Looking for cognition in the structure within the noise

Adam Johnson¹, André A. Fenton², Cliff Kentros³ and A. David Redish⁴

¹ Bethel University, St. Paul, MN 55112, USA
² SUNY, Downstate Medical Center, Brooklyn, NY 11203, USA
³ University of Oregon, Eugene, OR 97403, USA
⁴ University of Minnesota, Minneapolis, MN 55455, USA

Neural activity in the mammalian CNS is determined by both observable processes, such as sensory stimuli or motor output, and covert, internal cognitive processes that cannot be directly observed. We propose methods to identify these cognitive processes by examining the covert structure within the apparent ‘noise’ in spike trains. Contemporary analyses of neural codes include encoding (tuning curves derived from spike trains and behavioral, sensory or motor variables), decoding (reconstructing behavioral, sensory or motor variables from spike trains and hypothesized tuning curves) and generative models (predicting the spike trains from hypothesized encoding models and decoded variables). We review examples of each of these processes in hippocampal activity, and propose a general methodology to examine cognitive processes via the identification of dynamic changes in covert variables.

Introduction

The standard neurophysiological approach to understanding how neuronal activity encodes information is to examine neural activity while a subject is repeatedly presented with the same stimulus, or performs multiple trials of the same behavior. Typically, the relationship of a single neuron’s activity to an overt variable is described by the tuning curve of the cell, which is constructed under the assumption that noise sources are independent of the overt variable of interest and will, thus, average to zero in the tuning curve. In addition to this noise assumption, the standard approach also assumes that neurons independently contribute to the representation of the overt variable so that a neural representation can be understood by computing an average of the independent tuning curves.

Although the standard ‘encoding’ approach has been used to successfully characterize neural representations of overt variables in several areas of neuroscience, it has been more difficult to apply the encoding approach to identify representations of covert cognitive variables like attention, decisions and planning. One difficulty is found at the level of experimental design because such covert variables might not be expressed repeatedly and reliably, even given an identical set of external conditions. However, even beyond this experimental difficulty, the noise and independence assumptions that underlie the definition of tuning curves are at odds with the theoretical framework of ensemble representations (e.g. cell assemblies [1]). Within an ensemble framework, information is represented by conjoint neural activity distributed across a functionally defined neuronal subgroup or cell assembly composed of many neurons. Furthermore, each neuron in a cell assembly can participate in other cell assemblies that represent different overt and covert variables. According to this view, understanding a neural representation requires characterizing conjoint neural activity rather than just an average from individual neurons. It is also unlikely that a neuron will be tuned to a single variable. Both of these considerations are difficult to reconcile with the standard encoding approach based on tuning curves.

A growing body of evidence supports the cell assembly framework, especially in recent studies of the representation of location in hippocampal ensemble activity. Here, we consider evidence that what might seem to be noise in tuning curves actually results from operation of an experimentally unobserved (covert) variable or process. We focus on recent studies of hippocampal place cell representations of locations, which have gone beyond the encoding approach to provide compelling evidence of covert cognitive variables in hippocampal discharge. Based on these studies, we outline a general ‘generative’ approach for finding covert cognitive variables in the apparent noise of tuning curves.

An early example: decoding mental rotation in cortical discharge

An important study by Georgopoulos et al. [2] showed that coherent, cell assembly-like dynamics within a neuronal population were related to cognitive function. Monkeys were trained to reach towards a target selected on each trial from one of eight potential directions around a circle. However, on a subset of probe trials, a bright target indicated that the rewarded direction would be rotated 90 degrees from the signaled direction. Georgopoulos et al. [2] characterized the correspondence between observed spiking activity in the motor cortex immediately before reaching and the direction of the subsequent reach by computing tuning curves for a population of cells during the unrotated trials. Prior to reaching on the rotated probe trials, spiking within this population changed dynamically, such that the orientation represented by this population rotated 90 degrees. This study showed that apparent noise activity in the motor cortex was actually organized as
a coherent, dynamic representation of movements before reaching, and demonstrated that a covert cognitive process could be observed by characterizing the dynamics of spiking activity in a population of neurons (Figure 1).

