sub-cortical) including auditory processing (David, Mesgarani, Fritz, and Shamma, 2009), olfaction, hippocampus, cerebellum.

Examples of open questions:

- How do the "parameters" of these computations differ in different brain regions and how do these differences relate to the specific computational requirements? Do they differ according to the type of input?
- Are there fundamental differences between apparently similar brain regions. For example, is the nature of neuronal firing in auditory cortex indicative of a fundamentally different computation from the one performed in visual cortex?
- What computations are preserved across evolutionarily distant species, such as birds and mammals?
- What insights can be gained about brain function by understanding the circuitry that mediates these computations (e.g., cell type, connectivity, laminar position of cortical neurons)?
- What are the canonical computations outside of sensory systems?

Summary of discussion

We are in the midst of a golden era of neuroscience, with the integration of neurophysiological and theoretical techniques yielding rapid advances in our understanding of the neural mechanisms underlying perception and cognition. Neurophysiological experiments have provided a wealth of information on the response properties of neurons and how neuronal responses change as a function of the spatiotemporal properties of sensory stimuli (Rust et al., 2005, Solomon et al., 2006; Priebe & Ferster, 2008; Atiani et al., 2009). These data have led to the development of abstract models of single neuron responses, such as LNP and GLM models, as well as models of canonical network-level computations, such as normalization class of models (Heeger, 1992). Intracellular and extracellular recordings of the activity of small populations of neurons, local field potentials, calcium imaging and related measures, in conjunction with neuroanatomical data, have provided motivation for and constraints on models of the circuits that mediate these computations. Neurophysiological studies in animals trained to perform sophisticated behavioral tasks have characterized how population activity relates to sensation, multimodal integration, perceptual decision making, attention (Atiani et al., 2009) and cognition.

Canonical computations have served as the basic computational units in theories of higher order processes such as perceptual decision making (e.g., Corrado et al., 2005), and attentional selection (e.g., Shamma; Reynolds & Heeger, 2009), and sensory integration (e.g., Rust, et. al., 2006; Angelaki, Gu & Deangelis, 2009). These models have led to important simplifying insights into the relationship between neural computations and behavior. Studies combining neurophysiology and computational techniques have, for example, shown that a seemingly

heterogeneous variety of forms of attentional modulation can be understood as resulting from a simple model of contrast gain control. This suggests that evolution co-opted the circuitry that mediates this computation for use attentional selection (Reynolds & Heeger, 2009). Computational neuroscience has also revealed simplifying insights into the relationship between neuronal activity and perception. For example, quantitative studies examining the integration of vestibular and visual information have found that it is nearly optimal (Angelaki, Gu & Deangelis, 2009).

The advances that have been made through this integration of neurophysiology and theoretical approaches have helped define the challenges that, when met, will yield the next generation of advances. In our discussions, we outlined several key challenges.

(1) The Canonical Computation Toolbox

One of the challenges we all face in trying to synthesize and integrate results from a variety of sources is in the relatively basic realm of experimental design and analysis. All the grand notions of theory and computation in the world cannot gain traction if the data we collect do not distinguish the predictions of different theories. Yet we all have had the experience of reading a paper, observing that its outcome is "broadly consistent" with some theory, but also realizing that if only the experiments had been done or designed a bit differently, the work would have had much greater impact.

The idea is to develop a "toolbox" which consists of a collection of experimental design principles and code, intended to make experiments and results more useful to the community. Different toolbox elements could be designed for specific domains (unit physiology, EEG/MEG, fMRI, psychophysics) Some of the principles are obvious – we should all have learned them as first-year graduate students – but are often flouted. Others are a little more adventurous. Adoption of any or all of them would help.

Optimize data collection for model testing. This seems like a no-brainer, but how often do we actually fail to do this? We have all had the experience of fitting models post-hoc to data and finding that the data we actually collected were inadequate. So some kind of handbook or set of principles might be useful. One important innovation might be something that psychophysicists have done for decades using methods like PEST and QUEST: conduct experiments adaptively, placing experimental conditions where they provide the most information, continuously update model fits during the experiment, and terminate measurements when a criterion degree of precision is reached.

Formalize models to be tested. There should be a basic set of modeling tools that could and should be used to generate predictions, generic if need be, for particular experiments cast in the canonical framework. An example might be something like the code that Eero Simoncelli has on line for his motion model (http://www.cns.nyu.edu/~lcv/MTmodel/) and John Reynolds & David Heeger have for their model of attention (http://www.snl-r.salk.edu/_reynolds/Normalization_Model_of_Attention/), but somehow made more general so it could be applied to other cases of input and output.

Formalize gain control models. A subcase of the general problem would be a Matlab toolkit to explore gain control models at the circuit level. Experience with such models is invaluable in helping to determine whether particular forms of the model can be tested experimentally. This might be a much easier problem than the more general one.

A lab manual. Those of us who have done these experiments in different domains might usefully just sit down together and put together a manual of good practice. If nothing else, it would help

to train students if we had such a thing ready to hand them. There could be sections on each major methodological domain, on general statistical and signal processing approaches, etc.

(2) A set of standard methods for distinguishing the elements of the canonical circuit in awake behaving animals

A second challenge we face is how to differentiate the elements of the neural circuit in the behaving animal. An important observation about the neocortex is that very different brain functions, ranging from extracting orientation information in primary visual cortex to computations of economic value in parietal cortex and motor planning in the Frontal Eye Fields, are all computed by the a basic laminar circuit that is roughly duplicated throughout the brain with, so far as we can tell, only relatively modest variation (Douglas & Martin, 2004). Having identified canonical computations that mediate key aspects of perception and cognition, we are now poised to gain insight into the circuitry that mediates these computations. How are neural computations implemented in cortical circuits? That is, how do neurons of different types (different types of pyramidal neurons and different types of interneurons) in different layers of the cortical circuit interact with one another to give rise to identified neural computations? For example, what roles do different types of interneurons play in contrast gain control? How do normalization circuits transform attentional feedback signals into improved sensory processing?

