# From Wave Dynamics to Continuation Dynamics: Why Neural Oscillations May Be Realizations Rather Than Explanations

Flyxion
*Independent Researcher*

July 2026

**Abstract**

Miller, Brincat, and Roy's *Analog Cognition and Consciousness* argues that oscillatory brain waves are not merely correlated with flexible cognitive control but are computationally necessary for it, since synaptic and spiking mechanisms alone are too slow and too demanding to explain mixed selectivity and context-dependent routing. This essay shows that the necessity claim does not hold: recurrent architectures with no wave-based mechanism already reproduce the phenomena used to motivate it. That result does not settle the question in favor of any alternative; it reopens it. The essay then argues, without overreaching, that an existing body of manifold theory already explains a substantial class of neural stability phenomena on its own terms, before identifying a narrower class – representational drift with preserved behavior – that existing accounts only partially explain, by conflating two distinct properties: what can be decoded from a neural population state, and what remains reachable from it. A minimal formal object, the admissible continuation set $\Gamma_C(x)$, is introduced to state this distinction as a falsifiable empirical claim rather than a terminological one, and a discriminating experiment is specified that could resolve it in either direction. The essay's conclusion is deliberately layered rather than adversarial: manifold structure constrains the repertoire of available dynamics, admissible continuation structure characterizes which futures within that repertoire remain reachable, and oscillatory processes may participate in selecting among those futures without thereby defining the space of what is reachable. Two mathematical appendices situate the proposal relative to viability theory and demonstrate, via a minimal worked example, that decodability-preservation and reachability-preservation are not logically equivalent claims. The essay's central contribution is accordingly not the formalism itself but the identification of reachability preservation as a distinct, previously untested hypothesis about what is conserved when neural behavior survives representational change.

# I. Opening

Miller, Brincat, and Roy's *Analog Cognition and Consciousness* makes a specific argument, not a general one, and it is worth stating precisely before responding to it. Mixed selectivity — the finding that most cortical neurons participate in many tasks rather than one — creates a genuine control problem: something must determine which neurons contribute to which computation, moment to moment, without a system having to track individual synapses. The authors argue that synaptic mechanisms are too slow and too computationally demanding to solve this problem, and that spatially patterned oscillations — brain waves, acting as flexible, low-dimensional control signals — are the tractable alternative. On this view, waves are not merely correlated with cognition; they are a large part of what makes flexible cognition possible at all, and their organization across cortex is what consciousness, on their account, consists in.

This essay does not argue that waves are unimportant, that oscillations are epiphenomenal, or that Spatial Computing is wrong about what cortical rhythms do. It argues something narrower and, for that reason, more defensible: that the specific phenomena Miller et al. use to motivate a privileged role for waves — mixed selectivity, context-dependent routing, low-dimensional population dynamics — do not uniquely implicate a wave-based mechanism, because systems lacking any such mechanism already produce them. That observation does not settle the question in favor of any alternative. It reopens it, at a level of abstraction the original argument did not anticipate needing.

What follows is organized as a sequence of claims, each stated no more strongly than the evidence supports, several of which concede real ground to the position under examination. Section III shows that the necessity claim underlying Spatial Computing does not hold, using evidence the original paper itself half-cites without noticing the tension. Section IV then argues the opposite direction: that an existing, well-supported body of manifold theory already explains a substantial class of neural stability phenomena, to the point that any further theoretical apparatus is, at that point in the argument, unnecessary. Section V identifies the specific class of phenomena that theory does not explain — not by asserting a replacement, but by naming a distinction, between what a population currently encodes and what remains reachable from it, that no existing account tests. Sections VI and VII introduce the minimal formal object needed to state that distinction as a falsifiable prediction, and specify the experiment that would decide it, in either direction. Throughout, a further distinction will be drawn — one that becomes explicit only in Section VIII, but that is worth naming here so it can be tracked from the start: between evidence that waves help define the space of possible neural trajectories, and evidence that waves determine which trajectory is selected on a given occasion. Much of what follows turns on keeping these apart. Section VIII applies this distinction, together with the discipline built up in the sections before it, to the paper's anesthesia and consciousness evidence, and Section IX extends it to the wider field of consciousness theories the paper positions itself among.