Georgopoulos et al. [2] is an early demonstration that explanations of spiking activity in the neocortex by the overt stimulus and response properties of static tuning curves could be substantially improved by the addition of experimentally covert, dynamic cognitive variables. Although several more recent studies have shown spiking activity that is most readily explained as cell assembly dynamics in the neocortex [3–5], these investigations into cognition have been hampered by the long training times required for primate experiments (which tend to lead to the use of non-cognitive processes [6]). By contrast, many behavioral tasks that model cognition are quite amenable to unit recordings in rodents, often allowing for recordings throughout both acquisition and performance of the task [7,8]. The combination of these tasks, the highly selective spiking activity of hippocampal pyramidal cells and the well-established use of chronic recording technologies make the hippocampus a model system for studying the representation of cognitive variables in neural discharge.

**Dynamic spatial information in hippocampal ensembles**

The location of the animal relative to its environment is the clearest signal in the spiking activity of hippocampal pyramidal cells recorded in behaving rodents, leading to the adoption of the term ‘place cell’ [7–9]. Each pyramidal neuron primarily spikes when the animal is in a particular part of its environment (the cell’s ‘place field’), which indicates that the activity across the place cell population creates a map-like representation of the animal’s current position in an environment [7–9]. This idea led to the convention of constructing a spatial tuning curve for individual place cells as a time-averaged firing rate map of cell spiking as a function of position [10,11]. This time-independent firing rate map can be interpreted as a tuning curve of firing to spatial position. The strength of the spatial signal, the remarkable stability of these firing rate maps in constant environmental conditions and their response to well-defined changes of the spatial environment [8,10–13] led to the prevailing view that a given hippocampal pyramidal neuron is either a place cell or not active in a given environment [14]. Momentary spiking activity that deviated from the tuning curve of a cell was taken as noise and ignored by averaging across several minutes of recording.

Changes in place field firing (under, for example, environmental changes [12,13]) were identified as changes between multiple maps [15–18]. This hypothesis led to the suggestion that these transitory events in which spiking activity deviated from the tuning curve of a cell might reflect self-consistent information about other locations or other maps, presumably as the animal shifted its attention from its current position to other kinds of information. If this were true, then one should be able to detect these covert cognitive shifts between input streams by examining the ‘noise’ for statistical structure, even if the underlying parameters of that structure were unknown. Identifying cognitive function then entails finding the cognitive bases for the covert parameters: tying these often-fleeting sets of active neurons that coalesce and disappear to the psychological processes that they presumably underlie, such as memory, decision-making and attention. We review current progress towards these admittedly ambitious goals in the hippocampus and describe a general computational approach towards finding covert cognitive processes in the firing patterns of neural ensembles.

**Dynamics at long timescales**

Although we are mainly concerned with the dynamics of hippocampal neurons at the millisecond timescale, there is
abundant evidence for the dynamic influence of cognitive processes on the firing of hippocampal neurons at longer timescales, notably with regard to the property of place cells called ‘remapping’ [8,12,19]. Although a place cell can indeed have the same place field for months in a constant environment [20], unlike sensory neurons, place cells can also completely change their spatial tuning curves between experiences within an environment [13,19,21–23]. Although this remapping most commonly happens in response to changes in environmental cues (e.g. putting the animal in a novel environment), remapping is not a straightforward sensory transformation of the changed cues [8,13,23–30]. Moreover, place cells can remap without any change whatsoever in the available spatial cues, for instance when an animal changes its behavioral strategy [31], uses different coordinate systems [32–35] or as information about a context is acquired [36,37]. Often, the remapping can best be thought of as multiple stable states, with different environments or conditions associated with particular sets of place cells that are stably retrieved upon subsequent reintroductions to the appropriate environment, or as the animal switches between tasks or coordinate systems (Figure 2).

