UCLA
UCLA Previously Published Works

Title
Insights into decision making using choice probability

Permalink
https://escholarship.org/uc/item/6j16v06m

Journal
Journal of Neurophysiology, 114(6)

ISSN
0022-3077

Authors
Crapse, Trinity B
Basso, Michele A

Publication Date
2015-12-01

DOI
10.1152/jn.00335.2015

Peer reviewed

Review

CALL FOR PAPERS | *Decision Making: Neural Mechanisms*

Insights into decision making using choice probability

Trinity B. Crapse and Michele A. Basso
Joaquin Fuster Laboratory of Cognitive Neuroscience, Departments of Psychiatry and Biobehavioral Sciences and Neurobiology, The Semel Institute for Neuroscience and Human Behavior and the Brain Research Institute, University of California, Los Angeles, Los Angeles, California

Crapse TB, Basso MA. Insights into decision making using choice probability. *J Neurophysiol* 114: 3039–3049, 2015. First published September 16, 2015; doi:10.1152/jn.00335.2015.—A long-standing question in systems neuroscience is how the activity of single neurons gives rise to our perceptions and actions. Critical insights into this question occurred in the last part of the 20th century when scientists began linking modulations of neuronal activity directly to perceptual behavior. A significant conceptual advance was the application of signal detection theory to both neuronal activity and behavior, providing a quantitative assessment of the relationship between brain and behavior. One metric that emerged from these efforts was choice probability (CP), which provides information about how well an ideal observer can predict the choice an animal makes from a neuron's discharge rate distribution. In this review, we describe where CP has been studied, locational trends in the values found, and why CP values are typically so low. We discuss its dependence on correlated activity among neurons of a population, assess whether it arises from feedforward or feedback mechanisms, and investigate what CP tells us about how many neurons are required for a decision and how they are pooled to do so.

signal detection theory; sensation; perception; neurophysiology; vision; eye movements; correlated variability; multiple neuron recording

HOW ACTION POTENTIALS GENERATED from single neurons and groups of neurons give rise to our perceptions, emotions, decisions and actions is arguably the most vexing question in neuroscience. From the pioneering work of Adrian (1928), Hartline and McDonald (1947) and Barlow (1972) to Kiang et al. (1965) and Siebert (1970) in the auditory system, Mountcastle and Vallbo in the somatosensory system (Johansson and Vallbo 1979; Mountcastle et al. 1967), Hubel and Wiesel (1962) in the visual system, and Evarts (1968) in the motor system, the search for the neuronal correlates of our perceptions and actions remains a dominant focus of experiments in neuroscience. Without question, a significant breakthrough in efforts to understand the relationship between neuronal activity and perception came with the application of signal detection theory in psychophysics (Green and Swets 1966) to neurophysiology. Whereas most of the initial experiments focused on correlating individual neuronal activity to varying levels of sensory stimulation (examples cited above and Bradley et al. 1987; Britten et al. 1992; Newsome et al. 1989; Parker and Hawken 1985; Skottun et al. 1987; Tolhurst et al. 1983; Vogels

and Orban 1991), a major advance came when investigators began directly linking modulations in neuronal activity to an animal's reported percept (e.g., Logothetis and Schall 1989). Later work built on this insight and used signal detection theory approaches to quantify precisely the relationship between neuronal responses and behavioral responses of monkeys performing sensory discrimination tasks (Britten et al. 1992; Newsome et al. 1989). This new approach capitalized on the rigorous quantitative measures provided by signal detection theory, as had previous work, but now was being applied to internal, perceptual experiences. Specifically, this advance led to the ability to correlate variations in neuronal activity occurring with sensory signals to variations in behavioral choices based on those same sensory signals (Britten et al. 1996). Because the metric quantifies how well an ideal observer can predict a choice, the investigators coined the term choice probability (CP). CP as a measure provides a way to link trial-to-trial fluctuations in neuronal activity to fluctuations in internal percepts as indicated by behavior.

CP may be used to shed light on how the brain makes decisions. Here we focus on three issues: *1*) what brain areas contain neurons with activity correlated with perceptual decisions?; and *2*) how many neurons are required and how is their activity pooled to inform or reach a decision? Using CP to ask

Address for reprint requests and other correspondence: M. A. Basso, Fuster Laboratory of Cognitive Neuroscience, Dept. of Psychiatry and Biobehavioral Sciences and Neurobiology, UCLA, 635 Charles E. Young Drive Box 957332, Los Angeles, CA 90095 (e-mail: mbasso@mednet.ucla.edu).

the second question includes a tacit assumption that CP originates from feedforward mechanisms. Recent work on CP calls this assumption into question. Therefore, we also ask, *3*) what is the origin of CP?; does it result from feedforward pooling of neuronal activity or from feedback mechanisms such as attentional allocation? We should point out that CP can be applied to brain areas thought to report only the sensory evidence or to areas that directly mediate decisions. In either case, CP provides useful information about how well a brain area's activity is predictive of an animal's perceptual report. Whereas single neuron recording techniques applied to different brain regions while animals perform behavioral tasks is sufficient to address the first question, advances toward answering the second question are aided by the implementation of multiple neuron recording techniques. With the introduction of these techniques, investigators are able to eavesdrop on large populations of neurons and measure correlations in variability of activity between many neurons participating in decision making. Detailed information about the correlation structure in neuronal populations provides powerful constraints on our interpretation and understanding of CP and, as a result, on models describing how neuronal activity gives rise to our decisions.

After a brief introduction describing what CP is and how it is calculated, we discuss recent work that addresses the questions posed above. We review evidence revealing that many brain areas show significant CPs but that the values are typically low. Two important research directions developed from this curious observation: first, are there task or stimulus conditions in which CPs can be improved? Second, what do CPs tell us about the pooling mechanisms used by neurons to arrive at decisions? We review the evidence for two types of pooling models of decision making stemming from observations made using CPs and illustrate how measurements of correlated activity among neurons reveal new ways of thinking about these models. Pooling models assume feedforward mechanisms, signals converging from lower areas are integrated in higher areas, and CPs are tacitly assumed to reflect the process of feedforward pooling of neuronal activity. However, recent work calls into question this assumption and suggests that other mechanisms, notably feedback, may be at play in determining CP values and, therefore, decision making.

The Calculation of CP and Signal Detection Theory

The CP metric is based on signal detection theoretic methods and requires the calculation of a receiver operating characteristic (ROC) curve. The ROC is a nonparametric statistical measure that assesses how likely an event is, given random draws from two response distributions. In a typical experiment, an animal judges the presence, or quality, of a sensory stimulus, and reports its perceptual decision, as either "yes, I see the stimulus" or "no, I do not see the stimulus" or, as another example, the stimulus direction is "left" or "right." In either case, animals report their decision often by making an eye movement to a target that indicates which of two possible percepts it experienced. The difficulty of the task varies from trial-to-trial by experimental manipulations of the strength of the sensory stimulus, and neuronal activity is monitored as the animal arrives at and reports its decision. The resulting data are distributions of discharge rates associated with each of two choices in response to varying strengths of sensory stimuli.

One of the distributions is referred to as the "noise" distribution, and the other as the "signal" distribution. In a detection task, the "noise" distribution represents the discharge rates measured on trials in which no stimulus appeared, whereas the "signal" distribution represents discharge rates on trials in which a stimulus appeared. In a discrimination task, such as the random dot motion task, discharge rates occurring during one direction of motion, usually in the neuron's nonpreferred direction, make up the "noise" distribution, whereas discharge rates occurring in trials in which the motion direction is in the neuron's preferred direction make up the "signal" distribution (Fig. 1, *A* and *B*). The ROC curves quantify the differences between these distributions by determining how much of each distribution exceeds a given threshold or criterion level. A "hit" occurs when the neuron's activity on a "signal" trial exceeds the criterion and a "false alarm" occurs when the neuron activity on a "noise" trial exceeds the criterion. To calculate the entire ROC curve, hits and false alarms are calculated at each criterion level, a process that is iterated from a minimum (e.g., 0 sp/s) to a maximum discharge rate (Fig. 1, *B* and *C*). Plotting all the hit rates vs. all the false alarm rates yields the ROC curve (Fig. 1*C*, *right*) and integrating the area under the ROC curve (auROC) yields a single-value (ranging from 0–1) that quantifies the degree of separation between the two distributions. Values of 0.5 indicate that the two distributions overlap completely, and an ideal observer would be unable to determine from which distribution a discharge rate originated. Values >0.5 indicate separation between the two distributions. A value of 1 indicates no overlap and an ideal observer could determine from which distribution the draw originated precisely.