As detailed in Kevan Martin's talk, a great deal is known about the anatomical structure of the cortical laminar circuit. And, progress is being made in this direction using reduced preparations, such as the slice. For example, Massimo Scanziani described experiments suggesting how different types of interneurons are configured into very small circuits, or 'nanocircuits', involved in gain control. However, to understand how neural computations mediate perception and cognition, a necessary first step is to be able to differentiate between different classes of neurons in the awake behaving animals capable of performing tasks that enable the experimenter to control precisely information processing demands. Answering such questions is a challenge because it is difficult to differentiate among neural elements in behaving animals. We cannot visualize neurons *in vivo* as we can in the slice, because in order to ensure that the animal remains healthy throughout months of recording, we lower electrodes into the brain blindly through the dura, the protective tissue that surrounds the brain. Some steps have been made in this direction, primarily based on spike waveform (e.g., Swadlow, 2003; Mitchell, Sundberg & Reynolds, 2007), but additional methods, need to be developed to identify cell types and their positions in the laminar circuit in awake animals as they perform cognitive tasks. Promising approaches include: (1) targeted recording from genetically-defined neurons (e.g., Margrie et al., 2003; Liu et al., 2009); and (2) two-photon imaging using activity-dependent probes in genetically defined subsets of neurons (e.g., Sohya et al., 2007).

(3) Tools for controlling the canonical circuit in awake behaving animals.

Our understanding of the neural circuitry underlying canonical computations is also limited by a lack of tools to modulate the activity of specific neural circuits in the awake, behaving animal. Significant progress in testing hypotheses about the circuitry mediating neural computations, as well as the relationship between neural computations, perception and cognition, can be made if we can control the activity of different elements in the cortical circuit. Using electrical microstimulation, it is possible to directly activate neurons and test for changes in behavior. This technique makes it possible to alter activity with great temporal precision, but it has several major drawbacks. First, it is often unclear whether microstimulation mimics the normal activity of neurons or instead disrupts their activity. Second, in addition to causing orthodromic activation of neurons, microstimulation can also cause antidromic activation of neurons that project to the area of interest, as well as activation of fibers that simply pass through the area without making any synaptic contacts. Pharmacological techniques for altering neuronal activity can solve some of these problems but have other limitations. For example, injection of pharmacological agents

(e.g., muscimol, a $GABA_A$ agonist) can locally inhibit neurons without affecting fibers of passage. However, the specificity of this technique is limited because GABA receptors are present on many types of neurons. Thus, any behavioral effects observed with this technique cannot be attributed to a particular class of neurons, but can only be attributed to the general brain region affected by the injection. Finally, because the effects of such chemical agents generally last for many hours, the behavioral effects reflect the compensatory properties of the entire circuit, not the isolated function of the targeted neurons.

To probe the circuits that perform canonical computations, we need to be able to independently and precisely control the activity of different classes of neurons, and observe the consequences of this activation for computation as well as its effect on perception and cognition. To achieve this goal, we need to develop and apply molecular biological techniques, involving the use of viruses to cause selective expression of channels that will enable us to control the activity of specific types of neurons within the circuit (Luo, Callaway & Svoboda, 2008). These include "optogenetic" tools such as channelchannelrhodopsin-2 (ChR2). ChR2 is a transmembrane protein that contains a chromophore that, upon absorption of blue light, undergoes a conformational change, causing opening of the transmembrane channel. This results in an influx of cations, which leads to neuronal depolarization and generation of action potentials (Zhang et al., 2007). By selective expression of ChR2 in a particular type of neuron, one should be able to establish a causal link between the activity of the light-activated neurons, and resulting changes in the properties of the system within which they are embedded, such as changes in physiological properties or behavior.

Physiology: Mechanisms

Participants: Sarah Caddick, Matteo Carandini (coordinator), Michael Hausser, Kevan Martin, Nicholas Priebe, John Reynolds, Massimo Scanziani, Claudia Wiedemann, Rachel Wilson, Charles Yokoyama.

Introduction

Research in physiology has identified a number of canonical mechanisms and circuits, which are repeated across multiple brain regions for multiple purposes. The computational architecture of cortex, for example, is very similar from one area to another. The types, arrangements, and connections of cortical neurons are highly stereotyped. This suggests that each cortical area conducts calculations of the same form.

Examples of canonical circuits that have been proposed in cortex include the microcircuit for the cortical column (Douglas and Martin 2007a, 2007b), and the feedforward thalamocortical microcircuit (Gabernet et al. 2005). Examples of other canonical mechanisms whose computational roles are increasingly understood include synaptic plasticity (Abbott et al. 1997), dendritic computations (London and Hausser 2005), and the encoding of currents into spike trains (Priebe and Ferster 2008).

It may not always be possible to map mechanisms/circuits one to one onto neural computations, as several mechanisms and circuits may contribute to an individual computation. For example, divisive gain control in area V1 is likely to be due to a combination of multiple mechanisms: intracortical inhibition (Carandini et al. 1997), nonlinearities in thalamic responses (Priebe and Ferster 2006), depression at synapses from thalamus (Carandini et al. 2002) and within cortex (Abbott et al. 1997), and the integration of noisy inputs (Chance et al. 2002). An additional role may be played by mechanisms observed in other neural systems, such as lateral presynaptic inhibition observed in invertebrates (Olsen and Wilson 2008b).