What might govern these shifts in the information reflected by hippocampal neurons, and how might they be tied to cognition? Kentros et al. [22] found that place fields in mice spontaneously remapped far more often than those of rats under the same simple behavioral conditions (chasing after randomly dropped food pellets in a familiar environment). However, mice recorded while accurately

Figure 2. Remapping depends on covert, cognitive variables. (a) Observations of place field distributions between two experiences on a task show remapping under some conditions but not others. (b) These effects are significantly modulated by cognitive variables, such as the presence of a spatial task. (c) Diagram of the identification of structure in the place cell remapping. Tuning curves (green) are generated from observations of spiking data and behavioral variables (orange). The likelihood of remapping is dependent on covert processes of attention, task and behavioral relevance. Part (a) reproduced, with permission, from Ref. [21]. Part (b) reproduced, with permission, from Ref. [22].
performing a spatial task in the same familiar environment had place fields that were as stable as those of rats. Kentros et al. [22] suggested that the key difference was that the animals doing the spatial task had reason to pay attention to the available spatial cues, so they remembered them, which was reflected in the stability of their place fields. These studies support the hypothesis that covert cognitive processes (such as what the animal is paying attention to) influence the long-term dynamics of place fields. However, attention typically operates at much shorter timescales, on the order of seconds to milliseconds. This suggests the possibility of finding task-related shifts in network activity in hippocampal neurons at these timescales as well.

Dynamics at short timescales
Might the covert variables that control map retrieval upon entering an environment also operate on short timescales to repeatedly switch between multiple maps in constant conditions? If the tuning curve maps were sufficiently different, almost any repetitive map switching would appear as noise in place cell tuning curves [38]. Indeed, in the simple foraging tasks in which stable place fields are typically observed, a place cell has unexpectedly high levels of spiking variability inside its place field [39]. In these tasks, a robust place cell might emit 20 or more action potentials on a single pass through a place field, but fail to emit any action potentials seconds later on a pass that is behaviorally indistinguishable. The statistics of these deviations are incompatible with the hypothesis that place cell activity simply varies randomly about a mean described by a single spatial tuning curve [38,39], and instead support a hypothesis that place cell activity reflects a small number of spatial tuning curves that differ mainly in firing rate and are alternatively switched on and off with a period of about one second [38,40,41] (Figure 3).

This proposal is analogous to the suggestion that the hippocampus maintains multiple spatial maps of the environment and somehow switches between those maps very quickly. Support for this proposal comes from studies by Harris et al. [42] and Jackson and Redish [41]. Harris et al. [42] initially showed that predicting the spiking activity of hippocampal place cells using both position and the spiking activities of a set of simultaneously recorded place cells was significantly better than predicting hippocampal place cell activity using position information alone. Harris [43] argued that the covariation of place cell spiking activity was evidence for use of multiple

![Figure 3](https://example.com/figure3.png)

*Figure 3.* During simple tasks, there is noise in place fields that can be explained as changes in covert, cognitive variables. (a) The noise in place field firing differs between attentive and goal-related conditions. In simple foraging tasks, there is high variability. In tasks with a goal, the variability decreases. Taking only firing during goal-approach produces variability consistent with the predicted tuning curves. (b) Using generative models, it is possible to split place cell firing rate maps apart using hypothesized covert variables (Abbreviations: S, split by covert variable; R, random shuffled control). (c) Diagram of the identification of structure as covert variables. Cell assembly dynamics indicate the presence of covert variables (purple). Tuning curves (green) are generated by combining observations of spiking data and behavioral variables (orange) with those hypothesized covert variables (purple). The hypothesized covert variables explain large portions of the residual ‘noise’ (blue). Part (a) reproduced, with permission, from Ref. [38]. Part (b) reproduced, with permission, from Ref. [41].
cell assemblies within the hippocampus. Jackson and Redish [41] showed that coherent fast switching between multiple hippocampal cell assemblies could explain the excess variability observed within place cell spiking activity observed by Fenton and Muller [39]. Furthermore, Jackson and Redish [41] showed that fast switching between cell assemblies was clearly aligned to specific behavioral phases in certain tasks and produced multiple and distinct tuning curve maps. The cell assemblies observed on the linear track, for instance, were generally aligned with the running direction of the animal, and their projection onto spatial position was apparent as directional place fields. Although directional place fields have been previously explained as indicative of multiple reference frames (maps) in linear track tasks [10,17,18,32,41,44], Jackson and Redish [41] showed that reference frames are not specific to linear track tasks and can explain the excess variability first identified by Fenton and Muller [39].

The high levels of place cell spiking variability is unlikely to be noise because it has an across cell organization that can be explained as coordinated activity [41–43]. These observations further indicate that internal, unobservable or covert processes mediate the active tuning curve of the cell and, consequently, determine the discharge of the cell at that moment. Observations that these cell assembly dynamics are modulated by cognitive demands [22,38] and aligned to specific task components [41] indicate that cell assembly dynamics are better described as a reflection of covert cognitive processes than of noise.