ROC may be used to calculate the sensitivity of a neuron to stimuli varying parametrically or to quantify the relationship between the activity of a neuron and behavioral choice, the CP. The former, referred to as the neurometric function, was used in a now classic experiment involving neuronal recordings in area MT, a motion-selective area of the monkey brain, during performance of a motion direction discrimination task (Britten et al. 1992). Monkeys judged the direction of clouds of moving dots and the difficulty of the discrimination varied from trial-to-trial by experimentally manipulating the coherence of the moving dots; more coherent motion resulted in easier discrimination. Two distributions of discharge rates were acquired for each coherence level; one for leftward and one for rightward motion. ROC areas were plotted as a function of stimulus coherence level (or motion strength), and these data were fitted with a logistic function yielding the neurometric function. This function describes an individual neuron's sensitivity or ability to discriminate between the two stimulus directions across different motion strengths.

In contrast to the neurometric function, which characterizes a neuron's tuning or sensitivity, CP relates neuronal discharges directly to behavioral choices and thus provides information about how well a neuron tracks an internal percept. The calculation of CP begins with ROC analysis and the determination of auROC, but for CP, the two distributions are associated with the two possible choices for each stimulus level. For example, in tasks involving judgments of leftward or rightward directions of moving dots, one set of discharge distributions comes from trials in which the animal chose rightward, and the other comes from trials in which the animal chose leftward.
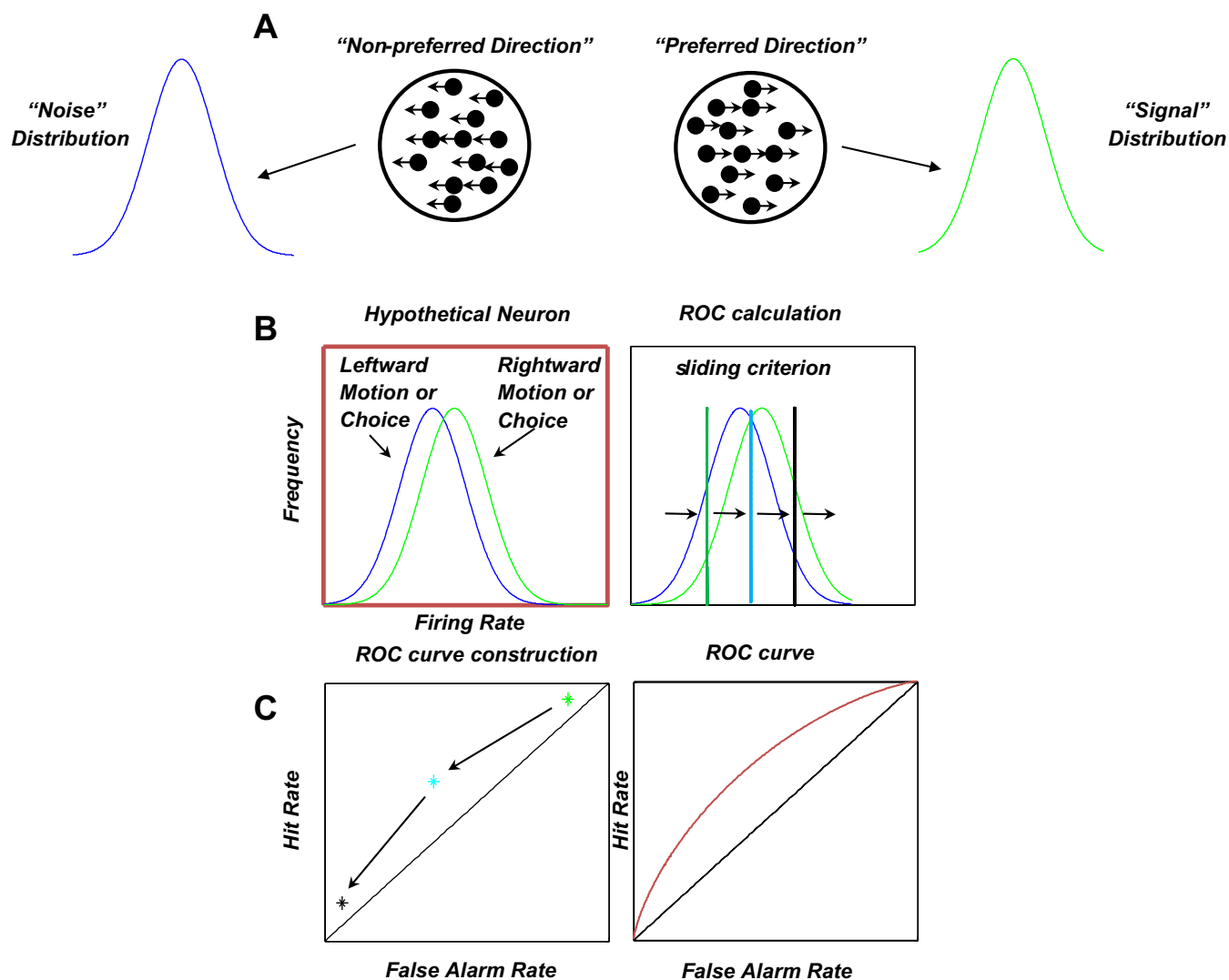
Fig. 1. Receiver operating characteristics (ROC) and its calculation. *A*: in a random dot motion discrimination task, 2 motion directions are possible (leftward or rightward) and the animal's task is to indicate which direction occurred. A hypothetical MT neuron monitored during this task would yield a distribution of discharge rates for each direction, 1 in the preferred direction (rightward) called the signal distribution and 1 in the nonpreferred direction (leftward), called the noise distribution. *B*, *left*: 2 distributions of firing rates for a hypothetical neuron monitored during the random dot motion task. The 2 distributions may represent leftward and rightward motion directions, respectively, relevant for neurometric calculation, or leftward and rightward choices, respectively, relevant for choice probability (CP) calculation. In either case, ROC is used to quantify the difference between the 2 distributions. *B*, *right*: ROC is calculated by systematically sliding a criterion across the 2 distributions and integrating the area of each distribution that exceeds each criterion level. Shown are 3 example criterion levels for the hypothetical neuron and the resulting ROC values (*C*, *left*) obtained at each level. The ROC curve is constructed by plotting all ROC points thus obtained (*C*, *right*).

Due to differences in firing rate associated with each stimulus strength, CP is typically calculated for each stimulus value. For discrimination tasks such as the random dot motion task, this means that CP would be calculated for each coherence level. Investigators may group all coherence levels and report a single CP, but the discharge distributions must be z-transformed first to remove differences in discharge rate related to the differing stimulus levels. CP values near 0.5 indicate chance performance and reveal that a neuron's activity has no detectable relation to behavior. Values closer to 1 indicate that the neuron's activity is correlated with an animal's perceptual experience as revealed by its behavioral choices. In any case, CP provides a simple and intuitive way of gauging the relationship between a neuron's activity and an animal's perceptual behavior.

CPs Appear in Many Brain Areas

CPs have been measured in many cortical areas during performance of a variety of perceptual decision-making tasks, for example MT, MST, lateral intraparietal area (LIP), somatosensory cortex areas S1 and S2 and premotor and motor cortical areas, as well as subcortical areas involved in eye movement selection (Britten et al. 1996; Celebrini and Newsome 1994; Cohen and Newsome 2009; de Lafuente and Romo 2005, 2006; Dodd et al. 2001; Gu et al. 2008; Kim and Basso 2008; Law and Gold 2008; Liu and Newsome 2005; Nienborg and Cumming 2006, 2009, 2014; Price and Born 2010; Romo et al. 2002; Sasaki and Uka 2009; Uka and DeAngelis 2004). What is evident from this literature is that CPs typically increase in value as one approaches the site of decision making, and CPs in early

sensory areas are often particularly low. For example, Romo and colleagues recorded from somatosensory and motor areas involved in the discrimination of vibrotactile stimuli (Hernandez et al. 2010). Monkeys compared two vibratory tactile stimuli sequentially applied to their fingertip and reported their decision by pressing a button indicating whether the second stimulus frequency was lower or higher than the first stimulus. Significant CPs were absent in S1 and emerged gradually starting in S2, with maximal CPs obtained in prefrontal cortex and movement-related premotor cortex.