Recent experimental advances promise to provide new insight into the hardware that mediates neural computations, and will likely lead to the discovery of unexpected properties of neural circuits. Neurophysiological methods now provide access to the circuitry that mediates neural computations. A particularly promising place to start these investigations is in invertebrates, where there is a reasonable hope to be able to obtain a complete characterization (Olsen and Wilson 2008a).

It has become practical to record signals from large populations of neurons using calcium imaging and electrode arrays, revealing the spatiotemporal profile of neuronal activity and the correlational structure of the neural population. It is also becoming increasingly feasible to record simultaneously from neurons in all cortical laminae, to characterize differences in the response properties of neurons in different cortical layers. It is feasible to distinguish among different anatomical classes of neurons (Mitchell et al. 2007) in awake primates as they perform cognitive tasks, providing insight into the different roles they play in mediating cognition. Finally, the emergence of new molecular-genetic approaches (Luo et al. 2008) enables experimentalists to control the activity of elements of the circuit.

Examples of open questions:

- What is the fundamental computation performed by a cortical column?
- What is the fundamental neural circuit underlying the computation?
- What roles do neurons in different layers of the cortical circuit play in this computation?
- What are the functions of different anatomical types of neurons in these computations?
- What are the cellular and synaptic mechanisms that implement that computation?
- How do the distinct biophysical properties of different classes of neurons subserve computation?
- What homeostatic and self-calibration mechanisms allow microcircuits to be adjusted in response to changes in environment, mental state, etc.
- How are the computational properties of local cortical circuits leveraged to mediate more complex forms of processing, e.g., working memory?

Summary of discussion

As we increase our understanding of canonical neural computations and of their normal and abnormal deployment, we need to map these computations onto mechanisms. Only by identifying these mechanisms we can hope to influence or correct the computations. For example, even if it were proven that a disease involves a malfunction in a divisive computation, one could not then go on and “treat the denominator”. Rather, one would need to identify and act upon the underlying mechanisms.

The mapping between computations and mechanisms will not generally be one-to-one. Some mechanisms may be fundamentally associated with a computation; for example, NMDA receptors for coincidence detection (e.g. of the kind seen in LTP), and the three types of cone receptors for trichromacy. However, in general a neural computation may well rely on multiple mechanisms. For example, the establishment of a receptive field has been shown in cat visual cortex to rely on a combination of at least 3 mechanisms: summation of appropriate thalamic inputs, spike threshold, and input from other cortical neurons. A similar case can be made for divisive normalization, which is likely to rest on multiple mechanisms.

One approach to identify the relevant mechanisms is to concentrate on well-understood instances of canonical computations and identify one by one the underlying mechanisms. For example, the establishment of orientation selectivity in layer 4 of cat area V1 can be seen as an instance of a canonical computation by which a brain region establishes receptive fields over the output of the preceding brain region. The inputs and outputs of this specific computation are

clear: spike trains of LGN neurons and spike train of a layer 4 neuron. Recent advances made in identifying the mechanisms that provide this transformation could generalize to other in other instances of the neural computation. For example, they could be at work in the transformation between layer 4 and layer 2/3 in primate visual cortex, or the connections between parallel fibers and Purkinje cells in cerebellum.

This functionally-driven approach puts the emphasis on areas where we understand the computations and on mechanisms that can be studies *in vivo*. Often, however, we want to understand areas where we don't fully know what information processing is being performed (e.g. hippocampus, higher visual areas), and often we are faced with mechanisms that are best understood in a preparation that is greatly simplified relative to a behaving animal (e.g. in vitro). In these cases we start from mechanisms and try to infer their function.

This approach has led to a mismatch between increasingly complex mechanisms that are being identified (e.g., the potential for complex computations that is available to even the smallest of the microcircuits) and the overall simplicity of the operations that we currently understand neural systems to be performing (e.g. summation, division, amplification). Scientists who follow the mechanism-led approach may tend to emphasize the potential complexity, whereas scientists who follow the computation-led approach may tend to identify as few and simple mechanisms as they can.

A hope is that neuroscience can achieve something similar to what was achieved in molecular biology, where the "secondary structure" has been identified as a fruitful level of abstraction to understand proteins, the level that provides the appropriate modules. The secondary structure specifies key protein substructures called domains (alpha-helix, random coil, beta sheet). These structurally distinct domains constitute canonical modules, which are simply repeated across proteins in different arrangements to obtain different function. Domains are not full explanation of proteins, e.g. they don't specify the protein's folding, but they are mightily useful in understanding how a protein operates. By contrast, a lower level of abstraction would be less useful. In particular, the "primary structure", i.e. the sequence of base pairs, constitutes a degenerate code: different combinations of base pairs can lead to similar protein domains. To the degree that understanding neural computation is similar to understanding protein function, we need to identify the modules that correspond to the secondary structure.

The main question is then what constitutes a building block, i.e. what level is the one where we can expect generalization. One way to conceptualize modularity is in terms of anatomy, and ask if the modules should be sought at the level of synapses, or dendrites, or whole neurons, or in combinations of neurons. It is likely, however, that the canonical mechanisms lie at multiple anatomical levels, and that their usefulness will lie in their being building blocks for systems-level models.