The result is not a defense of one theory against another. It is an argument that the two theories currently on offer — wave-based control and something like continuation-based control — have not yet been distinguished by any evidence either side has produced, together with a specific, falsifiable account of what would distinguish them. Readers looking for a confirmation of

continuation geometry will not find one here. What they will find is a narrower and, I think, more useful result: a demonstration of exactly where the current evidence runs out, and what experiment would need to be run to take the argument further than assertion currently allows.

## II. What the Paper Gets Right

Before turning to the necessity claim in Section III, it is worth stating plainly what is not in dispute. Miller et al.'s critique of strict connectionism is correct and, at this point, largely uncontroversial within systems neuroscience. The classical picture of cortex as a hierarchy of stable, specialized feature detectors — each neuron's spiking carrying one fixed meaning — cannot account for a finding that has been replicated across species, tasks, and cortical regions for over two decades: that a given task typically engages only a fraction of the neurons available to it, and that the same neurons participate in different computations under different conditions. This is not a marginal correction to the connectionist picture. If neurons were genuinely specialized in the way early models assumed, a cortical population would exhaust its representational capacity after learning only a handful of tasks, which is not what is observed.

Mixed selectivity is the right name for this phenomenon, and the control problem it creates is real. Something has to determine which neurons contribute to which computation, in which context, without the system needing to track individual synaptic assignments — Miller et al. are correct that this is a nontrivial coordination problem, and correct that any adequate theory of cortical function has to say something about how it is solved. Nothing in this essay disputes that mixed selectivity is pervasive, that it poses a genuine control problem, or that an explanation is owed.

What is in dispute is narrower and more specific: not whether a control problem exists, but what kind of mechanism is required to solve it. That is the question Section III takes up directly.

## III. The Necessity Claim, and Where It Breaks

Miller, Brincat, and Roy's argument for a privileged explanatory role for brain waves does not rest primarily on correlational evidence. It rests on a claim about computational necessity, stated explicitly in their introduction: synaptic mechanisms are "likely too cumbersome" for flexible top-down control, because such control would require a system to "know" which synapses and neurons were employed for any given representation and to coordinate them with subsecond precision — "a seemingly implausible computational demand." Everything that follows — Spatial Computing, the alpha/beta stencil mechanism, ephaptic routing, analog wave computation — is introduced as the solution to this problem. The argument has a simple structure:

1. Mixed selectivity creates a control problem: the same neurons must be flexibly recruited into different task-relevant populations depending on context. 2. Synaptic reconfiguration is

too slow and too computationally demanding to solve this problem on subsecond timescales. 3. Spatially patterned oscillations provide a tractable alternative — a control signal that can gate which neurons express information without requiring the system to track individual synapses. 4. Therefore, wave dynamics occupy a privileged, and in this specific sense necessary, explanatory position in accounts of flexible cognition.

The premise doing the real work is (2). If synaptic and recurrent dynamics genuinely cannot produce fast, flexible, context-dependent routing without an auxiliary control layer, then something like Spatial Computing becomes a reasonable inference to the best available mechanism. If they can, the entire argument's motivating urgency dissolves — not because waves are shown to be absent or unimportant, but because the specific problem they were introduced to solve turns out to already have a solution that doesn't require them.

