



Published in final edited form as:

*Annu Rev Neurosci.* 2018 July 08; 41: 77–97. doi:10.1146/annurev-neuro-080317-061936.

## Cognition as a Window into Neuronal Population Space

Douglas A. Ruff\*, Amy M. Ni\*, and Marlene R. Cohen

Department of Neuroscience and Center for the Neural Basis of Cognition, University of Pittsburgh, Pittsburgh, Pennsylvania 15213, USA

### Abstract

Understanding how cognitive processes affect the responses of sensory neurons may clarify the relationship between neuronal population activity and behavior. However, tools for analyzing neuronal activity have not kept up with technological advances in recording from large neuronal populations. Here, we describe prevalent hypotheses of how cognitive processes affect sensory neurons, driven largely by a model based on the activity of single neurons or pools of neurons as the units of computation. We then use simple simulations to expand this model to a new conceptual framework that focuses on subspaces of population activity as the relevant units of computation, uses comparisons between brain areas or to behavior to guide analyses of these subspaces, and suggests that population activity is optimized to decode the large variety of stimuli and tasks that animals encounter in natural behavior. This framework provides new ways of understanding the ever-growing quantity of recorded population activity data.

### Keywords

information coding; visual cortex; cognition; attention; spike count correlations; noise correlations

## INTRODUCTION

Cognitive processes that improve perception, such as attention, learning, and motivation or arousal, are unique tools in the study of the relationship between neurons and behavior.

Understanding what changes in the responses of sensory neurons when behavioral performance improves is key to understanding what aspects of the neural code are most important for encoding sensory stimuli and guiding behavior in the first place.

The effects of cognition on small numbers of sensory neurons are well studied. Numerous studies have shown that cognitive processes multiplicatively scale, or change the gain of, the trial-averaged responses of individual sensory neurons (Gilbert & Sigman 2007, Maunsell 2015). Recently, a growing number of studies have demonstrated that cognition also affects the response variability shared between pairs of neurons [termed spike count or noise correlations, or  $r_{SC}$ ] (for a review, see Cohen & Kohn 2011).

---

\*These authors contributed equally to this work.

### DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

However, linking these changes in small numbers of neurons to performance has been difficult. The relationship between single sensory neurons and behavior is weak and requires large amounts of data to quantify (Nienborg et al. 2012). Models have focused on whether response changes associated with cognition improve the amount of sensory information encoded by neuronal populations, but the relationship between single neurons, noise correlations, and information is complex (Kohn et al. 2016).

Most models designed to relate neuronal activity to perception are built on the idea that the individual neuron is the critical unit of neural computation. These models have been invaluable for generating hypotheses and integrating experimental data. However, while improvements in recording technology have made it possible to monitor the responses of larger populations of neurons, our understanding of neural coding has not kept pace with the technology. It has become clear that understanding the relationship between large populations and performance is not simply a matter of scaling up old analyses and models.

Instead, we propose that the relationship between sensory neurons and behavior requires a new conceptual framework and experimental and analytical strategy. We propose the following:

1. The relevant units of computation are arbitrary combinations of the responses of many neurons rather than average firing rates of groups of neurons with certain properties.
2. To gain insight into how these subspaces guide behavior, it is critical to compare population activity to either the activity of neuronal populations in other brain areas or the animal's behavior on a trial-to-trial basis.
3. Rather than optimally discriminating responses to a specific pair of stimuli in a laboratory task, animals naturally read out stimulus features using decoding strategies that would work for the large set of stimuli and behavioral tasks they encounter in natural vision.

Here, we describe the framework that has guided much of the research in this field, review insights from new experiments, and use novel simulations to investigate the implications of these new ideas.

## THE POOLING MODEL

The conceptual foundation for nearly all efforts to understand the relationship between sensory neurons and behavior is a simple model that is often referred to as the pooling model (Shadlen et al. 1996). This model has provided a rigorous basis for making and testing predictions about the relationship between neuronal activity and behavior, and it is arguably the primary reason that our understanding of the neural computations underlying perceptual decision making outpaces our understanding of the mechanisms underlying other systems and behaviors. Accordingly, the pooling model has had a broad influence across many fields of neuroscience (Brody & Hanks 2016, Carandini & Churchland 2013, Gold & Shadlen 2007, Heekeren et al. 2008).

The pooling model was developed to explain the responses of motion direction-selective neurons in the middle temporal area (MT) while rhesus monkeys performed a challenging perceptual discrimination task (Parker & Newsome 1998). In the basic conception of the model, the responses of groups of MT neurons encode sensory evidence in favor of each of the two possible behavioral choices. For example, in a left-right discrimination task, a pool of leftward-preferring MT neurons would provide evidence in favor of leftward choices and an analogous pool would provide evidence in favor of rightward choices. The model makes decisions by comparing the average responses of the neurons in the two pools.

For the study of cognition and behavior, the pooling model's greatest strength is that its simple conceptual framework generated many testable hypotheses about how neuronal activity might depend on cognition and how those changes might improve perception. In fact, although it is rarely cited for this inspiration, all the predominant hypotheses about how cognition improves perception can be thought of as coming from the framework of the pooling model.

## **PREDOMINANT HYPOTHESES ABOUT HOW COGNITIVE PROCESSES IMPROVE PERCEPTION**

### **Hypothesis 1: Cognitive Processes Improve Sensory Information Encoding**

**Support**—The first and, by far, most studied hypothesis is that cognitive processes improve performance by improving the amount of stimulus information that is encoded in the activity of a population of neurons. In the context of the pooling model, improving the signal-to-noise ratio of the mean rate of the neurons in each pool would create a larger difference between the responses of the two pools.

Essentially all studies that have focused on how cognitive processes affect the responses of neurons in one area of visual cortex have addressed this first hypothesis. Numerous studies spanning many visual cortical areas and tasks have revealed that directing attention to a particular location or feature increases the gain of the mean firing rates of neurons that are tuned for that location or feature (for reviews, see Anton-Erxleben & Carrasco 2013, Desimone & Duncan 1995, Maunsell 2015, Maunsell & Cook 2002, Maunsell & Treue 2006, Reynolds & Chelazzi 2004, Yantis & Serences 2003), and global cognitive processes such as arousal have similar effects (Boudreau et al. 2006). Attention (Cohen & Maunsell 2009, Mitchell et al. 2007) and learning (Raiguel et al. 2006) are also associated with modest decreases in the trial-to-trial variability of individual neurons. Both of these observations are consistent with the idea that cognition improves perception by increasing the signal-to-noise ratio of single neurons.

Additionally, the pooling model suggests that noise correlations can affect stimulus information because positively correlated noise cannot be averaged out by the pooled signal (Abbott & Dayan 1999, Averbeck et al. 2006, Shadlen et al. 1996). A seminal study by Zohary and colleagues (1994) made the now oft-replicated (Kohn et al. 2016) finding that noise correlations between pairs of MT neurons in the direction discrimination task tend to be small but positive.

Although recent studies suggest that the relationship between noise correlations and information coding is more complex than the simple idea that lower correlation is good (Kohn et al. 2016), many studies have shown that performance improves in situations in which noise correlations decrease. Predominantly, attention decreases noise correlations between neurons that have similar tuning properties (Cohen & Maunsell 2009, 2011; Gregoriou et al. 2014; Herrero et al. 2013; Luo & Maunsell 2015; Mayo & Maunsell 2016; Mitchell et al. 2009; Nandy et al. 2016; Ruff & Cohen 2014a; Verhoef & Maunsell 2017; Zénon & Krauzlis 2012). Other cognitive processes, such as arousal (Ruff & Cohen 2014b) and learning (Gu et al. 2011, Jeanne et al. 2013, Yan et al. 2014), are also associated with decreases in noise correlations.

The pooling model makes a second prediction about the role of noise correlations in decision making. In contrast to correlations between pairs of neurons in the same pool, positive noise correlations between neurons in different pools should help information coding, because shared variability can be subtracted out when the activity of the two pools is compared. We recently tested this hypothesis by recording from groups of neurons in area V4 that represented evidence in favor of both possible choices during a contrast discrimination task (Ruff & Cohen 2014a). Consistent with model predictions, we found that attention increased noise correlations between pairs of those same neurons when they were in different pools (while decreasing correlations within each pool). Together, these results suggest that attention and other cognitive processes change correlations in ways that are broadly consistent with the predictions of the pooling model.

**Limitations**—It is curious that so much evidence supports the prediction that cognitive processes improve information coding despite recent modeling efforts suggesting that only a very small subset of these changes to correlations should affect information at all (Kanitscheider et al. 2015b, Kohn et al. 2016, Moreno-Bote et al. 2014). This work shows that information coding is affected only by the small subset of correlated variability that aligns with the dimensions in population space in which the signal (e.g., motion direction) is encoded (called information-limiting correlations) when the neuronal population is large enough. This argument is based on the sensory information that could be gleaned by an optimal, high-dimensional decoder that has unlimited access to information about the stimuli being discriminated, as opposed to access to the average activity of pools of neurons. The intuition is that with large enough neuronal populations (and therefore very high-dimensional representations), correlated activity is likely to be orthogonal to (and easily separated from) the dimensions in which a particular stimulus feature is encoded. Whether animals can decode stimuli in this way is an open question that we address below.

Another probable limitation of the hypothesis that attention improves performance by improving information coding is that it relies on the idea that the amount of sensory information encoded in sensory cortex is what limits behavioral performance. However, numerous examples show that even small populations of neurons can encode far more information than a monkey appears to use (Parker & Newsome 1998). Therefore, cognition-related improvements in perception may instead, or may also, result from changes in other aspects of neural processing.