**Extra-field spikes during sleep, rest and directed behavior**

Although place fields described pyramidal cell firing during awake behavior, these same pyramidal cells fired during specific sleep states (e.g. during sharp-waves occurring within slow-wave sleep and during REM sleep [7,45]). This sleep-related firing was difficult to explain from a traditional tuning-curve perspective. Similarly, place fields fire extra spikes outside their place fields during rest, grooming, eating and other, non-attentive pausing behaviors [7]. Subsequent studies showed that during sleep states these extra-field spikes entail a reactivation of place cell firing sequences that is both reliable and coherent. Cell pairs, ensembles and the temporal order therein that were active during awake behavior are reactivated during subsequent sleep states [46–52] – hippocampal pyramidal cell spiking activity continues to be organized with respect to the representation of previously experienced space during sleep. The firing of hippocampal cells during sleep is better described as ‘replay’ than as noise (Figure 4).

Decoding algorithms applied to neural ensembles found that the decoded location during awake rest states deviated from the observed location of the rat [53], but Jensen and Lisman were not able to identify any structure in these processes and concluded that they were noise. Recent studies, however, have found that these extra-field spikes occurring during rest can be understood as reactivation of recently experienced behaviors [54–59]. These results indicate that the noise identified by Jensen and Lisman [53] actually contains structure and reflects a covert (cognitive) event.

Johnson and Redish [60] examined place cell firing at a decision point. At difficult decision-points, rats pause but remain attentive to their surroundings: they turn back and forth, orienting down potential choices in a process termed vicarious trial-and-error (VTE [61]). During these behaviors, the hippocampus remains in an active theta state (similar to running behavior [7,60,62]). Johnson and Redish [60] found that during these behaviors, place cells fired spikes even if the animal was outside of the cell’s place field, at a location where the tuning curve predicted zero spikes [60]. Decoding activity during these behaviors revealed a sequential sweep of positions from the current position of the animal to potential future positions on each arm. The non-local forward representations contained sequential structure, they were predominantly ahead of the animal and were related to the orientation of the animal during the VTE behavior. These data indicate that place cell activity that occurs outside of a cell’s place field signals a covert process related to the prediction of potential spatial positions available to the animal rather than simple noise (Figure 5).

**Discussion: encoding, decoding and generative approaches**

The studies reviewed earlier indicate a correspondence between covert, cognitive variables and deviation of neural activity from the tuning curve of a cell. The conclusions from these studies result from at least one of three distinct approaches to analysis, ‘encoding’, ‘decoding’ and ‘generative’. Each approach is a distinct way to evaluate the validity of a hypothesized neural code and determine how well a variable can be related to neural activity [59] (Box 1).

**Encoding approaches**

If a cell reliably changes its firing as a function of different stimulus conditions or differences in a behavioral parameter, then one can say that the cell is tuned to the parameter in question [63]. The tuning curve thus describes how activity represents information about the parameter [64]. Within the encoding approach, validity of a neural code is based on how much information the neural code provides about the variable in question [63–65]. The amount of information can be measured using either Shannon information [64] or Fisher information [63]. However, information measures depend on the concept of a signal separated from noise, and they all effectively measure the extent to which the signal can be differentiated from an unstructured, unexplainable noise component. Identifying covert, cognitive processes using encoding approaches depends on making covert processes overt by aligning cognitive events to behavioral tasks so that they always occur at the same time on each trial or by creating tasks that differ only in the hypothesized covert processes. The encoding approach has enjoyed some success within tightly controlled behavioral experiments on attention [22], categorization [66] and various other
Figure 4. Replay after behavior is an example of a covert, cognitive event identified by clear structure within extraneous spiking. (a) Direct examination of cell firing shows that cells fire in the same sequence during post-behavior sleep (bottom) as the place fields during behavior (top). (b) Direct observation of firing during sharp-waves during awake states shows sequential replay akin to that seen during sleep. (c) Diagram of the identification of structure in the noise. Tuning curves (green) are generated from observations of spiking data and behavioral variables (orange, top). When compared to observations of spiking data from sleep states (orange, bottom), these tuning curves can be used to decode represented position (blue). Structure in the decoding leads to the conclusion that there are covert variables (replay). (d) Diagram of the identification of structure in the noise during awake states. Tuning curves (green) are generated from observations of spiking data and behavioral variables (orange, top). During sharp-waves, there is a mismatch between the spiking predicted by the tuning curves and the observed spiking. When decoding represented position (blue), structure in the decoding leads to the conclusion that there are covert variables (replay). Part (a) reproduced, with permission, from Ref. [49]. Part (b) reproduced, with permission, from Ref. [57].
aspects of decision-making [67], including complex higher-order transformations of complex variables [68,69]. However, encoding approaches offer only limited and indirect methods for assessing the dynamic organization of neural populations that, according to cell assembly concepts, support cognitive function by coordinated changes across many cells within a single trial.