This premise is not simply undemonstrated. It has already been contradicted by a result the authors do not engage on this point. Mante, Sussillo, Shenoy, and Newsome (2013) recorded prefrontal activity in monkeys performing a context-dependent sensorimotor task and found that the same population of neurons could integrate and select between sensory inputs depending on task context, with no anatomical rewiring and no oscillatory gating mechanism of any kind. They then reproduced this behavior in a trained recurrent neural network — a system whose only computational resources are fixed connectivity and unfolding population dynamics — and showed that selection and integration emerge as two aspects of a single dynamical process within that fixed architecture. This is not a weaker or partial version of the phenomenon Miller et al. are trying to explain. It is the same explanandum: context-dependent, flexible, subsecond routing of information through a population of multifunctional units, without the system needing to "know" which units are currently doing what. A fixed recurrent architecture already does this. The "implausible computational demand" the wave mechanism was introduced to avoid is not, in fact, implausible; it has been demonstrated.

There is a further irony worth making explicit, because it is not an omission on the authors' part so much as an unnoticed seam in their own citations. When Miller et al. turn to subspace coding — the observation that population activity organizes into low-dimensional, partially orthogonal subspaces — they cite Kaufman et al.'s work on null-space coding and the broader population-dynamics literature (including Vyas et al.'s review of computation through neural population dynamics) as consistent with, and supportive of, their own account. But this is precisely the literature in which such organization is shown to arise from connectivity and recurrent dynamics alone, without appeal to any field-based control signal. The authors have absorbed a body of evidence into their framework that does not, on inspection, distinguish their mechanism from the alternative. It supports the existence of low-dimensional, context-structured population coding. It does not support waves as the reason that coding exists.

None of this licenses the conclusion that oscillations are unimportant, or that the Spatial Computing framework is wrong about what waves in fact do in cortex. What it licenses is a specific, narrower claim, and the essay's argument depends on stating it precisely rather than overstating it. The available evidence on brain waves and cognitive control sorts into a hierarchy of decreasing support:

- **Waves correlate with, and likely help organize, ongoing neural computation.** This is well-supported

by a large body of evidence the authors correctly marshal. - *Waves may make biological control more efficient, spatially scalable, or energetically economical than synaptic recurrence alone.* This remains plausible and is empirically open; nothing in the RNN literature bears directly on efficiency, only on capability. - *Wave interactions are the unique computational primitive that explains mixed selectivity, subspace coding, and context-dependent routing — phenomena unavailable to non-wave systems.* This is the claim the paper's introduction stakes out, and it is the claim the existing recurrent-network literature directly contradicts.

The significance of this is not that oscillatory theories are false. It is that the explanatory burden has moved. Once context-dependent selection, integration, and flexible reuse are shown to emerge in systems lacking a mesoscale ephaptic wave substrate, the question is no longer why waves *can* support these functions — they plausibly can, and may do so efficiently. The question becomes what organizational invariant is shared by every system, wave-based or not, that supports them. The remainder of this essay argues that the relevant invariant is not oscillation itself, and not low-dimensional population geometry as conventionally measured, but the preservation and selection of admissible continuations — a property that manifold theory partially captures and partially misses, and that the following two sections examine in turn.

## IV. What Manifold Theory Already Owns

The previous section drew on two literatures often grouped together under the heading of population dynamics, and the argument only works if they are kept apart. Task-relevant subspace and null-space analyses ask how information is organized within an existing neural population at a given moment — which directions in activity space carry behaviorally relevant signal, which are orthogonal to it, how interference between simultaneous computations is avoided. Intrinsic-manifold and repertoire theories ask a different question: what aspects of that organization remain stable as the system learns, adapts, or is perturbed. The first literature is what did the work in Section III — it shows that context-dependent routing does not require a wave-based control layer. The second literature is what this section is about, and it places a real constraint on anything proposed to replace or extend that account.

The constraint is this: neuroscience already has a working theory of what is preserved during adaptation, and it is not vague. Sadtler et al. (2014) trained monkeys to control a brain-computer interface and then perturbed the mapping between neural activity and cursor movement in two ways — within the animal's existing intrinsic manifold, or outside it. Within-manifold perturbations were learned quickly. Outside-manifold perturbations, requiring activity patterns the population had never generated, were not learned at all over the timescale of the experiment, even though nothing about the task itself was harder. The intrinsic manifold — a covariance structure reflecting the network's existing connectivity — was not incidental to learning. It was the thing that determined whether learning was possible.