## Hypothesis 2: Cognitive Processes Improve Communication Between Cortical Areas

**Support**—An alternative explanation for the improvements in performance associated with cognitive processes is that cognition reduces the information that is lost (Sprague et al. 2015) when it is communicated to downstream areas (i.e., the areas that perform the pooling; Figure 1). Quantifying the amount of stimulus- or task-related information that is communicated between areas is difficult, but evidence suggests that attention increases the interdependence of neural activity in different areas across a range of timescales from both recordings (Bichot et al. 2005, Bosman et al. 2012, Buschman & Miller 2007, Fries 2015, Fries et al. 2001, Gregoriou et al. 2009, Lakatos et al. 2008, Miller & Buschman 2013, Saalman et al. 2007, Saproo & Serences 2014, Womelsdorf & Fries 2007, Womelsdorf et al. 2006) and causal manipulations such as electrical microstimulation (Briggs et al. 2013; Dagnino et al. 2015; Klink et al. 2017; Moore & Armstrong 2003; Ruff & Cohen 2016a, 2017). Furthermore, spatial attention increases noise correlations between neurons in different cortical areas (Oemisch et al. 2015; Pooresmaeili et al. 2014; Ruff & Cohen 2016a,b).

**Limitations**—In general, task-related changes in communication between areas, across all timescales, are small. This potentially limits their role in the framework in a simple pooling model, in which synchrony affects communication between different pools of neurons. Furthermore, there is something fundamentally strange about the observed increase in noise correlations between pairs of neurons in different areas while there is simultaneously a decrease in noise correlations between pairs of neurons within an area. It is tempting to think that the decreases in correlation within an area serve to improve the information represented in that area, which, in turn, is then more faithfully communicated to a downstream area via increased correlations. However, mathematical and biological constraints may limit the strength and importance of correlation changes whose sign is different within and across areas.

## Hypothesis 3: Cognitive Processes Improve the Way Sensory Information Is Decoded from Neuronal Populations

**Support**—At its most basic, the pooling model suggests that the readout of sensory information by a downstream neuron or a decision-related brain area simply involves comparing the average response of neurons in the two competing pools (i.e., that the responses of all neurons that belong to a pool are given the same weight). One possibility is that cognitive processes improve performance by changing the weighting function to make readout closer to optimal.

One of the most common tools used to make inferences about readout is choice probability, which measures the relationship between choices and the trial-to-trial fluctuations in an individual neuron's responses (Britten et al. 1996, Nienborg et al. 2012). This measure is, by definition, correlative, and there has been ample debate surrounding the origins of choice-predictive activity in sensory neurons (Cohen & Newsome 2009, Cumming & Nienborg 2016, Nienborg & Cumming 2009, Wimmer et al. 2015).

However, there is evidence consistent with the idea that the weighting of different neurons for decision making depends on their tuning. It has commonly been observed that neurons or voxels that are best tuned for the task at hand are also those with highest choice probability (Nienborg et al. 2012). Furthermore, choice probabilities of the most informative neurons increase throughout learning (Law & Gold 2008). More generally, researchers have suggested that a neuron's contribution to behavior can be changed by training (Chowdhury & DeAngelis 2008, Liu & Pack 2017).

**Limitations**—In many ways, the idea that readout weights are flexible has to be true. For the direction discrimination task, it may seem straightforward for the brain to divide neurons into appropriate pools based on their direction tuning, a feature which is neatly organized in columns in MT (Albright & Desimone 1987, Born & Bradley 2005). However, several factors suggest that readouts must involve flexible, nonuniform weights. Neurons are tuned for multiple stimulus features, there are many tasks and features where a clean division into pools would be complicated, and the mapping from stimuli to behavior is extremely flexible, making it unlikely that the brain solves tasks by cleanly dividing neurons into groups. A more realistic possibility is that decisions are based on flexibly weighted combinations (or subspaces) of the activity of the entire population (Cunningham & Yu 2014). In this scenario, cognitive processes such as attention could improve performance by making weightings more optimal.

The biggest limitation of this idea is the lack of experimental evidence. Although several studies have discussed the idea that subspaces of population activity, rather than groupings of neurons, are the important units of computation for readout (Churchland et al. 2012, Cunningham & Yu 2014, Elsayed et al. 2016, Elsayed & Cunningham 2017, Kaufman et al. 2014, Miri et al. 2017, Semedo et al. 2016, Yuste 2015), few studies have analyzed the relationship between population subspace activity and either the activity of downstream neurons or behavior. Some insights about how population subspace activity guides behavior have come from brain-machine interfaces (BMIs) in the motor system (Golub et al. 2016). BMI studies have shown that the representation of different actions and motor plans is relatively low dimensional and that animals cannot learn to access neuronal activity outside of key subspaces (Ganguly et al. 2011, Golub et al. 2016, Law et al. 2014, Sadtler et al. 2014). In sensory systems, the idea that decisions are based on population subspaces, not pools, has been slower to take hold (but see DiCarlo & Cox 2007, DiCarlo et al. 2012, Kriegeskorte 2009, Jazayeri & Afraz 2017, Pitkow & Angelaki 2017, Quian Quiroga & Panzeri 2009), but we believe that it will be of critical importance. A successful example of this approach has been in the study of visual object processing where population recordings and deep-learning neural networks have begun to elucidate the nature of high-level object representations (Cadieu et al. 2014, Khaligh-Razavi & Kriegeskorte 2014, Pagan et al. 2013, Yamins et al. 2014, Yamins & DiCarlo 2016).

## MOVING BEYOND PAIRS OF NEURONS: SUBSPACES OF POPULATION ACTIVITY AS THE UNITS OF NEURAL COMPUTATION

While technological advancements have led to rapid growth in the number of studies using multineuron recordings in behaving monkeys, our understanding has lagged behind these new data. We propose that progress has been slow because our data analysis techniques take the pooling model too literally and focus on single neurons or average firing rates in a pool as the units of neural computation. As explained below, there is reason to believe that focusing on population subspaces (Figure 2) will allow us to understand the relative importance of the three hypothesized neuronal mechanisms underlying cognitive processes and can resolve several paradoxes in the current literature.

In the face of correlated variability, it is difficult to use population recordings to infer the role, or weighting, of each neuron in the decision-making process and therefore to infer the amount of sensory information that is communicated to downstream readout areas. Simulations have shown that an animal's choices can be predicted from the responses of neurons that play no role in the decision but whose responses are correlated with those of neurons that do (Cohen & Newsome 2009, Nienborg et al. 2012). Inferring the contribution of different neurons in a population or the extent to which correlated variability affects the amount of stimulus information encoded in a population of neurons would require simultaneous recordings from many thousands of neurons over an even larger number of behavioral trials (Kanitscheider et al. 2015b, Kohn et al. 2016, Moreno-Bote et al. 2014). Although technology for recording from more and more neurons is ever improving, the work ethic of experimental subjects is not, and the number of behavioral trials required to infer the role of all of those neurons is prohibitive [which theoretical work suggests is an order of magnitude larger than the number of neurons (Bishop et al. 2017)].

We performed a simple simulation to gain intuition about what can and cannot be learned from recordings from small subsets of large populations of neurons over hundreds or thousands of trials. We made many simplifying assumptions, and the result is almost certainly not an accurate account of perceptual decision making in the brain. Our goal was illustrative and to suggest directions for future work.

Our simulation takes as its basis the published pooling model (Shadlen et al. 1996), but instead makes decisions based on arbitrary combinations of many neurons instead of the average firing rate of groups of neurons with different properties. We imagined that perceptual decisions about a single sensory feature (e.g., motion direction discrimination) are based on a linear combination of the responses of 5,000 neurons. The neurons in our simulation had cosine tuning for two sensory features (e.g., motion direction and binocular disparity). We imposed trial-to-trial noise correlations so that the variance of the spike count equaled the mean rate. Using previously published methods (Cohen & Maunsell 2009, Shadlen et al. 1996) and consistent with published observations (Cohen & Kohn 2011, Kohn & Smith 2005), we imposed noise correlations that were proportional to signal correlations (where signal correlations reflected mean responses to all combinations of the two sensory features).

As in the original pooling model, our simulation made decisions based on the combined responses of many neurons. However, our simulation incorporated decisions based on the activity of subspaces of the population. Population subspaces can be thought of as linear combinations of large numbers of neurons. We therefore imposed flexible weighting: We assigned each neuron a weight (which could be positive, negative, or 0) using one of several candidate weighting functions and based decisions on whether the weighted average response of the population was positive or negative. We then selected random subsets of 100 neurons to determine what we could and could not learn from the responses of this subset of simulated neurons and the simulation's behavior on each trial.

Our simulations give good reason for optimism, and this optimism can be tested in future experiments. The responses of small populations can be used effectively to understand important qualities of the entire simulated neural population, can be used to evaluate the relative importance of the three hypothesized ways that cognitive processes improve perception, and can offer a resolution to a longstanding paradox about whether animals can optimally decode sensory information from neuronal populations. Put another way, the detailed weightings of each neuron in a large population are not important: Many weightings achieve similar performance in simple tasks. In this situation, small populations can be used to distinguish between different decoding strategies without identifying the precise neuron weights. Whether this will hold in systems with complex dynamics or nonlinear decoding strategies remains to be seen. But these simulations suggest that the hypothesis that subspaces of the responses of populations of neurons are the units of neural computation will be testable in experimentally feasible data sets.

### **Insights Related to Hypothesis 1: Small Populations Can Reveal the Stimulus or Behavioral Information that Is Encoded by or Communicated Between Neural Populations**

The first hypothesis posits that cognitive processes improve perception by improving the amount of sensory information encoded by a neural population. Calculating information in neural populations is tricky because it relies on assumptions about the readout functions that can be used in the brain (da Silveira & Berry 2014, Kanitscheider et al. 2015a, Kohn et al. 2016). We discuss this issue separately below.