A wealth of candidate mechanisms is to be found at the subcellular level in mechanisms such as synaptic plasticity, spike generation, and dendritic computation. Dendrites, in particular, could in principle perform multiple computations: sum, multiply, divide, subtract, and estimate speed and direction. Recent work in barn owls revealed that dendrites can perform exquisite coincidence detection. Though the bipolar morphology of the neurons involved is highly specialized, the underlying principle could be exploited by a range of neuron types and could be at work in multiple sensory modalities and species, including mammals. Other cellular mechanisms may in turn have the role of simply undoing the "side effects" of dendrites, e.g., backpropagating action potentials. Finally, evidence for the existence of two spike generation zones in layer 5 pyramidal neurons may reflect canonical mechanisms in hippocampus and in other areas.

Another possibility is that the appropriate level of abstraction is cellular. For example, if we understand the operation of a basket cell in hippocampus, we can then go on and use that

information in neocortex. Indeed, basket cells in these two brain regions resemble each other: they receive a comparable pattern of connections, which result in comparable synaptic conductances. The role of specific cell types in computation could be addressed with targeted experiments in which a few neurons of a specified type are knocked out temporarily from a local circuit. For example, in seeking to understand the role of chandelier cells, which inhibit the axon initial segment, one could silence chandelier cells, but only the ones that affect one neuron, and ask what happens to functional responses. One may discover a resulting loss in a very specific computation (e.g. direction selectivity for a visual neuron).

Finally, the appropriate level of abstraction could well be that of circuits. There is ample evidence for a modularity of circuits. For example, Douglas and Martin proposed that the organization of neocortex could be summarized around a fundamental canonical microcircuit. Similarly, the cerebellar circuit is stereotyped: an input layer (granule cells) involved in “pattern separation” and a higher order layer (Purkinje cells) involved in “pattern storage”. Some circuits, in fact, may be more general than others. In particular, inhibitory circuits may be organized along more general principles than excitatory ones. For instance, basket cells seem to have a connectivity footprint that is consistently wider than that of principal neurons. This is seen in hippocampus, layer 4, layer 2/3. A difficulty with identifying the function of canonical microcircuits is that we could know everything about “average connectivity”, and yet come short of understanding what is going on because what matters is the specific pattern of connectivity on a neuron by neuron basis.

At the very least, these proposals for canonical microcircuits have the value of providing null hypotheses. In sensory cortex, for instance, there is disagreement over whether all areas are specialized or are fundamentally the same. A canonical microcircuit for cortex should be seen as a target against which one can test a specific circuit, as a way to establish a possible deviation from the norm.

It is also possible that the more fruitful level of analysis is that of smaller circuits, which we dub “nanocircuits”. For example, instead of trying to understand all layers of neocortex, a more reasonable goal may be to understand a portion of the circuit, e.g. recurrent excitation in layer 4. There are multiple examples of successful analysis of nanocircuits, and strong indications that these nanocircuits are indeed canonical. These examples include: (1) the ensemble formed by a pyramidal neuron and the “accessory” GABAergic interneurons that complement its function facilitating excitatory inputs; (2) the feedforward inhibition circuit made of thalamic neuron, a layer 4 neuron and an inhibitory interneuron. (3) the feedforward inhibition circuit made of a layer 4 neuron, a layer 2/3 neuron, and an inhibitory interneuron; (4) the feedback inhibition circuit, which provides a shift of inhibition from one compartment to the next and ensures proportionality of excitation and inhibition.

Further clues to the nature of fundamental modules may come from studies of development. For example, the cortex develops as an initially uniform sheet, which is then differentiated thanks to patterns of input. Understanding the rules that confer anatomical identity and the creation of local circuits based on this pattern of inputs may reveal the nature of the modules particularly clearly.

Given all these possible options and approaches, looking for a strict “canonical” circuit or set of mechanisms repeated across brain regions may not appear to be that helpful. After all there is a reason why different brain structures evolved: first cerebellum, then hippocampus, then neocortex. The circuits that were available (canonical or not) were evidently not sufficient for the new tasks at hand. Nonetheless, we must operate on the assumption that there are a finite (hopefully small) number of fundamental or canonical mechanisms, and that once we understand these modules we have obtained the building blocks that are used to build all neural

computation. If this assumption were incorrect, in trying to understand the brain we would be faced with a task of staggering complexity.

Theory and computation

Participants: Larry Abbott, Frances Chance, Adrienne Fairhall, Dario Ringach, Odelia Schwartz, Eero Simoncelli (Coordinator), Claudia Wiedemann.

Introduction

Computation has always been of central importance in neuroscience, as an essential ingredient in the analysis of neural data. As experimental methodologies have advanced (e.g., toward multi-electrode measurements, and more sophisticated molecular and genetic manipulations), the role of computation has become even more indispensable. New statistical methods are being developed to efficiently estimate the response properties of neurons and related methods have emerged for estimating the most likely sensory conditions given the responses of a neural population. Development of new forms of stimuli, and new methods of analysis/fitting can play an important role in connecting models of canonical computation with experimental data.

Theory and computational modeling in neuroscience have been somewhat slower to develop, but the appreciation of their role has steadily grown in recent years, partly fueled by programs funded by the Sloan and Swartz Foundations, and by computationally-oriented meetings such as Cosyne. Computational modeling plays a vital role in bridging the gap between neural data, environmental properties and behavior. Theory also has an important role to play, in terms of formulating precise hypotheses about the computational capabilities and underlying purpose of neural circuits.