Golub et al. (2018) sharpened this further. Studying short-term BCI relearning, they found that animals did not generate new activity patterns to solve a newly perturbed task. They

reassociated an existing, fixed repertoire of patterns with new movements. The population's internal computational dynamics were largely unaltered by learning; what changed was which of the population's *pre-existing* patterns got mapped to which output. This is not a weak or metaphorical sense of "preserved structure." It is a specific, quantitative claim — the repertoire is fixed, and adaptation operates by reassignment within it, not by expansion of it.

Put plainly, and put in a way that should be uncomfortable for the argument this essay is building toward: if manifold theory already explains adaptation, recovery, and short-term learning as well as this, then a continuation framework proposed as a *replacement* for it is unnecessary. Sadtler and Golub between them already answer, with real data, the question "what remains invariant when a neural system is asked to do something new?" The answer, as far as these experiments show, is: the manifold. Not some more abstract object standing behind the manifold. The manifold itself, understood as a covariance structure imposed by fixed synaptic connectivity, does the explanatory work.

This is worth sitting with rather than moving past. A reader inclined to find the continuation framework attractive should feel the pull of this section — because if it were the whole story, there would be nothing left for Sections VI and VII to contribute beyond a change of vocabulary. Manifold preservation already predicts, quantitatively, when learning succeeds and when it fails. It already explains why some adaptations are fast and others are essentially impossible on a short timescale. Any framework introduced after this section has to clear a specific bar: it must identify something that happens in real neural systems that manifold preservation, in the Sadtler/Golub sense, does not predict.

That is now a concrete empirical requirement rather than a rhetorical one. The next section takes it up directly, by turning from short-term perturbation — where the manifold is stable and behavior tracks it — to long-term representational drift, where the manifold itself is not stable, and yet, in at least some of the relevant brain regions, behavior is.

## V. Where Manifold Theory Is Only Half an Answer

Section IV established that manifold preservation, in the Sadtler/Golub sense, successfully predicts a specific and important class of stability phenomena: when a neural system is perturbed on a short timescale, whether it can adapt depends on whether the required activity pattern already lies within its existing covariance structure. That result is not in question here. What is in question is whether the same object — the manifold, understood as a stable covariance structure over a population — is the right conserved quantity across a wider range of cases, including ones where the manifold itself does not stay put.

Driscoll, Pettit, Minderer, Chettih, and Harvey (2017) provide exactly such a case. Recording chronically from posterior parietal cortex in mice performing a repeated virtual navigation task, they found that individual neurons' activity–behavior relationships changed continually over days and weeks — cells gained and lost task correlations, altered their selectivity, entered and exited the population representation that carried the task. Over the course of several weeks, the task-related activity in PPC had, in their description, nearly entirely reconfigured. And yet

the animals' behavior and task performance did not measurably change. The representation drifted. The manifold, in any sense tied to which neurons carry which information with what covariance structure, did not stay fixed. Behavior did.

This is not a minor exception to the Sadtler/Golub picture; it is a different regime entirely. Sadtler and Golub describe systems where the manifold is stable and behavior tracks it. Driscoll describes a system where the manifold is not stable and behavior still does not change. If the preserved object were simply "the representation" — the specific pattern of population covariance the earlier sections treat as doing the explanatory work — this case should not be possible. Something is being conserved in PPC across weeks of near-total representational reorganization. It is not, on the available evidence, the representation itself.

The most direct existing response to this problem does not, on inspection, come from within the intrinsic-manifold/repertoire literature that did the work in Section IV. It comes from the other literature Section IV set aside at the outset — task-relevant subspace and null-space analysis — redeployed here for a different purpose. Rule et al.'s reanalysis of the same phenomenon, titled "stable task information from an unstable neural population," proposes that drift is not unconstrained: it occurs preferentially in directions of population activity orthogonal to the coding dimensions that carry task-relevant information — a null coding space, in the sense introduced by Kaufman et al. and discussed in Section III. This is worth flagging explicitly, because it means the rescue changes the kind of object being invoked. The repertoire in the Sadtler/Golub sense — a fixed covariance structure reflecting connectivity, tested by whether a required activity pattern lies within it or outside it — is not what Rule et al. show to be preserved during drift; on their own account it is not preserved. What is preserved, on their account, is something narrower: a decodable readout, extracted via the subspace methods of Section III, that survives even though the covariance structure it is read from does not.