A related hypothesis that is more straightforward to address is that cognitive processes improve the sensory information that is used to make decisions or that is communicated to downstream areas. We tested this idea in our simulation by using linear regression to infer the weights of the 100 selected neurons that would allow us to best predict the simulation's choices (which are based on the entire population) on one set of 500 trials and then calculating the simulation's performance at a different simulated stimulus coherence using those weights on a separate set of 500 trials.

This simulation showed that although the discrimination performance of small populations is worse than that of the entire simulated population, the performance of decoders based on small populations covaries with the performance of the population. Figure 3a plots the proportion of correct discriminations using the weighted sum of the responses of the 100 selected neurons and of the entire population. Although the full population outperformed the selected subset of neurons, the similar relationship between the full and subpopulations and



coherence suggests that recordings from small populations can be used to determine whether cognitive processes have improved the sensory information used to guide an animal's choices.

### **Insights Related to Hypothesis 2: Small Populations Can Reveal the Strength of Communication Between Subspaces of Different Neuronal Populations**

The second hypothesis posits that cognitive processes change the efficiency or strength of functional communication between brain areas.

Detecting such changes requires simultaneous recordings from multiple areas, and a measure of functional communication. Previous studies have used  $r_{SC}$  between neurons in different areas as such a measure (Ruff & Cohen 2016a, Klink et al. 2017, Poort et al. 2016).

To illustrate how changes in the number or activity of shared or independent inputs might change cross-area  $r_{SC}$ , we extended our simulation to include a second population of neurons. We simulated the activity of each neuron in the second population as a linear combination of the responses of neurons in the first population (shared inputs) and independent inputs (that could come from any other area).

In this scenario, when the number or activity of independent inputs is small (so most inputs come from the first population), cross-area  $r_{SC}$  is high (Figure 3b, small numbers on the  $x$ -axis). When the number or activity of independent inputs is large, cross-area  $r_{SC}$  is low (large numbers on the  $x$ -axis).

In real recordings, correlations between projections onto subspaces of population activity in each area may be even better measures of functional connectivity. Standard statistical methods like canonical correlation analysis may be especially useful for identifying subspaces of activity that are shared between populations (Semedo et al. 2016) and determining how the strength of functional communication depends on cognitive processes like attention.

### **Insights Related to Hypothesis 3: Small Populations Can Reveal the Weighting Function Used to Make Decisions but Not the Weights of Particular Neurons**

The third hypothesis posits that cognitive processes improve performance by changing the way that sensory information is read out to guide behavior (e.g., by changing the weighting, or contribution, of each neuron). Our simulation suggests that while recordings from small populations cannot be used to infer the weightings of any particular neurons, they can be used reliably to detect the distribution of weights and therefore whether that distribution changes. We used two weighting schemes to make our simulation make choices based on the responses of the full population. The first scheme is the one used in the pooling model: Each neuron had a weight of  $-1$ ,  $0$ , or  $1$  depending on whether its responses contributed to the first pool, no pool, or the second pool. In the second scheme, we used the weights that would comprise an optimal linear decoder for this stimulus feature (obtained using linear regression). We then tested subsets of 100 randomly selected neurons on 1,000 randomly selected trials and used linear regression to infer a set of weights that would allow us to best predict the simulation's choices from the responses of the 100 selected neurons as above.

The results suggest that recordings from small populations on reasonable numbers of trials can easily distinguish between the two weighting schemes. Figure 3c shows the distributions of inferred weights in the pooling and optimal schemes. A Hartigan's dip test found significant bimodality in the inferred weights in 88% of the 1,000 resampled small populations using the pooling scheme and only 4.8% of resampled populations using the optimal scheme. Furthermore, the inferred weights correlated with each individual neuron's ability to distinguish two adjacent stimuli (quantified as the  $d'$  between responses to the two stimuli over 1,000 trials) in the optimal but not the pooling scheme ( $R = 0.33$ ,  $P < 10^{-5}$  in the optimal scheme;  $R = 0.01$ ,  $P = 0.38$  in the pooling scheme).

However, our simulations indicate that recordings from small populations cannot be used to infer the actual weightings of individual neurons. The correlation between the inferred and actual weights was 0.04 ( $P = 0.12$ ) in the pooling scheme and 0.03 ( $P = 0.18$ ) in the optimal scheme.

## USING POPULATIONS TO FIGURE OUT THE ROLE OF NOISE CORRELATIONS

Together, our simulations suggest that recordings from small subsets of a large population can provide important insights about the mechanisms underlying cognitive improvements in perception. However, these simulations are gross oversimplifications of the complexity of real neuronal population data. We therefore provide an example from our own work of the insights that can be gleaned by considering population subspaces.

The role of correlated variability in limiting performance on perceptual tasks has been under heavy debate. Theoretical studies show that the bulk of correlated variability should not affect the amount of information encoded by a neuronal population because it is not part of the subspaces of activity (oriented along the same dimensions) that are read out by optimal stimulus decoders. However, this finding seems at odds with the large number of studies showing that cognitive processes such as attention, learning, and arousal change correlations (Cohen & Maunsell 2009, 2011; Gregoriou et al. 2014; Gu et al. 2011; Herrero et al. 2013; Jeanne et al. 2013; Luo & Maunsell 2015; Mayo & Maunsell 2016; Mitchell et al. 2009; Nandy et al. 2016; Ruff & Cohen 2014a,b, 2016a, Verhoef & Maunsell 2017, Yan et al. 2014, Zenon & Krauzlis 2012). We recently analyzed subspaces of simultaneously recorded V4 neuronal population activity to show that noise correlations are much more closely aligned with the population subspaces that matter for behavior than theoretical studies suggest they should be (Ni et al. 2017).

Noise correlations are by definition calculated over many trials. Determining the relationship between correlated variability and individual choices required us to derive a single trial measure of correlated variability from the activity of the populations of V4 neurons we recorded while monkeys performed an orientation change-detection task. We used principal component (PC) analysis on population responses to repeated presentations of the same stimulus to identify the axis in population space that accounted for the most correlated variability. The variance explained by the first PC during each recording session was highly correlated with the mean noise correlation of all neuron pairs recorded during that session,

consistent with recent observations that noise correlations are typically low dimension (Ecker et al. 2016; Goris et al. 2014, Kanashiro et al. 2017, Rabinowitz et al. 2015) (Figure 4a).

To determine the relationship between this correlated variability axis and the monkey's choices, we projected the population responses to the stimulus the monkey was charged with detecting onto the PCs calculated as described above. We then compared those projections between trials in which the monkey correctly versus incorrectly detected the stimulus.

We found that we could predict the monkey's choices using just the first PC as well as we could from the entire recorded population response (Figure 4b, *green line*). Put another way, projections onto the first PC, our single trial measure of correlated variability, explained essentially all the available choice-predictive activity in the recorded neuronal population.

This relationship between correlated variability and the performance of the choice decoder is particularly striking when compared to the relationship between correlated variability and the performance of an optimal stimulus decoder. The stimulus decoder (Figure 4b, *black line*) distinguished between the population responses to the two stimuli. Unlike the choice decoder (Figure 4b, *green line*), the stimulus decoder's performance improved significantly when based on more PCs, meaning that it was not as influenced by correlated variability. This is in line with a prior study that found that while correlated variability decreased with training on a perceptual task, those changes in correlated variability had little effect on the population coding efficiency of an optimal stimulus decoder (Gu et al. 2011).

The schematic in Figure 4c illustrates a potential scenario that could explain why correlated variability has a different relationship with choice and stimulus decoders. The choice decoding axis may be aligned with the correlated variability axis (the first PC), while the stimulus decoding axis may be based on multiple PCs. This finding suggests that monkeys are suboptimal in a very specific and perplexing way: Their decisions are aligned with the axis of correlated variability. Below, we discuss a potential explanation: The animals may be optimal, but for something other than the very specific task they performed in each trial in the lab.

## A NEW HYPOTHESIS: OPTIMALITY FOR GENERALITY

We propose an explanation for the relationship between the effects of cognitive processes on sensory neurons (which by all theoretical accounts should have a very limited effect on information coding) and behavior: that animals perform optimally, but for the much more general set of tasks they encounter in the natural world rather than the limited task they perform in the lab. We hypothesize that the relationship between signal and noise correlations that has been observed in many cortical areas means that this general sensory readout is truly affected by correlated variability and by the mean rates of sensory neurons.

The logic behind most conventional analyses [e.g., computing neurometric thresholds in single neurons (Britten et al. 1992, Parker & Newsome 1998) or decoding stimulus information from neuronal populations] is that there could be a new decoder (i.e., new weights) for each pair of stimuli. In the context of the motion direction discrimination task,

one decoder is set up to discriminate 3% coherence leftward motion from 3% coherence rightward motion, and a separate decoder is set up to discriminate 6% coherence leftward motion from 6% coherence rightward motion. (This is the way we set up the decoding in Figure 3a.) In a fine discrimination task, the decoders might be very different for each stimulus pair: The neurons that would be most useful for discriminating a difference between 30°- and 32°-oriented gratings are different from the neurons that would be most useful for discriminating between 120°- and 122°-oriented gratings.

This scenario is not suitable for natural vision. If you are about to cross a busy street, you need to determine whether a car is headed your way, no matter its starting position or its features. Features such as color, size, or shape are irrelevant for motion discrimination but still modulate the responses of the same MT neurons that guide motion discrimination. Likewise, the many-decoders scenario is not likely to be true even for simple laboratory tasks. For monkeys to use a decoder that depends on the stimulus implies that they use a two-step decision process: identifying the stimulus followed by doing the actual discrimination. This is extremely unlikely; if animals could successfully identify the stimulus, there would be nothing left to discriminate.

A much more realistic scenario is that animals use a more general decoder that is at least capable of discriminating arbitrary pairs of the stimuli used in any task. For example, animals performing the motion direction discrimination task might use a general motion decoder that can discriminate any moving object.