Decoding approaches
Decoding compares a sample of neural activity to an established tuning curve in an attempt to predict the value of the encoded behavioral, environmental or cognitive parameter [53,64,70,71]. Decoded results that deviate from the observation are interpreted as errors [53,70–72]. Average errors are used to distinguish between multiple decoding algorithms [70–72], but individual decoding errors are typically thought to indicate noise or the inadequacy of the neural code used for decoding.

Analysis of covert, cognitive processes present a distinct problem for the decoding approach to neural codes because the cognitive variables implied by these processes are not typically experimentally observable. In decoding approaches, these cognitive processes will probably appear as decoding errors. To avoid the problem of calling these deviations ‘errors’, decoding approaches to covert, cognitive processes have either highlighted unexpected structure within the distribution of decoding errors or compared the time series of predicted (decoded) cognitive variables with a hypothetical time series that is derived from subjective expectations about the cognitive processing.

The application of the decoding approach to cognitive processing has been successful in several studies on population vector rotation in motor cortex [2,4] and route replay in hippocampus [49,53,54,57–59]. However, the decoding approach to cognitive processing provides a weak form of statistical testing because it depends entirely on how the hypothesized cognitive time series has been defined. As a result, decoding approaches to covert cognitive processes are usually reducible to an extended cross correlation analysis that is informed by the tuning curve of each cell, which requires the potentially dubious assumption that the covert cognitive variable is constant across many trials [2,4,48,49,51,55,73].

Generative approaches
The studies of within session multi-stability and extra-field spiking used what we call the generative approach. The generative approach exploits the fact that a tuning curve is not merely descriptive; it can also be used as a basis for generating a prediction of neural activity from a given set of behavioral, environmental or cognitive parameters [39,41,42,59,64]. The generative approach enables direct comparisons between the predictions that a spike should occur and actual observations. The generative approach can be used in flexible tasks that are more likely to involve cognition. As a result, the generative approach provides a framework for detecting and quantifying the multi-neuron cell assembly dynamics that Hebb [1] originally proposed.

Support for the validity of a proposed neural code within the generative approach is determined by how well an...
observed set of neural activity can be predicted from behav-
ioral, environmental or cognitive variables [41,42,59,64]. At a super-
ficial level, the application of the generative approach to covert, cognitive processes faces the same prob-
lem encountered within decoding approaches (that a hypo-
thetical time series of a cognitive variable must be pro-
posed for evaluating validity). However, several appli-
cations of the generative approach have circumvented this

Box 1. Analyzing neural ensembles

Current technology enables the simultaneous recording of large neural ensembles from awake, behaving animals. The simultaneity of the large neural ensembles enables the identification of both decoded behavioral variables (reconstruction) and the prediction of spike trains (generative models) from hypothesized processes (Figure I).

Encoding
Tuning curves are generated from the correlations of behavioral variables and simultaneously observed spike trains. Behavioral variables can include sensory inputs, motor outputs or task-related behavioral variables (as shown here).

Decoding
By definition, a tuning curve encodes a description of an observed variable, such as a behavior, as a function of spike trains. Through standard methods, it is possible to invert this description to predict the behavioral variable from an observation of spike trains.

Generative models
Because the tuning curve is a description of the spike train as a function of the behavioral variable, it is possible to predict the spike trains from the observed behavior and/or the value of a hypothesized or decoded covert variable.