On this view, then, the population's intrinsic manifold is not exactly stable, but a *readout* of it is: a decodable task variable remains recoverable from population activity at every point in time, even as the raw geometry drifts around it.

This is a genuine partial rescue of the broader population-dynamics account, and it should be treated as one rather than dismissed, even though it succeeds by changing which object is claimed to be conserved rather than by defending the original one. It explains why decoding a task variable continues to work despite drift. But it answers a narrower question than the one the phenomenon actually raises. Stable decodability is a claim about the present moment: given the current population state, can a fixed or slowly-updating readout recover the task-relevant variable right now. It says nothing about whether the *set of futures reachable from that state* — what the system could still do, where it could still go, how it would respond to a given perturbation — is preserved across the same drift. A population can support stable decoding of "which direction to turn" at every timepoint while nonetheless differing, from one week to the next, in what would happen if the trial were interrupted, if the cue were ambiguous, or if the animal needed to recover from an error mid-trial. Decodability and reachability are not obviously the same property, and no result in this literature distinguishes them, because no one has yet asked whether they come apart.

So the conclusion this section is entitled to is narrower than it might appear, and narrower

than the continuation framework will eventually want it to be. It is not: behavior remains stable while representations drift, therefore continuations are the invariant. That inference outruns the evidence. The conclusion is only this: manifold preservation, in its strongest available form, predicts one class of stability phenomena very well — the short-timescale, perturbation-driven case examined in Section IV — and only partially predicts another — the long-timescale, drift-driven case examined here, where the partial rescue trades a claim about structure for a weaker claim about decodability, and where that substitution has not yet been tested against the alternative it quietly assumes away.

If behavior can remain stable while both individual neurons and the population-level representation itself reorganize, what exactly is being preserved?

## VI. Admissibility and Reachability

The previous section ended with a distinction that the existing literature does not currently make. A neural population may preserve the ability to decode a task variable while changing the set of trajectories that remain available to it. Decodability concerns what can be inferred from a state. Reachability concerns what can still be done from that state. These are not obviously the same property, and nothing in the drift or manifold literatures examined so far has tested whether they come apart.

To state this distinction precisely enough to test it, some minimal formal vocabulary is needed — not as an alternative theory of neural computation, but as the smallest object capable of expressing the difference between "this variable is currently decodable" and "this range of futures is currently reachable." Let a task system have a state space X, a set of task constraints C, and a transition law F governing how population activity evolves. Define the admissible continuation set from a state x as the collection of trajectories starting at x that obey F and continue to satisfy C:

$$\Gamma_C(x) = \{\, \gamma : [0, T] \to X \mid \gamma(0) = x,\ \gamma \text{ follows } F,\ \gamma(t) \text{ satisfies } C \text{ for all } t \,\}$$

This is a description of what remains possible from x under the task's constraints, not a description of what x currently encodes. That difference is the entire content of the proposal, and it is what lets a specific empirical claim be stated — a claim that, not the definition itself, is what the rest of this section exists to earn:

If two population states preserve the same decodable variables but differ substantially in their admissible continuation sets, the continuation account predicts different recovery behavior, different adaptation behavior, and different perturbation responses despite equivalent decoding performance.

This is a discriminating prediction in the strict sense used throughout this essay: it is a claim the null-space/decodability account, as currently formulated, does not make and has no resources

to make, since that account is defined entirely in terms of what is recoverable from a state at a single moment, not in terms of what remains reachable from it over time.

Two precisions are needed before this claim is usable rather than merely stated, and it is worth taking them in turn rather than folding them together.