A similar progression of increasing CP values appears in the visual system while monkeys report the direction of an apparent motion stimulus (Williams et al. 2003). In this experiment, the stimuli were designed carefully so that one of two percepts could be obtained from the same stimulus and neuronal recordings were made from cortical areas MT, MST, and LIP. Neurons in MT responded to the sensory stimulus but ultimately had little information about the reported percept and thus had no correlation with perceptual choice, as measured by CP. Areas MST and LIP, however, did exhibit significant CPs, with area LIP, a putative site of decision making, exhibiting both a higher fraction of neurons with significant CPs and a larger mean CP across neurons. Similar CP values are observed in the superior colliculus, a structure considered a final processing stage for eye movement generation, during a target selection task. There the values could be as high as 0.7 (Kim and Basso 2008). Consistent with the underlying idea that CPs predict an internal state related to behavioral choice, these results indicate that CPs increase further from the sensory periphery and closer to the putative sites of decision making and movement generation.

In sum, CPs appear in a variety of areas and they tend to increase in value further in the processing hierarchy as though decisions are made in a feedforward fashion. These signals could emerge from feedforward processing, or alternatively, they could reflect the contribution of feedback signals originating from structures further downstream. In the next section we explore the implications of CP values for possible pooling mechanisms assuming a feedforward architecture. In a later section we review the evidence that calls this assumption into question.

Does Size Matter?

In the apparent motion study described above, the mean CPs in MST and LIP were 0.53 and 0.58, respectively (Williams et al. 2003). These values are within the typical range of values that are observed in dorsal visual stream areas during motion discrimination tasks (Britten et al. 1996; Celebrini and Newsome 1994; Cohen and Newsome 2009; Law and Gold 2008; Price and Born 2010; Purushothaman and Bradley 2005; Sasaki and Uka 2009). An obvious question is why are these CP values so close to chance, i.e., 0.5? CP values at chance or near chance levels would be expected if the decision was based on the aggregate activity of large numbers of neurons conveying independent, nonredundant signals. In this scenario, each neuron would bare very little relation to the group signal determining the decision, and so one might well ask, "why are observed CP values so big?" Much of the information carried by neurons is correlated and redundant (Pitkow et al. 2015),

however, with many neurons exhibiting a strong relationship with the group signal. Therefore, here we focus on the fact that the values stay so close to chance despite this and so we ask "why so low?" Measuring a low CP from a sensory area may not be surprising when the sensory information on which the decision is based is weak. However, CP values tend to be relatively constant across all stimulus strengths, so why are CP values typically low for even those discriminations involving strong sensory information? Moreover, why are CPs so low in areas such as LIP? These areas are thought to be involved in converting sensory evidence into a decision, so even if an animal were to guess about a stimulus with no sensory evidence, these areas should show a high CP since the activity of these neurons is considered to be causal for the decision. One possibility explaining why CPs values are so low is that the epoch used for measuring the discharge rate matters. Decisions take time to evolve so the epoch used for calculating CP may be critical to ensure measurement of the most accurate CP value. To overcome this limitation, many investigators calculate running CPs by sliding a window beginning at the onset of a sensory stimulus and terminating at the time of the choice report. To return to our apparent motion example (Williams et al. 2003), CPs in MST and LIP calculated over time peak at ~0.58 and 0.68, respectively, values that are considerably higher than the reported means. Such studies reveal interesting temporal dynamics of CPs that must be considered when choosing an appropriate epoch of interest. As we will discuss below, changes in CP over time reveal important information about when an area is most likely involved in a decision. Furthermore, the timing of maximal CP values places constraints on models of whether a brain area is causally involved in decision making, since peaks occurring after the animal makes its decision would have no impact on the decision process itself.

A second possible reason for low CPs may be related to the type of stimulus used. Experiments measuring CPs in sensory areas MT, V1, and V2 hint at this possibility. Dodd et al. (2001) trained monkeys to make directional judgments about a rotational cylinder composed of randomly moving dots. At zero disparity the perceived direction of rotation undergoes spontaneous changes; sometimes it is perceived as rotating leftward and sometimes rightward. In other words, the stimulus is perceptually ambiguous but has identical retinal stimulation. The average CP for MT neurons measured during this task was ~ 0.67, appreciably larger than that reported in MT during motion direction discrimination tasks. A possibility for the discrepancy is that the motion stimulus used in this experiment had strong depth cues that may drive MT neurons better than motion direction cues alone. In another study, trained monkeys detected motion pulses induced by phase shifts of Gabor patches (Ghose and Harrison 2009). Mean CP values of 0.67, identical to that reported by Dodd et al. (2001), were observed. Similar to Dodd et al., the stimulus differed from the random dot motion stimulus. The Gabor stimuli used contained little noise, indicating that a low CP, as observed in other studies, could be related to noise in the stimulus itself rather than noise in the nervous system. Together, these studies bring to the fore the important point that the type of stimulus used and task requirements can influence the magnitude of CP values.

Work described in a pair of recent papers by Nienborg and Cumming (2006, 2014) further emphasizes the importance of

stimulus specificity for determining a maximal CP. In the first study, they recorded from V1 and V2 during a depth-disparity judgment task and found surprisingly, that V2 had significant CPs while V1 had no significant activity correlated with the animal's disparity judgment (Nienborg and Cumming 2006). This was surprising because both V1 and V2 neurons exhibit disparity tuning. In a followup paper, they discovered why this was the case (Nienborg and Cumming 2014). They returned to V1, but this time instead of a disparity task, they used an orientation judgment task. They reasoned that the presence or absence of significant CPs may depend on the presence of maps coding for the discriminandum used. This could explain why V2, with its strong disparity map, had significant CPs during a disparity judgment task. V1 has a strong orientation map, so they predicted that V1 would show significant CPs in an orientation task. Indeed, V1 neurons showed significant CPs while animals made judgments about orientation. These results, taken together with the higher CPs obtained in MT using a structure from motion task, or an "easy" phase change detection task, suggest that the type of stimulus used and the coding properties of the brain area are critical determinants of the size of the observed CPs.

A third reason that measured CPs are low could be related to pooling operations occurring downstream of the sensory processing stage. Noise introduced at such later stages, arising from stochastic cellular events related to random vesicle release and ion channel dynamics (Faisal et al. 2008), could effectively de-yoke the sensory pool signals from the decision stage, leading to lower correlations of sensory pool neuronal activity with behavior. This would translate into lower CPs.

It should be pointed out that the low CP applies typically to the mean CP. Brain areas possess neurons that exhibit a range of CP values, with some individual neurons having CP values as high as 0.8 and some as low as 0.5 (e.g., Britten et al. 1996). Therefore, a mean CP value is only informative if the signals from all neurons are weighted equally. That distributions of CP values are observed may indicate that signals from neurons are not weighted equally. Rather, some receive higher weights, some receive lower weights toward the decision. Neurons with signals more heavily weighted would exert a greater influence on the decision process and thus would show tighter correlations with behavior as measured by CP. This leads to the question of which neurons receive the highest weight. Many neurons with high CP tend to be very sensitive to sensory stimulation as evidenced by their neurometric functions (Britten et al. 1996; Gu et al. 2008; Parker et al. 2002; Purushothaman and Bradley 2005) suggesting that downstream areas may ignore signals from neurons with low sensitivity and only listen to those with high sensitivity. Indeed, a modeling exercise by Purushothaman and Bradley found that psychophysical thresholds measured in monkeys performing a fine perceptual discrimination task could be replicated only if the decision stage consulted the most informative neurons, those with greatest precision (Purushothaman and Bradley 2005). Although they did not calculate model neuron CPs, their result suggests that only neurons with high sensitivity are averaged together selectively to inform a decision. Such selective pooling would result in some neurons exhibiting relatively high CPs. These considerations suggest that an emphasis on the mean CP may be misleading since it tacitly assumes that all neurons are contributing equally to a decision.