As a decoder gets more general (i.e., it has to contend with stimuli that vary in more feature dimensions), its weights depend on more and more tuning properties. Consider, for example, two rightward-selective MT neurons in the context of a left-right motion direction discrimination task. If they have the same binocular disparity tuning, then a large response from them might indicate rightward motion, near disparity, or a combination of those two features. This is irrelevant if the decoder considers only stimuli with identical disparity, but a more general decoder would need to resolve this discrepancy by choosing weights of these and other neurons that take their disparity tuning into account.

Therefore, the weights in a more general decoder would depend on the tuning of all neurons to all stimulus features to which they respond, which means that decoding weights would depend on exactly the same factors as noise correlations. It is well established that noise correlations depend on tuning similarity for all features. For example, spike count correlations between pairs of MT neurons depend on tuning similarity for motion direction (Bair et al. 2001, Cohen & Newsome 2008, Solomon et al. 2015, Zohary et al. 1994), speed (Huang & Lisberger 2009), sensory normalization (Ruff et al. 2016), and, as has been shown in other areas, cortical distance (Smith & Kohn 2008, Smith & Sommer 2013).

This dependence of signal and noise on the same feature set means that the optimal decoder might be aligned with the population subspaces containing the most correlated variability (as in Figure 4). Put another way, correlated variability is in a position to have a large effect on the performance of the generalized optimal decoder.

The idea that the weights of a general decoder depend on the same features as noise correlations explains some confusing results in the literature. Noise correlations can either help or hurt the performance of a specific optimal decoder (Abbott & Dayan 1999, Averbeck et al. 2006), but the vast majority of studies report that attention and other cognitive processes decrease correlations overall (Cohen & Maunsell 2009, Gregoriou et al. 2014; Gu et al. 2011; Herrero et al. 2013; Jeanne et al. 2013; Luo & Maunsell 2015; Mayo & Maunsell 2016; Mitchell et al. 2009; Nandy et al. 2016; Ruff & Cohen 2014a, Verhoef & Maunsell 2017; Yan et al. 2014; Zenon & Krauzlis 2012). Similarly, attention and other cognitive processes typically increase the trial-averaged response gains of all neurons, regardless of their tuning for the specific stimuli to be discriminated (Maunsell 2015, McAdams & Maunsell 1999). Both results make sense if a generalized, rather than a specific, decoder is being optimized.

## DESIGNING EXPERIMENTS TO TEST THE IDEA OF OPTIMALITY FOR GENERALITY

What would experimental evidence for the idea of a general decoder look like? Our simulations suggest that experiments should incorporate a richer stimulus set or behavioral task, but that a small amount of richness can go a long way.

To illustrate this idea, we adapted our simulation to make the neurons tuned for two features [i.e., direction and disparity, two features whose tuning preferences in MT are independent and largely separable (DeAngelis & Uka 2003, Smolyanskaya et al. 2013)]<sup>1</sup>, and we had the simulation discriminate the direction of stimuli whose disparity varied. For example, the simulation needed to discriminate rightward from leftward motion in the face of stimuli with highly variable disparities. The neuron weights of the optimal decoder that has to contend with variable disparity are only weakly correlated with those of a decoder that works only on stimuli with constant disparity ( $R = 0.18$ ; Figure 5a). Furthermore, changing the mean noise correlation across the population from 0.1 to 0.05 (which is in the range of attention- and learning-related changes in mean  $r_{SC}$ ) had a much greater effect on the decoder that contended with multiple disparities than on the decoder that dealt only with constant disparity (compare the left and right sets of bars in Figure 5b). This large effect of changing noise correlations on the general decoder's performance is much more in line with our prior experimental findings than the effect of noise correlations on the decoder that only dealt with constant disparity. We found a robust, consistent relationship between changes in noise correlations and changes in behavioral performance, whether those changes occurred quickly with attention or slowly with learning (Ni et al. 2017). Our simulation suggests that this relationship between changes in noise correlations and behavioral performance is better explained by a general decoder.

These simulations suggest that adding a small amount of complexity to laboratory tasks might provide qualitatively new insights. It is not necessary to move to natural stimuli and uncontrolled behaviors. Instead, our simulations suggest that using stimuli that vary in a single, task-irrelevant stimulus feature or adding a small amount of complexity to the behavior is sufficient to distinguish between specific and general decoders. For example,

simple statistical techniques to infer the neuron weights that best explain activity in another brain area or the animal's choices would give very different answers if the true decoder were specific or general.

Our simulations suggest that general decoders may provide a superior account of the effects of cognitive processes like attention and learning on behavior. For this reason, determining whether animals in fact use general decoders and the extent to which their decoders depend on the particular stimuli and task conditions will be an important avenue for future work.

## EXPERIMENTAL EVIDENCE FOR AND AGAINST OPTIMALITY

Even though there are practical benefits of decoding a feature in a general enough way to work for all stimuli, it is reasonable to hypothesize that a trained animal might optimize its decoder for the set of stimuli or task conditions it encounters in the lab. Indeed, a few studies have shown that the effects of attention or learning on spike count correlations depend on the relationship between the tuning of the neurons and the stimulus dimension being discriminated (Jeanne et al. 2013, Ruff & Cohen 2014a, Verhoef & Maunsell 2017). These results suggest that the effects of cognitive processes on sensory neurons can be optimized for a specific task. However, attention was associated with overall decreases in average correlation in addition to the tuning-specific effects in two of those studies (Ruff & Cohen 2014a, Verhoef & Maunsell 2017), suggesting that the extent to which decoders can be optimized is limited.

A study by Clery and colleagues (2017) found that while neurons in area V2 exhibited decision-related activity in a fine disparity discrimination task, the relationship between the neuronal responses and the monkeys' psychophysical performance levels was not compatible with an optimal linear readout of available sensory information. Even when they restricted their analysis to sessions in which the psychophysical performance of the monkeys exceeded neuronal sensitivity, they monkeys did not appear to read out the sensory information in V2 optimally.

On the other hand, numerous studies have shown that in certain situations, subjects do appear to behave optimally based on the available sensory information (Drugowitsch et al. 2014, Ernst & Banks 2002, Fetsch et al. 2011, Jacobs 1999, Knill 2007, Pitkow et al. 2015). Provocatively, many of these studies suggesting optimal decoding strategies have used multisensory stimuli or otherwise involved the integration of multiple modalities of sensory information (Chandrasekaran 2017, Kording & Wolpert 2006). In a study by Fetsch and colleagues (2011), monkeys demonstrated the ability combine visual and vestibular inputs in a near-optimal manner, with the activity of multimodal neurons in the dorsal medial superior temporal area giving a fair account of this behavioral optimization.

One possibility is that animals always read out stimulus information as if the stimuli vary across multiple task-relevant and -irrelevant feature dimensions, regardless of whether the task requires them to do so. Consistent with this idea, in our simulation, merely accounting for a second additional feature dimension makes readout appear nearly optimal. In this case, behavior would approach optimality for studies using multisensory stimuli.

There might be a biological explanation for why even well-trained animals might settle on a strategy that uses a generalized decoder. We and our collaborators recently showed that the simplest way to modulate spike count correlations and rates in cortical circuit models in ways that are consistent with physiological data is to change the balance of excitation and inhibition (Huang et al. 2017, Kanashiro et al. 2017). For example, modulating inhibition more than excitation increases response gain and decreases correlated variability in ways that are consistent with data from attention tasks. Perhaps this sort of simple mechanism by which cognitive processes affect visual cortex is easier to implement in biological circuits than the complex weight changes required for stimulus- or task-specific cognitive effects.

## CONCLUSION

Much of our understanding of the neuronal basis of behavior is based on a framework that focuses on the activity of single neurons or pools of neurons as the units of neural computation. However, true population activity is far richer, and we propose that considering subspaces of population activity as the relevant units of computation and considering the wide variety of stimuli and tasks that animals contend with in natural behavior will allow us to gain a better understanding of neuronal computations. We suggest methods for finding the population subspaces relevant to behavioral performance: comparisons of population activity to trial-by-trial behavioral choices, or comparisons to the activity of neuronal populations in other brain areas. This conceptual framework for understanding how neuronal activity relates to behavior will allow our analytical techniques to keep pace with our ever-growing technical abilities.

This is an especially exciting time to be a neuroscientist interested in understanding the relationship between the activity of neurons and behavior. Standing on the shoulders of decades of important experimental and quantitative work, the field is ideally positioned to ask critical questions about this relationship. We believe, with the guidance provided by this new theoretical framework bolstered by the use of bleeding edge technology, that we have never been better positioned to find these answers.

## ACKNOWLEDGMENTS

We thank Faisal Baqai for helpful comments on an earlier version of this manuscript. M.R.C. is supported by US National Institutes of Health grants 4R00EY020844-03, R01 EY022930, and Core Grant P30 EY008098s; a Whitehall Foundation grant; a Klingenstein-Simons Fellowship; Sloan Research Fellowship; a McKnight Scholar Award; and a grant from the Simons Foundation. A.M.N. is supported by a fellowship from the Simons Foundation.