The first concerns C itself. If task admissibility is fixed only by whether a trial ultimately succeeds, then $\Gamma_C(x)$ becomes, by construction, "the set of trajectories that lead to success," and the claim above collapses into the unfalsifiable observation that recovery tracks capacity for success. C must therefore be built from constraints checkable before a trial resolves, not from its outcome. Three kinds of constraint do this across the datasets already discussed: task checkpoints fixed by design — in a BCI paradigm, the decoder mapping from neural activity to effector output is known in advance, so whether a trajectory produces decoder-consistent progress toward a target is verifiable at every timestep, independent of whether the trial is eventually correct; timing windows — a trajectory is admissible only if it reaches a decision-consistent state within the task's known deadline; and epoch-defined representational requirements — a trajectory through a working-memory delay period is admissible only if the relevant variable remains decodable throughout that epoch, a condition checkable epoch by epoch rather than only at trial's end. Defined this way, C is fixed by task structure the experimenter already controls, and whether a given $\gamma \in \Gamma_C(x)$ in fact leads to correct behavior becomes a testable question rather than a defining one.

The second concerns what kind of object $\Gamma_C(x)$ is claimed to be, and here the caution runs the other direction. The claim is not that admissible continuation sets are substrate-independent, floating free of whatever neural hardware realizes them. Different implementations may well realize different continuation structures, and nothing here assumes otherwise. The claim is narrower: if the scientific question at hand concerns recovery, adaptation, and perturbation response, a description in terms of reachable futures may be more predictive than a description in terms of instantaneous encoding, for a given system, at a given time — not that reachability is somehow prior to or independent of the substrate producing it. This is the same caution manifold theory itself observes; a covariance structure is also substrate-dependent, tied to a specific connectivity, and nothing in Sections IV or V treated it otherwise. $\Gamma_C(x)$ is not exempt from that dependence. It is simply a different quantity to compute from the same underlying system.

With both precisions in place, it is worth being explicit about what distinguishes $\Gamma_C(x)$ from several existing constructs it might otherwise be mistaken for, since the proposal is only interesting if it is not simply a relabeling of tools already available. An attractor basin describes the set of states that converge to a particular fixed point or limit cycle under unperturbed dynamics; it is defined by convergence, not by task success, and says nothing about which of the trajectories passing through it satisfy C. A controllability region, in the control-theory sense, describes which states can be driven to a target using available inputs; it is a property of the control system, not of the population dynamics evolving under a fixed policy, and does not incorporate task constraints as a first-class object. A viability kernel is closer in spirit — it describes the largest set of states from which some trajectory can be kept within a constraint region indefinitely — but viability theory is typically applied to the state space of a controlled system with an explicit control input, whereas $\Gamma_C(x)$ is defined directly on population activity

without presupposing which variables, if any, are being externally controlled. Reachable sets in control theory ask which states can be reached from x, full stop; $\Gamma_C(x)$ asks which trajectories from x remain admissible under C, which is a constraint on the path, not merely the endpoint, and is what distinguishes recovery that passes through task-appropriate intermediate states from recovery that merely arrives at a correct final state by an inadmissible route. A successor representation, in the reinforcement-learning sense, is closest of all in spirit — it also encodes expected future occupancy from a state — but it is typically a discounted, scalar-valued expectation over future state visitation under a fixed policy, collapsed into a single predictive representation, whereas $\Gamma_C(x)$ is the set of admissible trajectories itself, not a summary statistic over it, and carries no discounting or policy assumption. None of these existing constructs is wrong for the purpose it was built for. The claim here is only that none of them, individually, is the object Section V's distinction requires, and that $\Gamma_C(x)$ is closest to a viability kernel stripped of its control-theoretic apparatus and stated directly in terms of task admissibility over population trajectories. Readers who find this unconvincing should treat $\Gamma_C(x)$ as notation for "whichever of these existing constructs turns out to be the right one once the discriminating experiment in Section VII is run" — the essay's empirical claim does not depend on $\Gamma_C(x)$ being conceptually novel, only on reachability being a different predictor than decodability or covariance geometry.