How Many Neurons Does It Take To Make a Decision?

Another assumption in the interpretation of CP is that it reflects a feedforward process in which signals within an area are integrated and then passed on to subsequent stages where further integration occurs. The integration occurs among neurons with similar tuning preferences, a process referred to as pooling. Trial-to-trial fluctuations in the response of neurons are thought to give rise to trial-to-trial behavioral variability without the need for feedback signals. The idea that neurons are pooled (summed or averaged) in a feedforward fashion to mediate a decision was proposed in a now classic computational study (Shadlen et al. 1996). The investigators explored systematically many of the conditions required for the observed relationships between neuronal activity and perceptual decisions. In their model, the activity of two simulated pools of oppositely tuned MT neurons were averaged separately and then compared on each trial (Fig. 2A). The pool with the highest activity dictated the decision, rightward or leftward. The model was constrained by several variables, one of which was information about known interneuronal noise correlations. Noise correlations quantify the degree to which neurons covary in their responses to repeated presentations of a given stimulus and arise from shared afferent connections, and/or interconnectivity among neighboring neurons, and are possibly indicative of circuit connectivity (Cohen and Kohn 2011). Information about known CP values and known psychophysical thresholds, which quantify discrimination abilities, were the remaining constraints for the model. The investigators found that the best replications of the observed CP values and psychophysical thresholds occurred with noisy pools of ~100 weakly correlated neurons, many of which were suboptimally tuned for the motion stimulus. Noise correlations and pooling noise, were especially critical in simulating the observed CP values.

Figure 2B, *left*, shows the relationship the investigators uncovered as the essential determinant of CP. Neuron *k*'s CP was due to the strength of the correlation between its activity level and the average activity of the pool. As we discussed, many CP values are relatively low. Part of the explanation could be due to pooling noise, which effectively reduces the coupling strength between a pool's fluctuations and fluctuations in behavior. Correlated variability has the paradoxical implication that even neurons with no causal role in the decision could show a significant CP simply because their activity is correlated on a trial-to-trial basis to the activity of the larger pool (Fig. 2B, *left*) or to the activity of specific neurons that are causal to the decision. Such correlations between causal and acausal neurons may reflect associative connections between the neurons resulting from the perceptual history of the animal, which may benefit the animal in certain perceptual contexts (Parker 2013).

Correlated variability makes it difficult to determine whether a nonselective pooling scheme is used to reach a decision or whether a selective pooling scheme is used in which the most informative neurons are consulted for the decision. It should be noted that modifying the weights of the neuronal signals had no effect on CP values in their simulations, a result that may be related to the noise correlation structure imposed on the neuronal activity. Two types of noise were important in their simulations: *1*) correlated or dependent noise arising from

A   *Pooling Model*

$$\sup\{\mu_{\text{left}}, \mu_{\text{right}}\}$$

μ_{left} μ_{right}

Leftward-biased Pool   *Rightward-biased Pool*

B

$$CP_k \propto NC_k \mu_{\text{left}}$$

$$CP_k \propto \frac{1}{n}\sum_{j=1}^{n} NC_{kj} SC_{kj}$$

$$CP_k - \frac{1}{2} \propto \frac{1}{n}\sum_{j=1}^{n} (NC_{kj}\beta_j)$$

$$\propto \text{mean weighted } C_{l,l} \text{ - mean weighted } C_{l,r}$$

NC=0   NC > 0

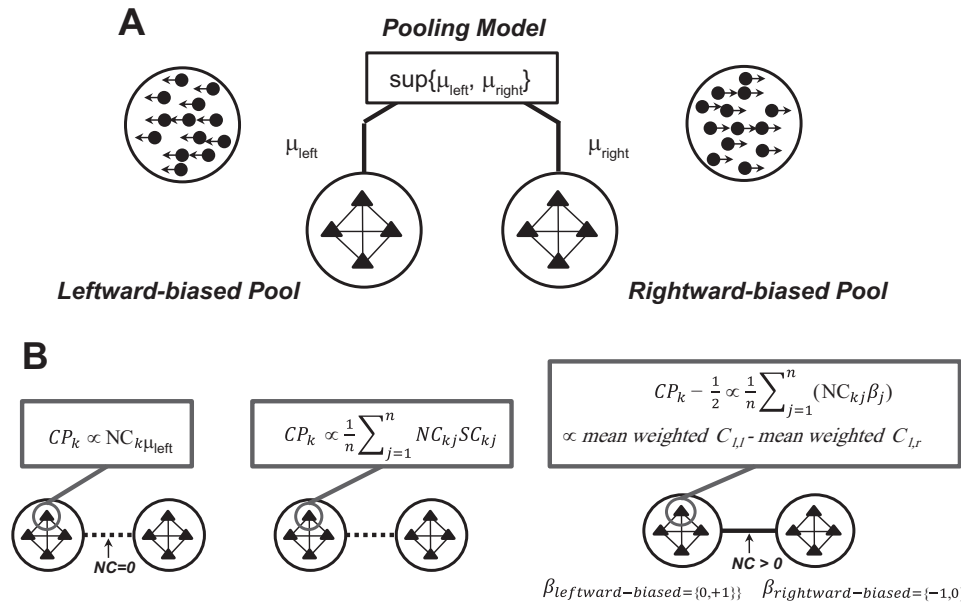$\beta_{leftward-biased=\{0,+1\}}$ $\beta_{rightward-biased=\{-1,0\}}$

Fig. 2. Pooling models, correlation structures and CP. *A*: schematic depiction of the classic pooling model informing most studies linking neuronal activity to decisions. In this scheme, illustrated here by MT activity during the random dot motion task, two rival pools of neurons (leftward-biased pool and rightward-biased pool) are averaged separately ($\mu_{\text{left-biased}}$ and $\mu_{\text{right-biased}}$) on each trial and then compared. The pool with the maximum activity ($\sup\mu_{\text{left-biased}}$, $\mu_{\text{right-biased}}$) dictates the choice. *B*: CP and noise correlation structure. *B, left*: in the classic model of Shadlen and colleagues, CP for an example neuron k, circled in gray, was a function of the noise correlation (NC) magnitude between neuron k's activity and the activity of the pool average ($\mu_{\text{left-biased}}$). Note that in this model, NC was zero between the rival pools. *B, middle*: a later model informed by Angelaki and colleagues (Liu et al. 2013) pointed out that in some cases CP is a function of not only the noise correlation between a neuron k and other members of the pool, but also the signal correlation structure between neuron k and other members of the pool. The dependence of CP on signal correlation structure indicates that selective connectivity among neurons with similar tuning curves is especially important in CP determination. *B, right*: a recent computational model derived a relationship between pooling weights, correlation structure, and CP. Pooling weights (β_j) appear prominently in the model, indicating that CP for neuron k is a function of the noise correlation strength between neuron k and each member of the pool and their respective weights (β_j) where j represents neuron j. Weights are a necessity given recent work indicating that NC between rival pools is nonzero. This suggests that CP is ultimately a function of the difference between the weighted distributions of the leftward-biased pool (here shown with positive weights) and the rightward-biased pool (shown with negative weights).

shared inputs and recurrent connectivity among neurons, and *2*) independent noise, arising from stochastic events related to neurotransmitter release and channel dynamics. The latter can be averaged away, whereas the former cannot. In their simulations, correlation coefficients, which represent the correlated variability or noise, were assigned randomly within a set range of values to neuronal pairs within each pool. We now know that noise correlations are a function of tuning curve similarity: neurons with similar tuning preferences tend to have higher noise correlations (Cohen and Kohn 2011). With this correlation structure, the suboptimally tuned neurons would have lower degrees of correlated activity relative to neurons tuned optimally for the stimulus direction of interest and thus show a minimal CP. The low CP, however, could be adjusted by manipulating the neuron's weight, effectively scaling up its low noise correlation with the pool. In this case, neuronal weights would be important. Below, we will discuss the importance of correlation structure and neuronal weights on CP determination.