Comparisons
The decoded behavioral variable can be compared with the actual behavioral variable. These differences (and similarities) can be examined for structure, which provides evidence for cognitive processes. Similarly, the predicted spike trains can be compared with the actual spike trains for differences and similarities, which can be examined for underlying structure and provides evidence for cognitive processes.

Figure I. Example of the encoding-decoding-generative models cycle. Encoding – tuning curves are generated from behavior variables (here, the location of the rat during foraging within a 1m cylinder) and from ensemble spike trains (here, showing diagrammatic action potentials from nine of 93 cells; spikes from each cell are indicated by a different color). Decoding – combining tuning curves and observed spike trains produces decoded variables. This example shows a Bayesian probability distribution of the location of the rat given the tuning curves and spike trains at a given time. The white star indicates the observed location of the animal. Generative models – combining decoded behavioral variables with tuning curves produces predicted spike trains.
problem by inferring the time series of a covert variable’s value on the basis of neural ensemble activity [41,42,59]. Instead of relying on assumed psychological processes, these variants of the generative approach either implicitly [42] or explicitly [41,59] used decoding algorithms to infer the time series of the covert variable. These studies showed that apparent noise at the single unit level was actually self-consistent within the set of simultaneously recorded neurons. The ensemble activity was coordinated even when it deviated from what the experimenter expected given the position of the animal in the maze and presumed spatial map; these experimentally unexpected deviations were consistent with switching between spatial maps [41], route replay [59] and other non-local spatial representations [60]. The success of the generative approach in identification of covert processes illustrates the power of basing validity tests of the neural code on a biological signal (ensemble activity) instead of relying on more indirect expectations that arbitrarily delimit cognition.

Conclusion
Cognitive processes fundamentally and often subtly change how information is represented in the brain. As a result, understanding the neural bases of cognition requires one to distinguish between noise and covert, internally generated changes in neuronal firing patterns. The generative approach provides a distinct advantage over approaches that stop at encoding or decoding because generative approaches can quantitatively differentiate between random noise and structured variability that is by definition not noise, but cannot be accommodated by the tuning curve of the neuron. As techniques for recording large neural ensembles become a more standard tool for investigating the neural basis of cognition across the brain, the theor- etical problem of how to analyze neural data under conditions in which constancy of the cognitive variable cannot be safely assumed becomes crucially important. Analysis of this ‘excess’ variability provides a point of departure for understanding how attention, memory consolidation and decision-making contribute to, and refine, traditional ideas of spatial representations present in hippocampal activity. Further technological advances in the quality and density of recording neural ensembles will offer unprecedented opportunities to understand the dynamics of neural population activity, with the promise of defining many of the largely covert, subjective processes we call cognition in terms of objective biological mechanisms.

Acknowledgements
We thank Matthijs van der Meer, Anoopum Gupta, Jadin Jackson, David Crowe, Matt Chafee, Dave Rowland, Aldis Weible and the Cognitive Sciences Center at the University of Minnesota for helpful discussions. This work was supported by NIH grant MH080318 (A.D.R.), NSF grant IOS0725001 (A.A.F.), NIH training grant (T32 HD007151) (A.J.) and a Dissertation Fellowship from the Graduate School at the University of Minnesota (A.J.).

References
1 Hebb, D.O. (1949) The Organization of Behavior, a Neuropsychological Theory. Wiley
8 Redish, A.D. (1999) Beyond the Cognitive Map: from Place Cells to Episodic Memory. MIT Press
10 McNaughton, B.L. et al. (1983) The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats. Exp. Brain Res. 52, 41–49
22 Kentros, C.G. et al. (2004) Increased attention to spatial context increases both place field stability and spatial memory. Neuron 42, 283–295
30 Wills, T.J. et al. (2005) Attractor dynamics in the hippocampal representation of the local environment. Science 308, 873–876
31 Markus, E.J. et al. (1995) Interactions between location and task affect the spatial and directional firing of hippocampal neurons. J. Neurosci. 15, 7079–7094
45 Pavlides, C. and Winson, J. (1987) Influences of hippocampal place cell firing on the awake state and the activity of these cells during subsequent sleep episodes. J. Neurosci. 9, 2907–2918
64 Rieke, F. et al. (1997) Spikes. MIT Press
71 Brown, E.N. et al. (1998) A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. J. Neurosci. 18, 7411–7425