## LITERATURE CITED

- Abbott LF, Dayan P 1999 The effect of correlated variability on the accuracy of a population code. *Neural Comput.* 11:91–101 [PubMed: 9950724]
- Albright TD, Desimone R 1987 Local precision of visuotopic organization in the middle temporal area (MT) of the macaque. *Exp. Brain Res.* 65:582–92 [PubMed: 3556486]
- Anton-Erxleben K, Carrasco M 2013 Attentional enhancement of spatial resolution: linking behavioural and neurophysiological evidence. *Nat. Rev. Neurosci.* 14:188–200 [PubMed: 23422910]
- Averbeck BB, Latham PE, Pouget A 2006 Neural correlations, population coding and computation. *Nat. Rev. Neurosci.* 7:358–66 [PubMed: 16760916]

- Bair W, Zohary E, Newsome WT 2001 Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J. Neurosci.* 21:1676–97 [PubMed: 11222658]
- Bichot NP, Rossi AF, Desimone R 2005 Parallel and serial neural mechanisms for visual search in macaque area V4. *Science* 308:529–34 [PubMed: 15845848]
- Bishop WE, Degenhart AD, Oby ER, Batista AP, Chase SM, et al. 2017 Extracting stable representations of neural population state from unstable neural recordings. *Cosyne Abstracts 2017*, Salt Lake City, UT
- Born RT, Bradley DC 2005 Structure and function of visual area MT. *Annu. Rev. Neurosci.* 28:157–89 [PubMed: 16022593]
- Bosman CA, Schoffelen J-M, Brunet N, Oostenveld R, Bastos AM, et al. 2012 Attentional stimulus selection through selective synchronization between monkey visual areas. *Neuron* 75:875–88 [PubMed: 22958827]
- Boudreau CE, Williford TH, Maunsell JHR 2006 Effects of task difficulty and target likelihood in area V4 of macaque monkeys. *J. Neurophysiol.* 96:2377–87 [PubMed: 16855106]
- Briggs F, Mangun GR, Usrey WM 2013 Attention enhances synaptic efficacy and the signal-to-noise ratio in neural circuits. *Nature* 499:476–80 [PubMed: 23803766]
- Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA 1996 A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* 13:87–100 [PubMed: 8730992]
- Britten KH, Shadlen MN, Newsome WT, Movshon JA 1992 The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* 12:4745–65 [PubMed: 1464765]
- Brody CD, Hanks TD 2016 Neural underpinnings of the evidence accumulator. *Curr. Opin. Neurobiol.* 37:149–57 [PubMed: 26878969]
- Buschman T, Miller E 2007 Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315:1860–62 [PubMed: 17395832]
- Cadiou CF, Hong H, Yamins DL, Pinto N, Ardila D, et al. 2014 Deep neural networks rival the representation of primate IT cortex for core visual object recognition. *PLOS Comput. Biol.* 10:e1003963
- Carandini M, Churchland AK 2013 Probing perceptual decisions in rodents. *Nat. Neurosci.* 16:824–31 [PubMed: 23799475]
- Chandrasekaran C 2017 Computational principles and models of multisensory integration. *Curr. Opin. Neurobiol.* 43:25–34 [PubMed: 27918886]
- Chowdhury SA, DeAngelis GC 2008 Fine discrimination training alters the causal contribution of macaque area MT to depth perception. *Neuron* 60:367–77 [PubMed: 18957227]
- Churchland MM, Cunningham JP, Kaufman MT, Foster JD, Nuyujukian P, et al. 2012 Neural population dynamics during reaching. *Nature* 487:51–56 [PubMed: 22722855]
- Clery S, Cumming BG, Nienborg H 2017 Decision-related activity in macaque V2 for fine disparity discrimination is not compatible with optimal linear readout. *J. Neurosci.* 37:715–25 [PubMed: 28100751]
- Cohen MR, Kohn A 2011 Measuring and interpreting neuronal correlations. *Nat. Neurosci.* 14:811–19 [PubMed: 21709677]
- Cohen MR, Maunsell JHR 2009 Attention improves performance primarily by reducing interneuronal correlations. *Nat. Neurosci.* 12:1594–600 [PubMed: 19915566]
- Cohen MR, Maunsell JHR 2011 Using neuronal populations to study the mechanisms underlying spatial and feature attention. *Neuron* 70:1192–204 [PubMed: 21689604]
- Cohen MR, Newsome WT 2008 Context-dependent changes in functional circuitry in visual area MT. *Neuron* 60:162–73 [PubMed: 18940596]
- Cohen MR, Newsome WT 2009 Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *J. Neurosci.* 29:6635–48 [PubMed: 19458234]
- Cumming BG, Nienborg H 2016 Feedforward and feedback sources of choice probability in neural population responses. *Curr. Opin. Neurobiol.* 37:126–32 [PubMed: 26922005]



- Cunningham JP, Yu BM 2014 Dimensionality reduction for large-scale neural recordings. *Nat. Neurosci.* 17:1500–9 [PubMed: 25151264]
- da Silveira RA, Berry MJ 2014 High-fidelity coding with correlated neurons. *PLOS Comput. Biol.* 10:e1003970
- Dagnino B, Gariel-Mathis M-A, Roelfsema PR 2015 Microstimulation of area V4 has little effect on spatial attention and on perception of phosphenes evoked in area V1. *J. Neurophysiol.* 113:730–39 [PubMed: 25392172]
- DeAngelis GC, Uka T 2003 Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *J. Neurophysiol.* 89:1094–111 [PubMed: 12574483]
- Desimone R, Duncan J 1995 Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18:193–222 [PubMed: 7605061]
- DiCarlo JJ, Cox DD 2007 Untangling invariant object recognition. *Trends Cogn. Sci.* 11:333–41 [PubMed: 17631409]
- DiCarlo JJ, Zoccolan D, Rust NC 2012 How does the brain solve visual object recognition? *Neuron* 73:415–34 [PubMed: 22325196]
- Drugowitsch J, DeAngelis GC, Klier EM, Angelaki DE, Pouget A 2014 Optimal multisensory decision-making in a reaction-time task. *eLife* 3:e03005
- Ecker AS, Denfield GH, Bethge M, Tolias AS 2016 On the structure of neuronal population activity under fluctuations in attentional state. *J. Neurosci.* 36:1775–89 [PubMed: 26843656]
- Elsayed GF, Cunningham JP 2017 Structure in neural population recordings: an expected byproduct of simpler phenomena? *Nat. Neurosci.* 20:1310–18 [PubMed: 28783140]
- Elsayed GF, Lara AH, Kaufman MT, Churchland MM, Cunningham JP 2016 Reorganization between preparatory and movement population responses in motor cortex. *Nat. Commun.* 7:13239 [PubMed: 27807345]
- Ernst M, Banks M 2002 Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415:429–33 [PubMed: 11807554]
- Fetsch CR, Pouget A, DeAngelis GC, Angelaki DE 2011 Neural correlates of reliability-based cue weighting during multisensory integration. *Nat. Neurosci.* 15:146–54 [PubMed: 22101645]
- Fries P 2015 Rhythms for cognition: communication through coherence. *Neuron* 88:220–35 [PubMed: 26447583]
- Fries P, Reynolds JH, Rorie AE, Desimone R 2001 Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291:1560–63 [PubMed: 11222864]
- Ganguly K, Dimitrov DF, Wallis JD, Carmena JM 2011 Reversible large-scale modification of cortical networks during neuroprosthetic control. *Nat. Neurosci.* 14:662–67 [PubMed: 21499255]
- Gilbert CD, Sigman M 2007 Brain states: top-down influences in sensory processing. *Neuron* 54:677–96 [PubMed: 17553419]
- Gold JI, Shadlen MN 2007 The neural basis of decision making. *Annu. Rev. Neurosci.* 30:535–74 [PubMed: 17600525]
- Golub MD, Chase SM, Batista AP, Yu BM 2016 Brain-computer interfaces for dissecting cognitive processes underlying sensorimotor control. *Curr. Opin. Neurobiol.* 37:53–58 [PubMed: 26796293]
- Goris RL, Movshon JA, Simoncelli EP 2014 Partitioning neuronal variability. *Nat. Neurosci.* 17:858–65 [PubMed: 24777419]
- Gregoriou GG, Gotts SJ, Zhou H, Desimone R 2009 High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* 324:1207–10 [PubMed: 19478185]
- Gregoriou GG, Rossi AF, Ungerleider LG, Desimone R 2014 Lesions of prefrontal cortex reduce attentional modulation of neuronal responses and synchrony in V4. *Nat. Neurosci.* 17:1003–11 [PubMed: 24929661]
- Gu Y, Liu S, Fetsch CR, Yang Y, Fok S, et al. 2011 Perceptual learning reduces interneuronal correlations in macaque visual cortex. *Neuron* 71:750–61 [PubMed: 21867889]
- Heekeren HR, Marrett S, Ungerleider LG 2008 The neural systems that mediate human perceptual decision making. *Nat. Rev. Neurosci.* 9:467–79 [PubMed: 18464792]