A recent study by Angelaki and colleagues (Liu et al. 2013) reveals the importance of neuronal correlation structure for understanding CP. They recorded from vestibular nuclei (VTN), cerebellar nuclei (CN), and area MSTd, while monkeys performed a vestibular heading discrimination task and found that CP was much higher in the subcortical vestibular nucleus than in area MSTd. This is a surprising result given the tacit assumption that CP is determined by a feedforward integration of neuronal activity, since the vestibular nucleus is at the front

end of sensory perception and, therefore, should have the lowest CP values. The critical difference in CP value between the cortical and subcortical areas occurred because of the relationship between noise correlations and tuning curve correlations (also known as "signal correlations"). Noise correlations quantify how two neurons covary in their responses to repeated presentations of an identical stimulus and likely arise from inherited noise from common driving inputs and synaptic noise from their mutual connectivity. Tuning curve correlations quantify how similar the responses of two neurons are with respect to their stimulus preference. Two neurons with high tuning curve correlation means that their tuning curves are very similar and they likely receive similar driving inputs. In net, tuning curve correlations reveal input similarity and noise correlations reveal how much of those inputs are shared between the neurons.

Noise correlations and tuning curve correlations are closely related, implying that neurons with similar tuning either receive common driving inputs or are more strongly connected with each other, or both. Through simulations, Angelaki and colleagues found that higher CPs for subcortical neurons could be explained by a greater dependence of noise correlations on the degree of tuning similarity, as indicated by steeper slope relationships between noise correlations and tuning curve correlations. What this means is that, at least for subcortical nuclei, neurons receiving similar driving input assessed by tuning curve correlations also exhibit strong shared variability in trial-to-trial responses, indicating that these neurons are

strongly interconnected with each other and/or receive common driving inputs. These relations are shown schematically in Fig. 2B, *middle*. Neuron k's high CP is a result of the noise correlation structure between its activity and that of other neurons in the pool (NC_{kj}, where j represents the activity of neuron j), which in turn is conditioned by the strength of the tuning curve correlation between it and each member of the pool (SC_{kj}). A tight linkage of activity among similarly tuned neurons in the form of noise correlations is thus an important factor in determining high CP. This observation was supported further by the same authors in a causal followup study in which they compared vestibular nuclei CPs, noise correlations, and tuning curve correlations both before and after a unilateral lesion to the vestibular labyrinth, the principal input to the vestibular system. They hypothesized that the unusually high dependence on noise correlation-tuning curve relationships in the vestibular nuclei arises from the bilaterally converging inputs from each vestibular labyrinth. A unilateral lesion should affect these relationships. They found indeed that CPs and noise correlation-tuning curve correlation dependencies both decreased following the lesion. These results raise the possibility that a mechanism by which similarly tuned neurons, featuring high noise correlations, and by inference, putative strong interconnectivity, organize into functional micropools that are consulted during the decision process. These micropools would be weighted heavily, translating into high CPs, a conjecture supported by the work of Purushothaman and Bradley (2005). Computational studies together with the development and increasingly widespread use of multiple neuron recording techniques will provide critical data to address the relationships between neuronal weights and CP values. A recent computational study, which we discuss next, offers a first step toward inferring a population's weight distribution.

With the advent of multiple electrode recording techniques, it is now possible to acquire rich information about a neuronal population's correlation structure. Together with CP values calculated from the activity of all members of the population, we have two pieces of the puzzle for distinguishing between possible pooling models. The critical missing piece of information is the distribution of neuronal weights, that is, how is the activity of individual neurons weighted when forming a decision? With this information we can determine whether or not the activity of neurons is pooled selectively. A recent computational study sheds light on this by deriving an analytical relationship between correlation structure, neuronal weights and CP (Haefner et al. 2013). The first order approximation is:

$$CP_k \approx \frac{1}{2} + \frac{\sqrt{2}}{\pi} \frac{(C\beta)_k}{\sqrt{C_{kk}\beta^T C\beta}}, \qquad (1)$$

where,
$$(C\beta)_k = \sum_{j=1}^{n} C_{kj}\beta_j$$
$$\beta^T C\beta = \sum_{k=1}^{n} \sum_{j=1}^{n} \beta_k C_{kj}\beta_j$$

Equation 1 shows that the CP value for neuron k is a function of both the noise covariance matrix C, which characterizes how the response of neuron k covaries with that of all other neurons, and the weight distribution β, which describes how the activity of neurons in the population is pooled. The resulting numerator $(C\beta)_k$ is normalized by $C_{kk}\beta^T C\beta$, which is the total summed variance of the activity of all neurons in the population, scaled by their weights, plus the weighted covariance of the activity of all possible pairs in the population. The numerator provides information about whether all neurons have the same CP or whether they vary depending on the degree of tuning. In other words, it provides information about the total population's CP "shape." The denominator provides a normalization factor that determines the magnitude of this shape by factoring in the total variance of the population, the total weight distribution, and the total covariance of all possible neuronal pairs. To provide some intuition about this, *Eq. 2* shows the linear portion of *Eq. 1*:

$$CP_k - \frac{1}{2} \propto \frac{1}{n} \sum_{j=1}^{n} C_{kj}\beta_j \qquad (2)$$

Equation 2 states that the CP for neuron k is proportional to the average of neuron k's correlation with all neurons in the population (C_{kj}), where j represents neuron j, and their respective weights (β_j). This implies that the larger the population, the less neuron k's weight contributes to the decision. Therefore, CP is a function of population size, correlation structure, and the weight distribution (shown schematically in Fig. 2B, *right*). Haefner et al. (2013) assessed this model of CP through simulations involving neurons with varying tuning curves, response variances, interneuronal correlations and belonging to populations of different sizes with different weight distributions. For each set of parameters, a psychophysical decision based on the random dot motion task was simulated on a trial-by-trial basis by taking the sum of the activity of all neurons in the population, scaled by each neuron's weight. This decision rule, used commonly in the field, e.g., Law and Gold (2009), is linear and appears below:

$$\text{decision} = \sum_{k=1}^{n} \beta_k r_k \qquad (3)$$

where β is the weight of neuron k and r is the response of neuron k. Within the context of the random dot motion task in which the animal is discriminating rightward from leftward motion, all neurons with a greater bias toward leftward tuning would belong to a leftward-biased pool and receive a positive weight, and all neurons with a greater bias toward rightward tuning would receive a negative weight within a rightward-biased pool. The magnitude of each neuron's weight depends on the pooling scheme used (uniform, selective, or optimal). For a uniform scheme, all neurons belonging to the leftward-biased pool receive identical positive weights and all neurons belonging to the rightward pool receive identical negative weights. For a selective scheme, neurons with tuning matching exactly the two possible motion directions receive maximal weights, and all other neurons receive a weight of zero. For an optimal pooling scheme, the weight magnitude is informed by a neuron's tuning similarity to the motion direction and also the noise correlation structure of the population. If the total weighted sum is positive, then the model reports a leftward choice and vice versa for a negative weighted value. Through these simulations the authors were able to generate CP values while manipulating such variables as population size, correlation structure and weights. They found that CP was dependent on correlation structure and population size. Consistent with recent work (Cohen and Newsome 2009), they imposed positive noise correlations between rival pools. This means that a neuron preferring rightward motion could have a non-zero noise correlation with a neuron preferring leftward motion

indicating that shared variability cannot be averaged out. Given these facts, and assuming a differencing operation between rival pools (but see Heeger et al. 1999 for a different model), the sign and magnitude of the weights do matter in determining CP, since the sign of β_j provides information about how neuron j's noise correlation with neuron k should be weighted. Ultimately, then, *Eq. 1* tells us:

$$CP_k - \frac{1}{2} \propto \text{mean weighted } C_{k,\text{leftward-biased}(+)}$$
$$- \text{mean weighted } C_{k,\text{rightward-biased}(-)} \quad (4)$$

This means that CP for neuron k, with its leftward-bias (positive weight), is proportional to the difference between the mean weighted correlation structure of neuron k and all neurons with a leftward bias (positive weights), and the mean weighted correlation structure of neuron k and all neurons with a rightward bias (negative weights) (Fig. 2*B*, *right*). Relative to neuron k's leftward stimulus preference, all neuron k's weighted correlations with leftward-biased neurons have positive values, whereas each of neuron k's weighted correlations with the rightward-biased neurons have negative values. In principle, a strong noise correlation between the activities of two oppositely tuned neurons would hinder a decision, since the activity of each neuron would contribute to either decision and result in a reduced CP. The relationships between weights, noise correlations and CP are explored in more detail in a recent review (Nienborg et al. 2012).