In particular, normative theories for "why" neural circuits behave as they do provide important guidelines for the development of specific models. For example, normalization models have been proposed as serving purposes of optimal coding of sensory signal statistics (Ruderman and Bialek, 1994; Schwartz and Simoncelli, 2001; Fairhall et. al., 2001; Ringach and Malone, 2007). These ideas can be extended to learning and plasticity. Neurons in the brain are constantly adjusting their response characteristics based on their response history, and this adjustment occurs at a wide range of timescales (e.g., development, learning, adaptation, refractoriness). In order for the system to maintain stability (and usefulness) of computation, it would seem essential that these adjustments be coordinated according to some fundamental principles (Schwartz and Dayan, 2008; Wark et. al., 2009). A side benefit of these normative theories is that they have the potential to lead to development of a new generation of man-made machines and computing devices endowed with the flexibility found in biological processing, learning, and reasoning.

Some examples of open questions:

- In what sense are neural circuits "optimal"? Are they organized according to identifiable computational goals? What are the fundamental constraints that limit their capabilities (e.g., metabolic costs, wiring length or volume)?
- What computational operations should be included in a canonical circuit model, in order to strike a balance between mimicking the capabilities of neurons, and computational/analytical tractability?
- What are the computational capabilities of a network of such circuits? What are the factors that limit these capabilities?

- What principles are used during evolution/development/adaptation/homeostatis to optimize the parameters of these circuits? How much of this is "bottom-up", driven by the statistics of afferent activity, and how much of it is goal-specific?
- Given a model for a canonical circuit, what experimental measurements should be used to achieve a robust and reliable estimate of the parameters? What experiments should be used to distinguish between models, or to falsify a model?

Summary of discussion

Throughout the meeting, it was clear that theory and computation are playing (and must play) essential roles in advancing our understanding of canonical neural computation, and more specifically, in elucidating the nonlinear properties known as gain control, normalization, and adaptation. These roles include:

- Distillation of data, including data from different species, and different brain areas, into canonical computational forms. An example is the original publication of the normalization model (Heeger, 1992).
- Development of new stimuli and analysis methodologies that can be used by experimentalists to provide robust and reliable characterization of canonical model parameters. Nearly all of the experimentalists, and many of the theorists, at the meeting described work related to these issues.
- Analysis of the computational capabilities of canonical forms, and cascades of canonical forms. Most of the theory talks at the meeting addressed these kinds of issues. In addition, current activity in machine learning (by Geoff Hinton, Yann LeCun, etc) is aimed at understanding how to learn to solve complex recognition tasks using cascades of canonical populations.
- Development of principled (normative) theories for "why" neural circuits behave as they do. Talks by Abbott, Fairhall, Ringach, Heeger, Movshon, Schwartz, and Simoncelli all hit upon these issues.
- Development of principled theories for homeostatic adjustment (including gain control and adaptation phenomena) by which the brain maintains the computational integrity of canonical circuits in the face of noise and drift. Several talks addressed these issues, including Fairhall, Abbott, Chance, Movshon, and Simoncelli.
- Analysis methodologies for new experimental tools (multi-electrode measurements, imaging, molecular and genetic manipulations). Although this was not a central topic in the meeting, several participants are working on such problems.

The discussion within the theory/computation subgroup, which began in an after-lunch session on the first day, and continued on the second day, was initiated with the topic of definitions. In particular, Frances Chance asked "What is gain?", noting that the definition seems to vary substantially, and is often tied to a particular set of experimental protocols or stimuli! The same can be said for many other concepts that were central to the meeting: "gain control", "normalization", "adaptation", and even notions as ubiquitous as "noise" and "receptive field".

The last of these provides a nice example of how concepts in neuroscience are defined, and has provided a particularly good opportunity for participation by the theory community. The receptive field was introduced as a way of describing the spatial region of the retina over which a spot of light would elicit a response in a neuron. The definition was gradually elaborated to include other attributes, like the temporal region of the impulse response function, the frequency and phase characteristics of the response, etc. But the measurement of these attributes is done differently by each experimental lab, making it difficult to compare or unify results. Much of this elaboration is based implicitly on a *linear* model. This was eventually made explicit (and augmented with a cascade of a point nonlinearity and a Poisson spike generator), and a set of

tools (frequency analysis, and white-noise analysis) were developed to characterize this linear model.

The problem is that neurons are not linear (and the addition of a memoryless nonlinearity and spike generator are not sufficient to address the deficiencies). As a result, we find that the best-fitting linear kernel depends on the stimulus ensemble used to characterize the cell. Intuitively, we can say that the characterization is solving for a Taylor approximation of the cells response, averaged over the stimulus ensemble. Since the cell is nonlinear (more specifically, its nonlinearities are more complicated than the point nonlinearity in the model), the answer depends on the stimulus ensemble.

So the history of the receptive field is both a success story -- the definition of receptive fields became more precise over time, culminating in the embodiment in a quantitative model, coupled with characterization tools -- but also a cautionary tale. Although a quantitative model serves an essential role in creating a universal definition that allows comparison and compilation of experimental results, the limitations of the model can lead to failures of consistency. The same basic issues arise in the definitions of normalization, gain control, and adaptation. Many definitions and descriptions have been given, by both theorists and experimentalists, but we do not yet have a single universal form.

Much of our discussion focused on the need for precise models, along with characterization tools. A grand goal would be to define a set of fundamental computational modules, which could be assembled to mimic the behavior of any neural circuit, and an accompanying set of characterization tools, by which the structure and parameters of any given neural circuit could be estimated from experimental data. This theme was continued on the second day of discussion, when the theory/computation group merged with the physiology group.

Human and clinical

Participants: Steven C Dakin, David J Heeger (Coordinator), Concetta Morrone, David McAlpine, Anthony M Norcia, Joshua A Solomon, David Burr, Stefano Baldassi, Andrew F Rossi.