- Herrero JL, Gieselmann M, Sanayei M, Thiele A 2013 Attention-induced variance and noise correlation reduction in macaque V1 is mediated by NMDA receptors. *Neuron* 78:729–39 [PubMed: 23719166]
- Huang C, Ruff DA, Cohen MR, Doiron B 2017 Modeling within and across area neuronal variability in the visual system. *Cosyne Abstracts 2017*, Salt Lake City, UT
- Huang X, Lisberger SG 2009 Noise correlations in cortical area MT and their potential impact on trial-by-trial variation in the direction and speed of smooth-pursuit eye movements. *J. Neurophysiol.* 101:3012–30 [PubMed: 19321645]
- Jacobs R 1999 Optimal integration of texture and motion cues to depth. *Vis. Res.* 39:3621–29 [PubMed: 10746132]
- Jazayeri M, Afraz A 2017 Navigating the neural space in search of the neural code. *Neuron* 93:1003–14 [PubMed: 28279349]
- Jeanne JM, Sharpee TO, Gentner TQ 2013 Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron* 78:352–63 [PubMed: 23622067]
- Kanashiro T, Ocker GK, Cohen MR, Doiron B 2017 Attentional modulation of neuronal variability in circuit models of cortex. *eLife* 6:e23978
- Kanitscheider I, Coen-Cagli R, Kohn A, Pouget A 2015a Measuring Fisher information accurately in correlated neural populations. *PLOS Comput. Biol.* 11:e1004218
- Kanitscheider I, Coen-Cagli R, Pouget A 2015b Origin of information-limiting noise correlations. *PNAS* 112:E6973–82 [PubMed: 26621747]
- Kaufman MT, Churchland MM, Ryu SI, Shenoy KV 2014 Cortical activity in the null space: permitting preparation without movement. *Nat. Neurosci.* 17:440–48 [PubMed: 24487233]
- Khaligh-Razavi SM, Kriegeskorte N 2014 Deep supervised, but not unsupervised, models may explain IT cortical representation. *PLOS Comput. Biol.* 10:e1003915
- Klink PC, Dagnino B, Gariel-Mathis MA, Roelfsema PR 2017 Distinct feedforward and feedback effects of microstimulation in visual cortex reveal neural mechanisms of texture segregation. *Neuron* 95:209–20.e3 [PubMed: 28625487]
- Knill DC 2007 Robust cue integration: a Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *J. Vis.* 7:5
- Kohn A, Coen-Cagli R, Kanitscheider I, Pouget A 2016 Correlations and neuronal population information. *Annu. Rev. Neurosci.* 39:237–56 [PubMed: 27145916]
- Kohn A, Smith MA 2005 Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J. Neurosci.* 25:3661–73 [PubMed: 15814797]
- Kording KP, Wolpert DM 2006 Bayesian decision theory in sensorimotor control. *Trends Cogn. Sci.* 10:319–26 [PubMed: 16807063]
- Kriegeskorte N 2009 Relating population-code representations between man, monkey, and computational models. *Front. Neurosci.* 3:363–73 [PubMed: 20198153]
- Lakatos P, Karmos G, Mehta A, Ulbert I, Schroeder C 2008 Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320:110–13 [PubMed: 18388295]
- Law AJ, Rivlis G, Schieber MH 2014 Rapid acquisition of novel interface control by small ensembles of arbitrarily selected primary motor cortex neurons. *J. Neurophysiol.* 112:1528–48 [PubMed: 24920030]
- Law C-T, Gold JI 2008 Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat. Neurosci.* 11:505–13 [PubMed: 18327253]
- Liu LD, Pack CC 2017 The contribution of area MT to visual motion perception depends on training. *Neuron* 95:436–46.e3 [PubMed: 28689980]
- Luo TZ, Maunsell JHR 2015 Neuronal modulations in visual cortex are associated with only one of multiple components of attention. *Neuron* 86:1182–88 [PubMed: 26050038]
- Maunsell JHR 2015 Neuronal mechanisms of visual attention. *Annu. Rev. Vis. Sci.* 1:373–91 [PubMed: 28532368]
- Maunsell JHR, Cook EP 2002 The role of attention in visual processing. *Philos. Trans. R. Soc. B* 357:1063–72

- Maunsell JHR, Treue S 2006 Feature-based attention in visual cortex. *Trends Neurosci.* 29:317–22 [PubMed: 16697058]
- Mayo JP, Maunsell JHR 2016 Graded neuronal modulations related to visual spatial attention. *J. Neurosci.* 36:5353–61 [PubMed: 27170131]
- McAdams CJ, Maunsell JHR 1999 Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J. Neurosci.* 19:431–41 [PubMed: 9870971]
- Miller EK, Buschman TJ 2013 Cortical circuits for the control of attention. *Curr. Opin. Neurobiol.* 23:216–22 [PubMed: 23265963]
- Miri A, Warriner CL, Seely JS, Elsayed GF, Cunningham JP, et al. 2017 Behaviorally selective engagement of short-latency effector pathways by motor cortex. *Neuron* 95:683–96 [PubMed: 28735748]
- Mitchell JF, Sundberg KA, Reynolds JH 2007 Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron* 55:131–41 [PubMed: 17610822]
- Mitchell JF, Sundberg KA, Reynolds JH 2009 Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron* 63:879–88 [PubMed: 19778515]
- Moore T, Armstrong KM 2003 Selective gating of visual signals by microstimulation of frontal cortex. *Nature* 421:370–73 [PubMed: 12540901]
- Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A 2014 Information-limiting correlations. *Nat. Neurosci.* 17:1410–17 [PubMed: 25195105]
- Nandy AS, Nassi JJ, Reynolds JH 2016 Laminar organization of attentional modulation in macaque visual area V4. *Neuron* 93:235–46 [PubMed: 27989456]
- Ni AM, Ruff DA, Alberts JJ, Symmonds J, Cohen MR 2017 Learning and attention reveal a general relationship between neuronal variability and perception. *bioRxiv* 137083. 10.1101/137083
- Nienborg H, Cohen MR, Cumming BG 2012 Decision-related activity in sensory neurons: correlations among neurons and with behavior. *Annu. Rev. Neurosci.* 35:463–83 [PubMed: 22483043]
- Nienborg H, Cumming BG 2009 Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature* 459:89–92 [PubMed: 19270683]
- Oemisch M, Westendorff S, Everling S, Womelsdorf T 2015 Interareal spike-train correlations of anterior cingulate and dorsal prefrontal cortex during attention shifts. *J. Neurosci.* 35:13076–89 [PubMed: 26400938]
- Pagan M, Urban LS, Wohl MP, Rust NC 2013 Signals in inferotemporal and perirhinal cortex suggest an untangling of visual target information. *Nat. Neurosci.* 16:1132–39 [PubMed: 23792943]
- Parker AJ, Newsome WT 1998 Sense and the single neuron: probing the physiology of perception. *Annu. Rev. Neurosci.* 21:227–77 [PubMed: 9530497]
- Pitkow X, Angelaki DE 2017 Inference in the brain: statistics flowing in redundant population codes. *Neuron* 94:943–53 [PubMed: 28595050]
- Pitkow X, Liu S, Angelaki DE, DeAngelis GC, Pouget A 2015 How can single sensory neurons predict behavior? *Neuron* 87:411–23 [PubMed: 26182422]
- Pooresmaeili A, Poort J, Roelfsema PR 2014 Simultaneous selection by object-based attention in visual and frontal cortex. *PNAS* 111:6467–72 [PubMed: 24711379]
- Poort J, Self MW, van Vugt B, Malkki H, Roelfsema PR 2016 Texture segregation causes early figure enhancement and later ground suppression in areas V1 and V4 of visual cortex. *Cereb. Cortex* 26:3964–76 [PubMed: 27522074]
- Quian Quiroga R, Panzeri S 2009 Extracting information from neuronal populations: information theory and decoding approaches. *Nat. Rev. Neurosci.* 10:173–85 [PubMed: 19229240]
- Rabinowitz NC, Goris RL, Cohen MR, Simoncelli EP 2015 Attention stabilizes the shared gain of V4 populations. *Elife* 4:e08998
- Raiguel S, Vogels R, Mysore SG, Orban GA 2006 Learning to see the difference specifically alters the most informative V4 neurons. *J. Neurosci.* 26:6589–602 [PubMed: 16775147]
- Reynolds JH, Chelazzi L 2004 Attentional modulation of visual processing. *Annu. Rev. Neurosci.* 27:611–47 [PubMed: 15217345]
- Ruff DA, Alberts JJ, Cohen MR 2016 Relating normalization to neuronal populations across cortical areas. *J. Neurophysiol.* 116:1375–86 [PubMed: 27358313]

- Ruff DA, Cohen MR 2014a Attention can either increase or decrease spike count correlations in visual cortex. *Nat. Neurosci.* 17:1591–97 [PubMed: 25306550]
- Ruff DA, Cohen MR 2014b Global cognitive factors modulate correlated response variability between V4 neurons. *J. Neurosci.* 34:16408–16 [PubMed: 25471578]
- Ruff DA, Cohen MR 2016a Attention increases spike count correlations between visual cortical areas. *J. Neurosci.* 36:7523–34 [PubMed: 27413161]
- Ruff DA, Cohen MR 2016b Stimulus dependence of correlated variability across cortical areas. *J. Neurosci.* 36:7546–56 [PubMed: 27413163]
- Ruff DA, Cohen MR 2017 A normalization model suggests that attention changes the weighting of inputs between visual areas. *PNAS* 114:E4085–94 [PubMed: 28461501]
- Saalmann Y, Pigarev I, Vidyasagar T 2007 Neural mechanisms of visual attention: how top-down feedback highlights relevant locations. *Science* 316:1612–15 [PubMed: 17569863]
- Sadtler PT, Quick KM, Golub MD, Chase SM, Ryu SI, et al. 2014 Neural constraints on learning. *Nature* 512:423–26 [PubMed: 25164754]
- Saproo S, Serences JT 2014 Attention improves transfer of motion information between V1 and MT. *J. Neurosci.* 34:3586–96 [PubMed: 24599458]
- Semedo J, Zandvakili A, Machens C, Yu BM, Kohn A 2016 Predicting V2 activity from V1 population activity. *Cosyne Abstracts 2016*, Salt Lake City, UT
- Shadlen MN, Britten KH, Newsome WT, Movshon JA 1996 A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J. Neurosci.* 16:1486–510 [PubMed: 8778300]
- Smith MA, Kohn A 2008 Spatial and temporal scales of neuronal correlation in primary visual cortex. *J. Neurosci.* 28:12591–603 [PubMed: 19036953]
- Smith MA, Sommer MA 2013 Spatial and temporal scales of neuronal correlation in visual area V4. *J. Neurosci.* 33:5422–32 [PubMed: 23516307]
- Smolyanskaya A, Ruff DA, Born R 2013 Joint tuning for direction of motion and binocular disparity in macaque MT is largely separable. *J. Neurophysiol.* 110:2806–16 [PubMed: 24089395]
- Solomon SS, Chen SC, Morley JW, Solomon SG 2015 Local and global correlations between neurons in the middle temporal area of primate visual cortex. *Cereb. Cortex* 25:3182–96 [PubMed: 24904074]
- Sprague TC, Saproo S, Serences JT 2015 Visual attention mitigates information loss in small- and large-scale neural codes. *Trends Cogn. Sci.* 19:215–26 [PubMed: 25769502]
- Verhoef BE, Maunsell JHR 2017 Attention-related changes in correlated neuronal activity arise from normalization mechanisms. *Nat. Neurosci.* 20:969–77 [PubMed: 28553943]
- Wimmer K, Compte A, Roxin A, Peixoto D, Renart A, de la Rocha J 2015 Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area MT. *Nat. Commun.* 6:6177 [PubMed: 25649611]
- Womelsdorf T, Fries P 2007 The role of neuronal synchronization in selective attention. *Curr. Opin. Neurobiol.* 17:154–60 [PubMed: 17306527]
- Womelsdorf T, Fries P, Mitra PP, Desimone R 2006 Gamma-band synchronization in visual cortex predicts speed of change detection. *Nature* 439:733–36 [PubMed: 16372022]
- Yamins DL, DiCarlo JJ 2016 Using goal-driven deep learning models to understand sensory cortex. *Nat. Neurosci.* 19:356–65 [PubMed: 26906502]
- Yamins DL, Hong H, Cadieu CF, Solomon EA, Seibert D, DiCarlo JJ 2014 Performance-optimized hierarchical models predict neural responses in higher visual cortex. *PNAS* 111:8619–24 [PubMed: 24812127]
- Yan Y, Rasch MJ, Chen M, Xiang X, Huang M, et al. 2014 Perceptual training continuously refines neuronal population codes in primary visual cortex. *Nat. Neurosci.* 17:1380–87 [PubMed: 25195103]
- Yantis S, Serences JT 2003 Cortical mechanisms of space-based and object-based attentional control. *Curr. Opin. Neurobiol.* 13:187–93 [PubMed: 12744972]
- Yuste R 2015 From the neuron doctrine to neural networks. *Nat. Rev. Neurosci.* 16:487–97 [PubMed: 26152865]