Future work will undoubtedly make use of the computational scheme introduced by Haefner and colleagues (2013) to determine whether the activity of large numbers of neurons with disparate tuning preferences and variable weights is consulted at the decision stage or whether the activity of select subsets of highly weighted neurons is the real determinant of choice. With multielectrode recording technology, rich information about the correlation structure can be obtained and the computation of CPs for hundreds of individual neurons is possible, providing the critical data needed. As Haefner and colleagues show, this information makes it possible in principle to distinguish between uniform, selective, and optimal pooling schemes since each scheme predicts a specific CP profile given the existing correlation structure [but see Pitkow et al. (2015) for some possible complications]. Below we review further evidence that supports a selective pooling scheme for decision making.

The Origin of CP: Feedforward or Feedback?

A prominent model of CP assumes that neuronal activity correlated with behavior is a consequence of feedforward processing. According to this model, CP values are intrinsic to an area and are characterized by the pooling of activity from neurons within that area alone. One prediction of the feedforward model is that the correlation structure critical for CP should be driven by feedforward inputs. A recent causal study put this prediction to the test by reversibly inactivating areas V2/V3 and monitoring MT single units while monkeys alternately performed a motion direction detection task or a depth change detection task (Smolyanskaya et al. 2015). Area MT receives rich disparity information but no explicit motion information from areas V2/V3. The prediction was that during the depth task MT neurons would show a reduction in noise correlation structure and therefore CP following V2/V3 cooling due to the removal of highly structured and correlated disparity tuned inputs from V2/V3. Motion detection task correlation structure, however, should remain intact since V1 provides strong directionally selective information to MT directly, bypassing the site of inactivation. The investigators found that CP values were reduced during the depth task but not the motion task, a result that they traced to a reduction in the correlation structure in MT during depth change detection. These results provide evidence that some portion of CP and the correlations it depends on is derived from feedforward inputs from early processing stages. As we discuss next, however, some of these feedforward inputs may be caused by feedback signals from downstream processing stages that ultimately determine the CP activity.

A second model prediction is that the temporal course of CP activity should track how a subject weighs information in a stimulus. For example, weighing stimulus information heavily in the early phase of a trial should be associated with high CP values early in the course of a trial and weighing stimulus information less strongly at later phases of the trials should be associated with lower CP values at the same time points in the trial. This hypothesis was tested directly in a recent study by Nienborg and Cumming (2009). Using a white-noise analysis, the investigators correlated neuronal activity in V2 with frame to frame changes in disparity in a random dot stereogram. Trained monkeys indicated whether the stimulus was "near" or "far." This task design enabled the investigators to construct psychophysical kernels indicating which portion of the trial the monkey weighted most heavily to arrive at its decision. They found that monkeys weighted information early in the trial more heavily than information later in the trial. Surprisingly, an analysis of the temporal course of CP revealed that CP started out low and gradually increased as the trial progressed. This result indicates that at least in V2 high CPs appear later than would be expected from a feedforward model, as if driven by a delayed feedback signal. According to this interpretation, V2 choice-related activity is not causal to the decision but represents a time-lagged copy of the decision process occurring elsewhere. An earlier study observed similar CP dynamics in area MST during a motion discrimination task. CP values peaked during the second half of the stimulation epoch rather than the first half (Celebrini and Newsome 1994). These data provide evidence against a feedforward model of CP generation.

A third prediction of the feedforward model is that CP should vary in magnitude depending on the strength of the sensory evidence on which the animal bases its decision. In most perceptual decision tasks, the stimulus varies in difficulty or sensory strength, and the strength of the sensory stimulus influences neuronal activity differently; strong signals will drive neuronal activity well, whereas weak signals will drive neuronal activity poorly. Neurometric functions show that most neurons respond maximally to strong sensory signals and less so to weak sensory signals. The feedforward model of CP predicts that this same trend should extend to CP, since the information leading to a decision varies in strength and therefore ambiguity. Weak stimuli should result in low CP values and strong stimuli should lead to high CP values. However, CP values tend to be relatively constant across varying stimulus

difficulty levels, suggesting that something other than feedforward pooling of neuronal activity influences CP.

A second and polar opposite view of CP is one based on feedback. In this model, CP results from signals related to attention or decision making in other areas. Romo and colleagues found evidence for this idea in their recordings from the somatosensory and motor system while monkeys performed a vibrotactile discrimination task (Hernandez et al. 2010). Secondary somatosensory area S2 had significant CP values, but the latency of onset of the significant CP values lagged the latency measured in prefrontal and premotor cortical areas by tens of milliseconds, suggesting that CP developed in later areas before earlier areas. Similarly, experiments performed in primary auditory cortex (area A1) revealed significant ramps in CP just before the behavioral report, presumably after the decision was made (Niwa et al. 2013).The late occurrence of significant CP values in a primary sensory area point toward a role for feedback in the generation of CP.

Equation 1 indicates that CP is a function of a population's correlation structure and weight distribution. That a feedback signal could cause CP implies that the correlation structure and pooling scheme is under strong control by attention or a feedback signal. Indeed, recent work reveals that attention significantly influences the structure of correlations among neurons. When animals deploy attention to targets, interneuronal noise correlations between the activity of neurons belonging to the same pool decrease (Cohen and Maunsell 2009; Mitchell et al. 2009; Ruff and Cohen 2014) while noise correlations between the activity of pairs of neurons belonging to rival pools increase (Ruff and Cohen 2014). The latter finding means that with attention during the motion direction decision task, noise correlations would increase between the activity of neurons preferring leftward and rightward motion. Figure 3 shows these attention-induced effects for hypothetical data obtained from two rival pools of neurons recorded during the motion task, one codes for leftward-biased motion and the other codes for rightward-biased motion. According to *Eq. 1*, the attention-induced alteration has the paradoxical prediction that CPs should be relatively low among neurons belonging to a leftward-biased or rightward-biased pool, since the noise correlation structure (NC$_{leftward-biased,leftward-biased}$; Fig. 3) is weakened among neurons belonging to the leftward-biased pool, and strengthened between the leftward-biased pool and rightward-

biased pool (NC$_{leftward-biased,rightward-biased}$; Fig. 3). This relation is a consequence of *Eq. 1*, which indicates that a neuron's CP is determined by the strength of its noise correlation with each member of the population and the weights of the remaining neurons within the population. Since the noise correlation strength between the leftward-biased pool and the rightward-biased pool is non-zero, this introduces negative weights into the calculation since the rightward-biased pool neurons have negative weights relative to a leftward stimulus. This, together with the reduction in correlation strength among neurons in the same pool, would result in a reduced CP value.

The above reasoning may explain why observed CP values are typically low. Attention alters the correlation structure and interferes with a neuron's relationship to behavior. However, there are many examples of high CP values. The hypothetically low CP therefore is contrary to the actual CP values observed. Examination of *Eq. 1* reveals that there are several ways that a network could reconfigure itself to generate a high CP given the alteration in correlation strength induced by attention. Each of the hypothetical mechanisms we illustrate has implications about the possible pooling mechanism used by neuronal populations when correlational structures are modified by cognitive signals such as attention. One possibility is that attention dynamically alters the weights that offset the alteration in correlation structure. Neurons are known to undergo changes in discharge rate with attentional allocation (Maunsell and Cook 2002), so the increase in neuronal activity could be the implementation of enhanced weights. A second possibility is that attentional allocation engages the most sensitive neurons. Such selective pooling would adjust the weight distribution of the population by carving out a select subset of it (a micropool) and thereby compensate for the alteration of interneuronal noise correlation structure. This would result in a distribution of CP values, some low, some high, with the highest presumably reflecting those neurons that are receiving the highest weight in the decision. The wide range of CP values reported in many studies and the correlations between sensitivity and CP lend support to this second possibility and highlight the importance of modifiable weight distributions in compensating for the effects of altered correlation structure.