Introduction

To generalize from the canonical neural computations in animal models to humans requires methods for assessing these computations non-invasively. Normalization (cross-orientation and surround suppression), for example, has been assessed and characterized in human visual cortex using EEG and fMRI, in both infants and adults (Burr and Morrone 1987; Candy et al. 2001; Moradi and Heeger 2008; Zenger-Landolt and Heeger 2003). The behavioral consequences of such computations have been characterized as well (e.g., Watson and Solomon 1997).

It has been proposed that dysfunction or imbalance in canonical neural computations might cause the deficits underlying certain developmental disabilities (e.g., amblyopia, autism), neurological diseases (e.g., epilepsy) and mental illnesses (e.g., schizophrenia). If the canonical neural computation hypothesis is correct, then elucidation of the mechanisms and microcircuits underlying these neural computations is a fundamental mission whose success would simultaneously mark a new era in the scientific study of the brain and lead to new diagnostic procedures and treatments for disorders of the nervous system. For example, it has been hypothesized that autism spectrum disorders may be caused by a global imbalance of excitation and inhibition in cortical circuits (Rubenstein and Merzenich 2003; Markram et al. 2007; Jamain et al. 2008). Likewise, schizophrenia is associated not only with deficits in executive function but also with deficits in visual gain control and integration (Butler et al. 2008; Tadin et al. 2003, 2006).

It is important to note the fundamental difference between this approach and the prevailing view for how to characterize the neural basis of developmental disabilities and mental illnesses. Instead of focusing on deficits in a particular brain area or system (such as the "mirror system" in autism or the prefrontal cortex in schizophrenia), the idea here is that there may be a global (although perhaps subtle) deficit in neural computation that leads to a (perhaps dramatically) altered developmental trajectory. The end result may (or may not) be most evident in a particular brain area or system, but the underlying cause is a computational failure.

Examples of open questions:

- What measurements can be made (perceptual, MEG, fMRI, etc) to estimate the parameters of canonical computations in a human?
- How do canonical computations fail in various states of disease, and how do these computational failures contribute to the etiology of these brain diseases?
- What are the mechanisms by which these canonical computations fail?
- How do failures in canonical computations interact with development?
- Can we demonstrate improved or differential diagnoses based on this paradigm, and can we track efficacy of various (behavioral and/or pharmacological) interventions?

Summary of discussion

In our discussions, we outlined several challenges.

The first challenge is to bridge from healthy human brain function to physiology in animal models. The non-invasive methods for measuring human brain physiology (fMRI, EEG, MEG), although extremely useful, bear an uncertain relationship with invasive measurements of neural activity in animal models. This has been a topic of intense debate and controversy over the past decade, and substantial controversy remains. Suffice it to say that each of our methods, whether non-invasive in humans or invasive in animals, is subject to some bias. The path toward meeting this challenge surely will involve the application of multiple techniques, replicating results across techniques in human and animals at various spatial and temporal scales.

A second challenge is to bridge from the healthy human to disease/dysfunction. Characterizing a deficit in behavioral performance and/or brain activity is not enough. If it is not a specific deficit then it will not be possible to associate it with a specific computation. It is necessary as well to find some function that is spared, or even a task for which performance is improved. Dakin's work with schizophrenia represents a nice example of this, showing that schizophrenic patients exhibit less bias (on average) than controls at some visual appearance tasks (simultaneous contrast-contrast). This is hypothesized to be because of an underlying deficit in suppression (a weak denominator in the normalization model). Although this is a bad thing in general, it happens to lead to improved performance for some particular tasks. Such a sensory deficit might be at the root cause of disorder (e.g., the intense world hypothesis of autism), or it might be a marker of a global deficit that causes the disorder primarily because of its effect on computation in non-sensory brain areas (e.g., the sensory deficit in schizophrenia is just a marker for a computational failure in prefrontal cortex). Regardless, such sensory deficits could be useful markers of intervention.

A third challenge, alluded to above, concerns the heterogeneity of brain diseases and disorders. A particular deficit in neural computation might lead through development to any of a number of possible developmental trajectories with potentially very different outcomes. This may contribute to the great heterogeneity evident in many developmental disabilities (e.g., autism spectrum disorders, ASD) and mental illnesses (e.g., anxiety disorders). Might a common computational

deficit be the root cause of many of the variations of ASD or is the heterogeneity simply a failure of precise diagnostic criteria? Note that a common computation deficit might itself be caused by a very large number of possible genetic differences. For example, any number of genetic mutations that weaken cortical inhibitory networks might all lead to the same computational deficit (e.g., a weak denominator) which in turn could lead to a large number of behavioral dysfunctions because of the many possible developmental trajectories. Or is it the case that the heterogeneity is fundamentally due to underlying genetic differences. The path toward meeting this challenge will surely involve tracking the developmental trajectory. Even in late-onset disorders such as schizophrenia, it seems critical to perform longitudinal studies, characterizing neural computations at each stage of development.

A fourth challenge is to bridge to mechanisms. For disease intervention, it will be necessary to target the right mechanism. "Changing the exponent a bit" is not going to work. This is an even greater stretch than bridging from non-invasive measures of human brain function to physiology in animal models. Not only is it a greater distance conceptually but there may be multiple mechanisms contributing to the same piece of the computation (see summary of the "Mechanisms" group). Behavioral measures and non-invasive functional measures of brain activity may useful biomarkers, e.g., for diagnosis and/or to assess efficacy of intervention, even if we don't know the mechanism. But a connection to mechanisms will be needed to "design" interventions.