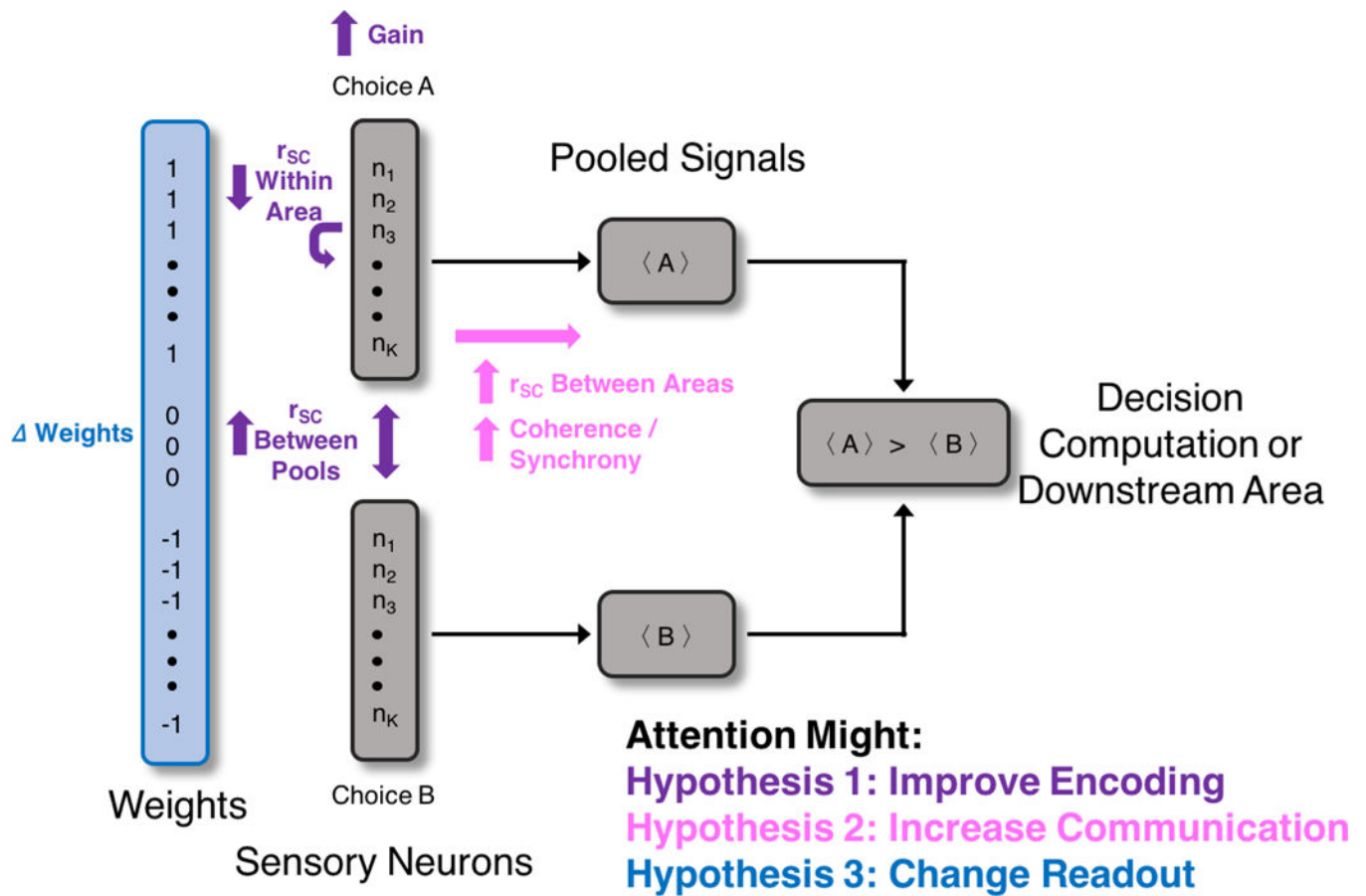
- Zénon A, Krauzlis R 2012 Attention deficits without cortical neuronal deficits. *Nature* 489:434–37 [PubMed: 22972195]
- Zohary E, Shadlen M, Newsome W 1994 Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370:140–43 [PubMed: 8022482]

Author Manuscript

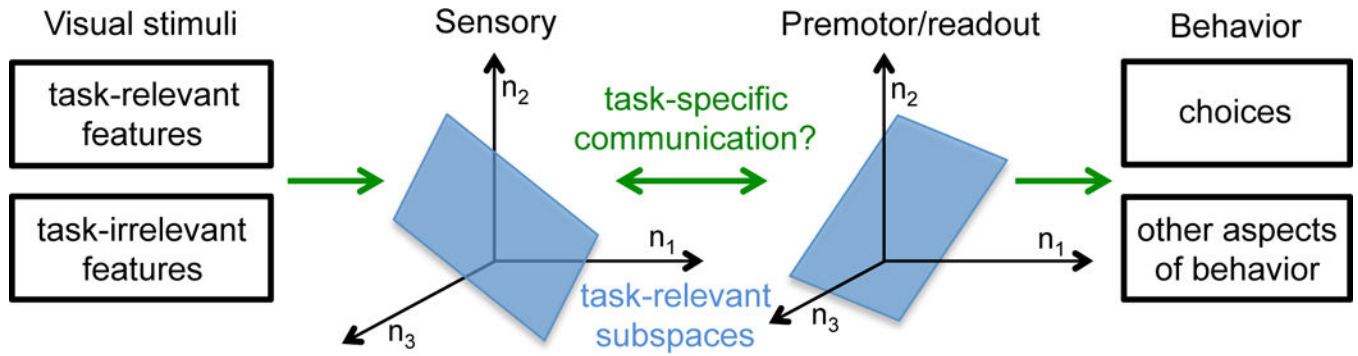
Author Manuscript

Author Manuscript

Author Manuscript

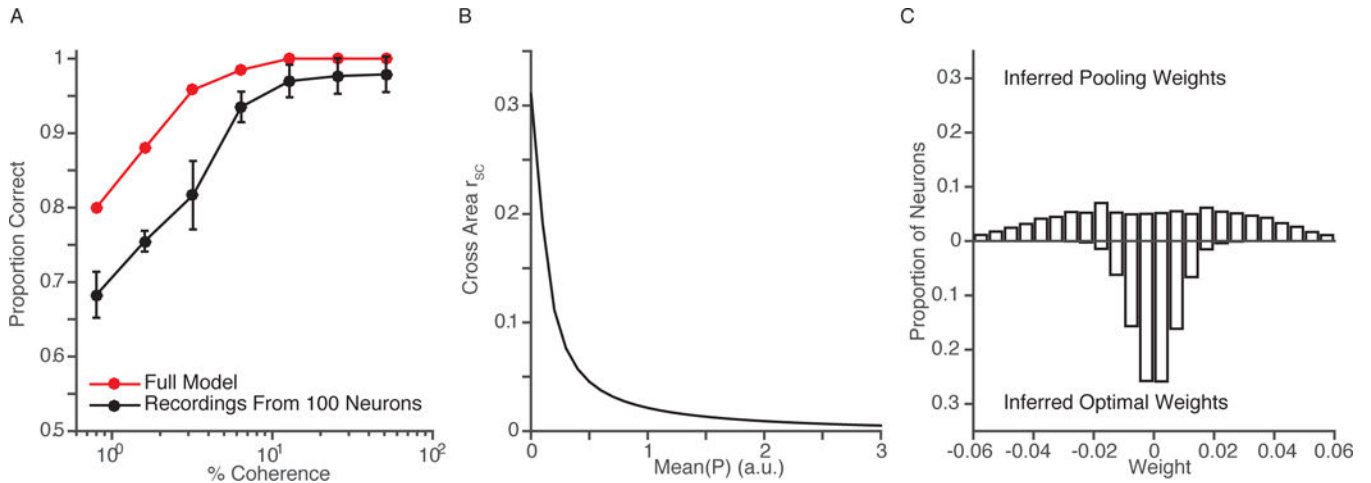


**Figure 1.** Schematic of the pooling model (adapted from Shadlen et al. 1996), highlighting cognition-related changes to neuronal responses that would provide evidence in favor of each of the three hypothesized underlying mechanisms (color coded for each mechanism). Abbreviation:  $r_{SC}$ , spike count or noise correlation.