A third possibility focuses on the well-known relations between tuning curve correlations and noise correlations. As we described above, the findings of Angelaki and colleagues
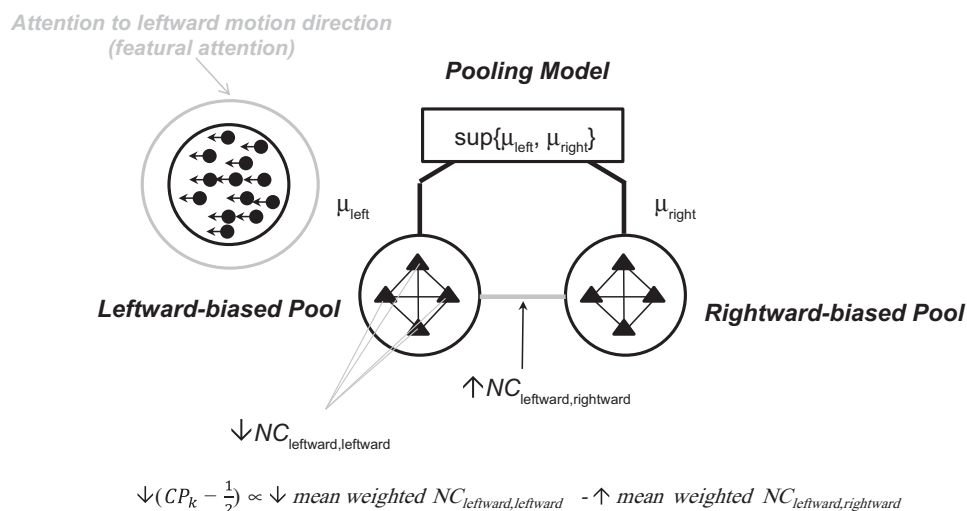

Fig. 3. Attention, correlation structures, and CP. Attentional allocation to a particular stimulus, here to leftward motion (represented by the gray circle), has significant effects on correlation structure and consequently on CP. Attention has been shown to decrease noise correlations between neurons belonging to the same pool (e.g., leftward-biased pool; ↓NC$_{leftward-biased,leftward-biased}$) and to increase correlations between neurons belonging to rival pools (leftward-biased pool and rightward-biased pool; ↑NC$_{leftward-biased,rightward-biased}$). According to *Eq. 1* (see text and Fig. 2B, *right*), this has the effect of decreasing the expected CP values of neurons belonging to the leftward-biased pool, during trials in which leftward motion direction is shown.

Attention to leftward motion direction (featural attention)

Pooling Model

$$\sup\{\mu_{left}, \mu_{right}\}$$

μ_{left} μ_{right}

Leftward-biased Pool   **Rightward-biased Pool**

$\uparrow NC_{leftward,rightward}$

$\downarrow NC_{leftward,leftward}$

$$\downarrow \left(CP_k - \frac{1}{2}\right) \propto \downarrow mean\ weighted\ NC_{leftward,leftward} - \uparrow mean\ weighted\ NC_{leftward,rightward}$$

(Liu et al. 2013) indicate that CP is dependent on the relationship between the tuning curve correlation structure and the noise correlation structure of the population. Computational work shows that neuronal populations convey more information when the tuning curve correlation structure and noise correlation structure move in opposite directions: a relatively high tuning curve correlation coupled with a relatively low noise correlation leads to better discrimination (for a review see, Averbeck et al. 2006). A possibility, then, is that with attention the decrease in noise correlation structure is accompanied by an increase in tuning curve correlation. According to the computational work, this would result in better discrimination capabilities of the population and lead to the oft-observed improved perceptual performance. There is evidence that attention can shift tuning curves towards values closer to the attended stimulus value (Connor et al. 1996, 1997; David et al. 2008; Womelsdorf et al. 2006, 2008). This type of alteration would lead to increases in signal correlations among the activity of neurons since the tuning curves are shifting towards a common value. A direction for future inquiry would be to examine not only the noise correlation structure but also the signal correlation structure during attentional allocation to understand the mechanistic underpinnings of CP generation.

The Future of CP

CP is a valuable metric and its introduction led to a boon in knowledge of how the brain makes decisions. However, as with any scientific advance, the introduction of CP also raised many questions. Recent technological advances in systems neuroscience, such as multiple neuron recording, make these questions infinitely more tractable. For example, our knowledge of correlations between the activity of neurons and how they change with task demands is incomplete. Yet, this information is the key to understanding CP. What we know now suggests that much of the correlation structure required for CP results from well-structured and formatted signals generated locally or inherited from other areas that code explicitly for the stimulus feature used (Nienborg and Cumming 2014; Smolyanskaya et al. 2015). Is this similar for all areas with CP? Does the structure of activity correlations leading to CP vary with each brain area? A related question is the role of attention in causing CP, which presumably does so by altering the correlation structure on which CP strongly depends. We know attention reduces noise correlations, but this appears paradoxical given the importance of strong noise correlation structure on CP. Does attention increase tuning curve correlations to compensate for noise correlation reduction? Other questions concern the locational and magnitude trends of CP. For example, why do CPs peak early in the vestibular system only to plummet at later stages? Is the early peak related to the explicit coding format of the vestibular system, where explicit signals related to heading discrimination are already present in the vestibular nuclei? CPs in the vestibular pathway decrease by the time of MT. Is this because MT lacks a topographic map for heading discrimination? Neurons in nontopographically arranged pools share fewer common inputs with neighboring neurons and thus presumably share less correlated information with each other through recurrent connectivity. This could result in a relatively weak correlation structure, translating into lower CPs.

A final question concerns the origin of CP: is it due to feedforward or feedback signaling? The answer likely resides somewhere in the middle, reflecting contributions of both feedforward and feedback signals in CP determination (Wimmer et al. 2015). The relative contribution of each signal could be assessed by recording from multiple areas simultaneously while animals engage in perceptual discrimination tasks. This would provide trial-by-trial information about the temporal dynamics and relationships between areas as CP evolves. Causal experiments such as the one performed by Born and colleagues discussed above should be performed to assess the relative importance of feedback signals on CP calculated from activity recorded from upstream areas. One possible experiment would involve stimulating in a downstream area closer to the decision process while an animal makes perceptual decisions and noting the changes in CP that occur in upstream areas in consequence. This would be an explicit causal test of the role of feedback signaling on CP. In sum, these are just a few of the questions that are left open, many more will likely arise as investigators continue to utilize and explore the CP metric in the quest to understand how neuronal activity gives rise to perception.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: T.B.C. and M.A.B. drafted manuscript; T.B.C. and M.A.B. edited and revised manuscript; T.B.C. and M.A.B. approved final version of manuscript; T.B.C. prepared figures.

REFERENCES

Adrian E. *The Basis of Sensation: The Action of the Sense Organs*. London: Christophers, 1928.

Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. *Nat Rev Neurosci* 7: 358–366, 2006.

Barlow HB. Single units and sensation: a neuron doctrine for perceptual psychology? *Perception* 1: 371–394, 1972.

Bradley A, Skottun BC, Ohzawa I, Sclar G, Freeman RD. Visual orientation and spatial frequency discrimination: a comparison of single neurons and behavior. *J Neurophysiol* 57: 755–772, 1987.

Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis Neurosci* 13: 87–100, 1996.

Britten KH, Shadlen MN, Newsome WT, Movshon JA. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12: 4745–4765, 1992.

Celebrini S, Newsome WT. Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. *J Neurosci* 14: 4109–4124, 1994.

Cohen MR, Kohn A. Measuring and interpreting neuronal correlations. *Nat Neurosci* 14: 811–819, 2011.

Cohen MR, Maunsell JH. Attention improves performance primarily by reducing interneuronal correlations. *Nat Neurosci* 12: 1594–1600, 2009.

Cohen MR, Newsome WT. Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *J Neurosci* 29: 6635–6648, 2009.