Although we are excited about the canonical computation hypothesis, research in this area will need to proceed cautiously. We have only a handful of parameters in our canonical computations (e.g., the strength of the denominator and the size of the exponent in the normalization model) but these various disorders and disabilities clearly express themselves very differently from one another. What is special/different about medial temporal lobe (MTL), for example, that makes it particularly susceptible to epilepsy. Although there is a high co-morbidity of ASD and epilepsy, why is it that some individuals have focal epilepsy without ASD (perhaps due to local imbalances in excitation and inhibition only in a circumscribed region of MTL) whereas as others develop ASD (perhaps due to widespread imbalances in excitation and inhibition) along with epilepsy.

We agreed that it would be worthwhile developing a suite of standardized assays to assess computational integrity across development, that could be applied to any different diseases and disorders. This would be most straightforward by focusing on sensory systems, for example, using behavioral measures of surround suppression to assess the denominator and super-additivity to assess the exponent. But it will also be important to extend beyond sensory systems. One possibility will be to capitalize on decision-making research in Paul Glimcher's lab at NYU in which they have been utilizing the normalization model to fit behavioral choice data in monkeys and single-unit firing rates in the lateral intraparietal area (LIP) of the macaque cortex.

References

Abbott LF, Varela JA, Sen K, and Nelson SB. Synaptic depression and cortical gain control. *Science* 275: 220-224, 1997.

Angelaki DE, Gu Y, Deangelis GC. Multisensory integration : psychophysics, neurophysiology and computation. *Curr Opin Neurobiol*, 2009.

Atiani S, Elhilali M, David SV, Fritz JB, Shamma SA. Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron* 61: 467-80, 2009.

Butler PD, Silverstein SM, and Dakin SC. Visual perception and its impairment in schizophrenia. *Biological psychiatry* 64: 40-47, 2008.

Burr DC, and Morrone MC. Inhibitory interactions in the human visual system revealed in pattern visual evoked potentials. *Journal of Physiology (London)* 389: 1-21, 1987.
Cadieu C, Kouh M, Pasupathy A, Connor CE, Riesenhuber M and Poggio M. A Model of V4 Shape Selectivity and Invariance. *J Neurophysiol* 98: 1733-1750, 2007.
Candy TR, Skoczenski AM, and Norcia AM. Normalization models applied to orientation masking in the human infant. *J Neurosci* 21: 4530-4541, 2001.
Carandini M. Amplification of trial-to-trial response variability by neurons in visual cortex. *PLoS Biol* 2: E264, 2004.
Carandini M, Heeger DJ, and Movshon JA. Linearity and normalization in simple cells of the macaque primary visual cortex. *J Neurosci* 17: 8621-8644, 1997.
Carandini M, Heeger DJ, and Senn W. A synaptic explanation of suppression in visual cortex. *J Neurosci* 22: 10053-10065, 2002.
Chance FS, Abbott LF, and Reyes AD. Gain modulation from background synaptic input. *Neuron* 35: 773-782, 2002.
Corrado GS, Sugrue LP, Seung HS, Newsome WT. Linear-Nonlinear-Poisson models of primate choice dynamics. *J Exp Anal Behav* 84: 581-617, 2005.
Dean I, Harper NS, and McAlpine D. Neural population coding of sound level adapts to stimulus statistics. *Nat Neurosci* 8: 1684-1689, 2005.
Depireux DA, Simon JZ, Klein DJ, and Shamma SA. Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *J Neurophysiol* 85: 1220-1234, 2001.
DiCarlo JJ, and Johnson KO. Receptive field structure in cortical area 3b of the alert monkey. *Behavioural brain research* 135: 167-178, 2002.
Douglas RJ, Martin KA. Neuronal circuits of the neocortex. *Annu Rev Neurosci* 27: 419-51, 2004.
Douglas RJ, and Martin KA. Mapping the matrix: the ways of neocortex. *Neuron* 56:226-238, 2007a.
Douglas RJ, and Martin KA. Recurrent neuronal circuits in the neocortex. *Curr Biol* 17: R496-500, 2007b.
Gabernet L, Jadhav SP, Feldman DE, Carandini M, and Scanziani M. Somatosensory integration controlled by dynamic thalamocortical feed-forward inhibition. *Neuron* 48: 315-327, 2005.
Heeger DJ. Normalization of cell responses in cat striate cortex. *Vis Neurosci* 9: 181-197, 1992.
Jamain S, Radyushkin K, Hammerschmidt K, Granon S, Boretius S, Varoqueaux F, Ramanantsoa N, Gallego J, Ronnenberg A, Winter D, Frahm J, Fischer J, Bourgeron T, Ehrenreich H, and Brose N. Reduced social interaction and ultrasonic communication in a mouse model of monogenic heritable autism. *Proc Natl Acad Sci U S A* 105: 1710-1715, 2008.
Kouh M, Poggio T. A canonical neural circuit for cortical nonlinear operations. *Neural Comput* 20:1427-1451, 2008.
London M, and Hausser M. Dendritic computation. *Annu Rev Neurosci* 28: 503-532, 2005.
Liu B, Li P, Li Y, Sun Y, Yanagawa Y, Obata K, Zhang LI, and Tao HW. Visual receptive field structure of cortical inhibitory neurons revealed by two-photon imaging guided recording, *J Neurosc*i 29: 10520-10532, 2009.
Luo L, Callaway EM, and Svoboda K. Genetic dissection of neural circuits. *Neuron* 57: 634-660, 2008.
Lyu S, and Simoncelli EP. Nonlinear image representation using divisive normalization, *Proc Comp Vision and Patt Rec* 1-8, 2008.
Mante V, Bonin V, and Carandini M. Functional mechanisms shaping lateral geniculate responses to artificial and natural stimuli. *Neuron* 58: 625-638, 2008.
Margie TW, Meyer AH, Caputi A, Monyer H, Hasan MT, Schaefer AT, Denk W, Brecht M. Targeted whole-cell recordings in the mammalian brain *in vivo*. *Neuron* 39: 911-918, 2003.
Markram H, Rinaldi T, and Markram K. The intense world syndrome - an alternative hypothesis for autism. *Frontiers in neuroscience* 1: 77-96, 2007.