**Figure 2.**

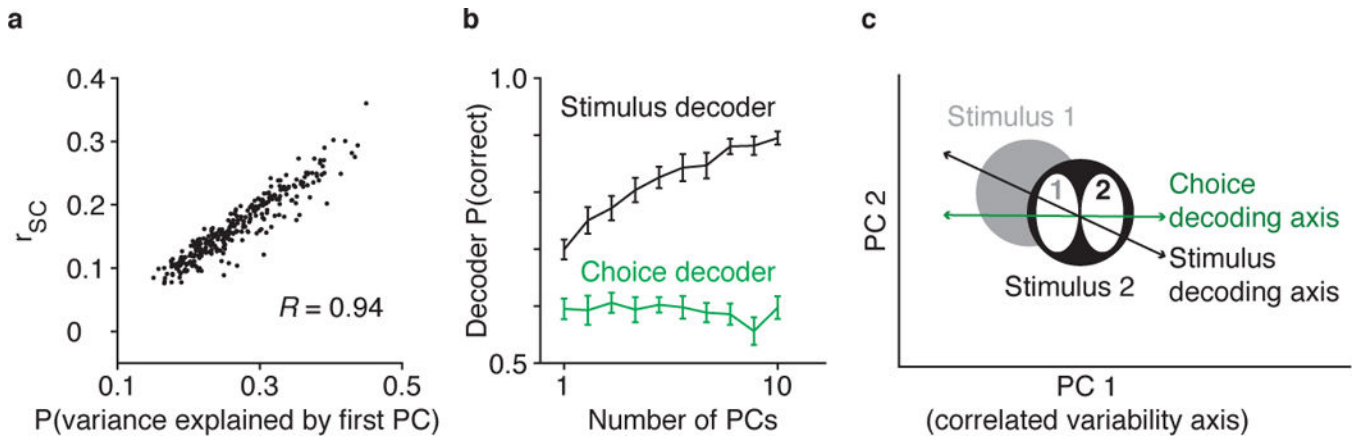
Schematic of a population subspace framework. Visual stimuli contain many features, a subset of which are relevant for a given perceptual task. For example, populations of MT neurons encode both motion direction and spatial frequency, even when the task concerns only motion direction. The activity of  $k$  MT neurons (neurons  $n_1, n_2, n_3, \dots, n_k$ ) can be plotted in a  $k$ -dimensional space, but the subspace corresponding to motion direction (*blue*) is likely lower dimensional. Similarly, neuronal populations in downstream areas encode a variety of stimulus, cognitive, and premotor factors, a subset of which are relevant for perceptual choices. We hypothesize that communication between the areas (*double-ended arrow*) will also be low dimensional and task specific. The subspaces of neural activity that are relevant for each task can be thought of as linear (or nonlinear) combinations of the responses of all of the neurons in the population.



**Figure 3.**

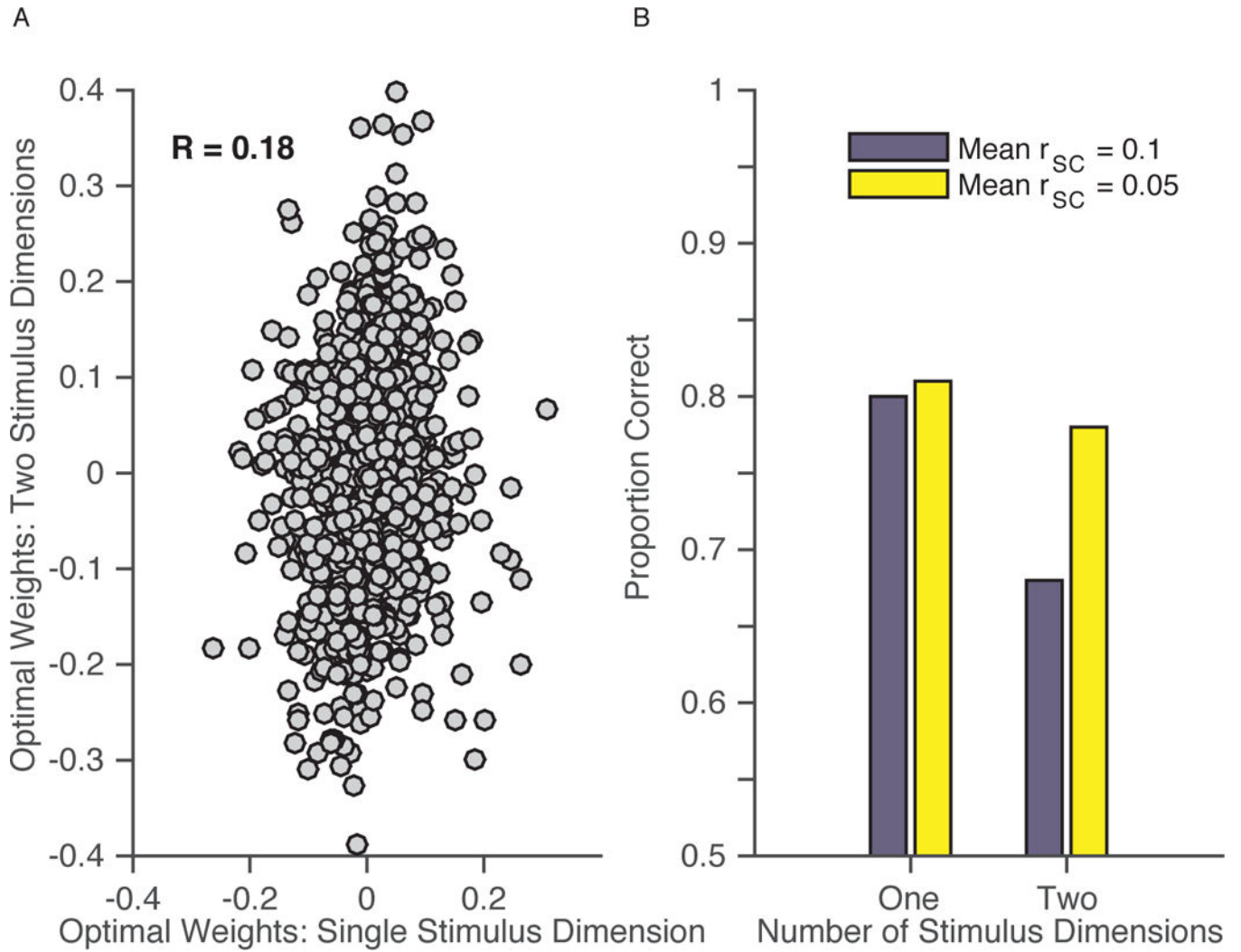
Insights from a set of simple simulations. (a) Proportion of correct discriminations of the full population (*red*) and of subsets of 100 selected neurons (*black*) using weights inferred from the simulation's choices (which are based on the full population). Error bars represent 95% confidence intervals based on 1,000 random draws (without replacement) of 100 neurons. The full population outperforms the subsets, but performance using small populations covaries with that of the full population. If this property is true in real neuronal populations, modest improvements in the sensory information that the animals use should be observable from recordings of realistic numbers of recorded neurons. (b) When neuronal responses are based on linear combinations of inputs that are either shared or private to each neuron, correlations between simulated populations are inversely related to the strength of private inputs. In this scenario, changes in the proportion of private and shared inputs could be observed by measuring  $r_{sc}$  between areas. (c) Differences in weighting functions can be detected using recordings from small populations. The distribution of inferred weights of a conventional pooling model (*top*) is bimodal and broad. The distribution of inferred weights from a model that uses an optimal readout scheme (*bottom*) is narrow and unimodal.





**Figure 4.**

Correlated variability is related to performance (Ni et al. 2017). (a) The variance explained by the first PC, based on PCA of the recorded population's responses to repeated presentations of the same stimulus, was highly correlated with the mean  $r_{SC}$  of all neuron pairs in the recorded population across experiments. Thus, the first PC provided a single trial measure of  $r_{SC}$ . (b) The first PC (and thus correlated variability) explained essentially all the choice-related population activity, as the choice decoder performed just as well based on the first PC as it did with additional PCs. In contrast, the stimulus decoder's performance improved with additional PCs. (c) Population responses illustrated for the first and second PCs only. The first PC ( $x$ -axis) is by definition the axis that explains the most correlated variability. The stimulus decoding axis (*black line*) detects differences between neuronal responses to stimulus 51 and stimulus 2, while the choice decoding axis (*green line*) detects differences between when the subject made the correct versus the incorrect choice [e.g., target 2 would be the correct choice, and target 1 the incorrect choice, when stimulus 2 was presented (*ovals* 2 and 1)]. Abbreviations: PC, principal component; PCA, principal component analysis;  $r_{SC}$ , spike count or noise correlation.



**Figure 5.**

An optimal general decoder uses very different weights than an optimal decoder tailored for each stimulus and accounts for known experimental results. (a) In simulation, the weights of a decoder that contends with two stimulus features (one task-relevant and one task-irrelevant feature; *y*-axis) are only weakly correlated with the weights of a decoder that contends with stimuli that vary only in a single, task-relevant feature. (b) Changing the mean  $r_{SC}$  has a much larger effect on the general decoder that contends with variability in two stimulus features (*right*) than the decoder for stimuli that vary only in a single stimulus dimension (*left*).