Connor CE, Gallant JL, Preddie DC, Van Essen DC. Responses in area V4 depend on the spatial relationship between stimulus and attention. *J Neurophysiol* 75: 1306–1308, 1996.

Connor CE, Preddie DC, Gallant JL, Van Essen DC. Spatial attention effects in macaque area V4. *J Neurosci* 17: 3201–3214, 1997.

David SV, Hayden BY, Mazer JA, Gallant JL. Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. *Neuron* 59: 509–521, 2008.

de Lafuente V, Romo R. Neuronal correlates of subjective sensory experience. *Nat Neurosci* 8: 1698–1703, 2005.

de Lafuente V, Romo R. Neural correlate of subjective sensory experience gradually builds up across cortical areas. *Proc Natl Acad Sci USA* 103: 14266–14271, 2006.

Dodd JV, Krug K, Cumming BG, Parker AJ. Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *J Neurosci* 21: 4809–4821, 2001.

Evarts EV. Relation of pyramidal tract activity to force exerted during voluntary movement. *J Neurophysiol* 31: 14–27, 1968.

Faisal AA, Selen LP, Wolpert DM. Noise in the nervous system. *Nat Rev Neurosci* 9: 292–303, 2008.

Ghose GM, Harrison IT. Temporal precision of neuronal information in a rapid perceptual judgment. *J Neurophysiol* 101: 1480–1493, 2009.

Green DM, Swets JA. *Signal Detection Theory and Psychophysics.* New York: Wiley, 1966.

Gu Y, Angelaki DE, Deangelis GC. Neural correlates of multisensory cue integration in macaque MSTd. *Nat Neurosci* 11: 1201–1210, 2008.

Haefner RM, Gerwinn S, Macke JH, Bethge M. Inferring decoding strategies from choice probabilities in the presence of correlated variability. *Nat Neurosci* 16: 235–242, 2013.

Hartline HK, McDonald P. Light and dark adaptation of single photoreceptor elements in the eye of Limulus. *J Cell Physiol* 30: 225–253, 1947.

Heeger DJ, Boynton GM, Demb JB, Seidemann E, Newsome WT. Motion opponency in visual cortex. *J Neurosci* 19: 7162–7174, 1999.

Hernandez A, Nacher V, Luna R, Zainos A, Lemus L, Alvarez M, Romo R. Decoding a perceptual decision process across cortex. *Neuron* 66: 300–314, 2010.

Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160: 106–154, 1962.

Johansson RS, Vallbo AB. Detection of tactile stimuli. Thresholds of afferent units related to psychophysical thresholds in the human hand. *J Physiol* 297: 405–422, 1979.

Kiang NY, Pfeiffer RR, Warr WB, Backus AS. Stimulus coding in the cochlear nucleus. *Ann Otol Rhinol Laryngol* 74: 463–485, 1965.

Kim B, Basso MA. Saccade target selection in the superior colliculus: a signal detection theory approach. *J Neurosci* 28: 2991–3007, 2008.

Law CT, Gold JI. Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat Neurosci* 11: 505–513, 2008.

Law CT, Gold JI. Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nat Neurosci* 12: 655–663, 2009.

Liu J, Newsome WT. Correlation between speed perception and neural activity in the middle temporal visual area. *J Neurosci* 25: 711–722, 2005.

Liu S, Gu Y, DeAngelis GC, Angelaki DE. Choice-related activity and correlated noise in subcortical vestibular neurons. *Nat Neurosci* 16: 89–97, 2013.

Logothetis NK, Schall JD. Neuronal correlates of subjective visual perception. *Science* 245: 761–763, 1989.

Maunsell JH, Cook EP. The role of attention in visual processing. *Philos Trans R Soc Lond B Biol Sci* 357: 1063–1072, 2002.

Mitchell JF, Sundberg KA, Reynolds JH. Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron* 63: 879–888, 2009.

Mountcastle VB, Talbot WH, Darian-Smith I, Kornhuber HH. Neural basis of the sense of flutter-vibration. *Science* 155: 597–600, 1967.

Newsome WT, Britten KH, Movshon JA. Neuronal correlates of a perceptual decision. *Nature* 341: 52–54, 1989.

Nienborg H, Cohen MR, Cumming BG. Decision-related activity in sensory neurons: correlations among neurons and with behavior. *Annu Rev Neurosci* 35: 463–483, 2012.

Nienborg H, Cumming BG. Macaque V2 neurons, but not V1 neurons, show choice-related activity. *J Neurosci* 26: 9567–9578, 2006.

Nienborg H, Cumming BG. Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature* 459: 89–92, 2009.

Nienborg H, Cumming BG. Decision-related activity in sensory neurons may depend on the columnar architecture of cerebral cortex. *J Neurosci* 34: 3579–3585, 2014.

Niwa M, Johnson JS, O'Connor KN, Sutter ML. Differences between primary auditory cortex and auditory belt related to encoding and choice for AM sounds. *J Neurosci* 33: 8378–8395, 2013.

Parker A, Hawken M. Capabilities of monkey cortical cells in spatial-resolution tasks. *J Opt Soc Am A* 2: 1101–1114, 1985.

Parker AJ. A micro-pool model for decision-related signals in visual cortical areas. *Front Comput Neurosci* 7: 115, 2013.

Parker AJ, Krug K, Cumming BG. Neuronal activity and its links with the perception of multi-stable figures. *Philos Trans R Soc Lond B Biol Sci* 357: 1053–1062, 2002.

Pitkow X, Liu S, Angelaki DE, DeAngelis GC, Pouget A. How can single sensory neurons predict behavior? *Neuron* 87: 411–423, 2015.

Price NS, Born RT. Timescales of sensory- and decision-related activity in the middle temporal and medial superior temporal areas. *J Neurosci* 30: 14036–14045, 2010.

Purushothaman G, Bradley DC. Neural population code for fine perceptual decisions in area MT. *Nat Neurosci* 8: 99–106, 2005.

Romo R, Hernandez A, Zainos A, Lemus L, Brody CD. Neuronal correlates of decision-making in secondary somatosensory cortex. *Nat Neurosci* 5: 1217–1225, 2002.

Ruff DA, Cohen MR. Attention can either increase or decrease spike count correlations in visual cortex. *Nat Neurosci* 17: 1591–1597, 2014.

Sasaki R, Uka T. Dynamic readout of behaviorally relevant signals from area MT during task switching. *Neuron* 62: 147–157, 2009.

Shadlen MN, Britten KH, Newsome WT, Movshon JA. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J Neurosci* 16: 1486–1510, 1996.

Siebert W. Frequency discrimination in the auditory system: place or periodicity mechanism. *Proc IEEE* 58: 723–730, 1970.

Skottun BC, Bradley A, Sclar G, Ohzawa I, Freeman RD. The effects of contrast on visual orientation and spatial frequency discrimination: a comparison of single cells and behavior. *J Neurophysiol* 57: 773–786, 1987.

Smolyanskaya A, Haefner RM, Lomber SG, Born RT. A modality-specific feedforward component of choice-related activity in MT. *Neuron* 87: 208–219, 2015.

Tolhurst DJ, Movshon JA, Dean AF. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res* 23: 775–785, 1983.

Uka T, DeAngelis GC. Contribution of area MT to stereoscopic depth perception: choice-related response modulations reflect task strategy. *Neuron* 42: 297–310, 2004.

Vogels R, Orban GA. Quantitative study of striate single unit responses in monkeys performing an orientation discrimination task. *Exp Brain Res* 84: 1–11, 1991.

Williams ZM, Elfar JC, Eskandar EN, Toth LJ, Assad JA. Parietal activity and the perceived direction of ambiguous apparent motion. *Nat Neurosci* 6: 616–623, 2013.

Wimmer K, Compte A, Roxin A, Peixoto D, Renart A, de la Rocha J. Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area MT. *Nat Commun* 6: 6177, 2015.

Womelsdorf T, Anton-Erxleben K, Pieper F, Treue S. Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nat Neurosci* 9: 1156–1160, 2006.

Womelsdorf T, Anton-Erxleben K, Treue S. Receptive field shift and shrinkage in macaque middle temporal area through attentional gain modulation. *J Neurosci* 28: 8934–8944, 2008.