Mitchell JF, Sundberg KA, and Reynolds JH. Differential Attention-Dependent Response Modulation across Cell Classes in Macaque Visual Area V4. *Neuron* 55: 131-141, 2007.
Moradi F, and Heeger DJ. Inter-ocular contrast normalization in human visual cortex. *J Vis* 9(3):13, 1–22, 2009.
Morgan ML, Deangelis GC, and Angelaki DE. Multisensory integration in macaque visual cortex depends on cue reliability. *Neuron* 59: 662-673, 2008.
Morrone MC, Ross J, and Burr D. Saccadic eye movements cause compression of time as well as space. *Nat Neurosci* 8: 950-954, 2005.
Movshon JA, Thompson ID, and Tolhurst DJ. Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *Journal of Physiology (London)* 283: 53-77, 1978.
Olsen SR, and Wilson RI. Cracking neural circuits in a tiny brain: new approaches for understanding the neural circuitry of Drosophila. *Trends Neurosci* 31: 512-520, 2008a.
Olsen SR, and Wilson RI. Lateral presynaptic inhibition mediates gain control in an olfactory circuit. *Nature* 452: 956-960, 2008b.
Priebe NJ, and Ferster D. Mechanisms underlying cross-orientation suppression in cat visual cortex. *Nat Neurosci* 9: 552-561, 2006.
Priebe NJ, and Ferster D. Inhibition, spike threshold, and stimulus selectivity in primary visual cortex. *Neuron* 57: 482-497, 2008.
Reynolds JH, and Chelazzi L. Attentional Modulation of Visual Processing. Ann Rev Neurosci 27:611-647, 2004.
Reynolds JH, and Heeger DJ. The normalization model of attention. *Neuron* 61: 168-185, 2009.
Ringach DL, and Malone BJ. The operating point of the cortex: neurons as large deviation detectors. *J Neurosci* 27: 7673-7683, 2007.
Rubenstein JL, and Merzenich MM. Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes Brain Behav* 2: 255-267, 2003.
Rust N, Schwartz O, Movshon JA, and Simoncelli EP. Spatiotemporal elements of macaque V1 receptive fields. *Neuron* 6(6): 945-956, 2005.
Rust N, Mante V, Simoncelli EP, and Movshon JA. How MT cells analyze the motion of visual patterns. *Nat Neurosci* 9: 1421-1431, 2006.
Rust N, and DiCarlo JJ. Increases in selectivity are offset by increases in tolerance ("invariance") to maintain sparseness across the ventral visual pathway, *Society for Neuroscience Annual Meeting*, Washington DC, Nov. 2008.
Schwartz O, and Simoncelli EP. Natural signal statistics and sensory gain control. *Nat Neurosci* 4: 819-825, 2001.
Simoncelli EP, and Heeger DJ. A model of neuronal responses in visual area MT. *Vision Res* 38: 743-761, 1998.
Solomon SG, Lee BB, and Sun H. Suppressive surrounds and contrast gain in magnocellular-pathway retinal ganglion cells of macaque. *J Neurosci* 26: 8715-8726, 2006.
Sohya K, Kameyama K, Yanagawa Y, Obata K, and Tsumoto T. BABAergic neurons are less selective to stimulus orientation than excitatory neurons in layer II/III of visual cortex, as revealed by *in vivo* functional Ca^{2+} imaging in transgenic mice, J Neurosci 27: 2145-2149, 2007.
Swadlow, H.A. Fast-spike interneurons and cerebral feedforward inhibition in awake sensory neocortex. *Cereb Cortex* 13: 25–32, 2003.
Tadin D, Lappin JS, Gilroy LA, Blake R. Perceptual consequences of centre-surround antagonism in visual motion processing. *Nature* 424: 312-315 2003.
Tadin D, Kim J, Doop ML, Gibson C, Lappin JS, Blake R, Park S. Weakened center-surround interactions in visual motion processing in schizophrenia. *J Neurosci* 26: 11403-11412, 2006.
Wainwright M, Schwartz O and Simoncelli EP. Natural image statistics and divisive normalization: Modeling nonlinearities and adaptation in cortical neurons. In *Statistical Theories of the Brain*, eds. R Rao, B Olshausen and M Lewicki, MIT Press. Spring 2002.
Wark B, Fairhall A, and Rieke F. Timescales of inference in visual adaptation. *Neuron* 61: 750-761, 2009.

Watson AB, and Solomon JA. Model of visual contrast gain control and pattern masking. *J Opt Soc Am A* 14: 2379-2391, 1997.
Zenger-Landolt B, and Heeger DJ. Response suppression in V1 agrees with psychophysics of surround masking. *J Neurosci* 23: 6884-6893, 2003.
Zhang F, Aravanis AM, Adamantidis A, de Lecea L, Deisseroth K. Circuity-breakers: optical technologies for probling neural signals and systems. *Nat Rev Neurosci* 8: 577-581, 2007.
Zoccolan D, Cox DD, and DiCarlo JJ. Multiple object response normalization in monkey inferotemporal cortex. *J Neurosci* 25: 8150-8164, 2005.