Two states can now be compared directly. Define x and y as continuation-equivalent, written $x \sim_C y$, when their admissible continuation sets are sufficiently similar — when the range of task-successful futures available from one is approximately the range available from the other, regardless of whether their instantaneous population geometry or decodable readout is similar or not. This relation is deliberately weaker than geometric equivalence and deliberately independent of decodability. Two states may be geometrically close, and equally decodable, yet fail to be continuation-equivalent if a perturbation from one recovers easily and a perturbation from the other does not. Two states may be geometrically distant, drawn from what a covariance analysis would treat as different manifolds entirely, and still be continuation-equivalent if their downstream recovery behavior is indistinguishable.

Two things should be said about what this section is not claiming, because both are easy to claim by accident and neither is earned by anything above. This is not a claim that admissible continuations are more fundamental than manifold structure, or that they describe some deeper reality underlying decodable representations. Sections IV and V gave manifold theory and decodability accounts their full due; nothing here revokes that. Nor is this a claim that $\Gamma_C(x)$ has already been shown, empirically, to be the conserved quantity in any of the drift or perturbation datasets discussed earlier. It has not. The datasets that would test it — Driscoll's drift recordings, the Sadtler and Golub perturbation datasets — have not yet been analyzed with this comparison in mind.

What can be said is narrower and, for that reason, more defensible: if the open question left by Section V is reachability rather than decodability, then answering it requires an object that describes reachability rather than decodability. $\Gamma_C(x)$ is proposed as such an object, on the grounds that it is the minimal construct capable of stating the distinction, not on the grounds that it has already resolved it. The claim that gives this section whatever force it has is empirical and falsifiable: recovery after perturbation should track preservation of admissible continuation sets more closely than it tracks preservation of covariance geometry, phase relationships, or

instantaneous decoding performance. If that turns out to be false — if recovery tracks manifold geometry or decodability just as well, or better, once the comparison is actually run — the proposal in this section fails on its own terms.

The proposal, then, is not that admissible continuations have already been shown to be the conserved quantity. The proposal is that they define a conserved quantity whose predictions differ from those of existing manifold and decoding accounts, and that this difference can be tested against them. What that test would need to look like is the subject of the next section.

## VII. What Would Distinguish the Accounts

Section VI's prediction is only worth making if it can be tested against data that already exists, and only worth taking seriously as a research program if it also specifies what new data would settle the question the existing data cannot. Both are addressed here, in that order — the reanalysis first, because it is cheap and could in principle already refute the proposal, and the new experiment second, because it is the test that would actually be decisive.

**The reanalysis test.** Two existing datasets are suited to testing whether recovery tracks continuation preservation better than it tracks manifold geometry or decodability, and neither has been analyzed this way.

The Sadtler and Golub perturbation datasets record, trial by trial, which BCI mappings a given population of neurons could and could not learn, and how quickly. The existing analysis explains this in terms of distance from the intrinsic manifold: within-manifold perturbations succeed, outside-manifold perturbations do not. The same data would support a second analysis, orthogonal to the first: for each perturbation, estimate the admissible continuation set available from the pre-perturbation state — the range of task-successful trajectories the population could still execute — and ask whether learning success is better predicted by continuation-set overlap between the required and the available repertoire than by manifold membership as conventionally defined. In most trials these two predictors will likely agree, since within-manifold perturbations are also, plausibly, continuation-preserving ones. The interesting trials are the disagreements: cases where a perturbation is technically within the manifold but requires a continuation the population cannot in practice execute, or vice versa. If disagreements are rare or nonexistent, the reanalysis will have shown that $\Gamma_C(x)$ reduces to manifold membership in this dataset, and the proposal will have failed its first test cheaply, which is the right way for it to fail if it is going to fail.

Driscoll's drift dataset is the more direct test, because it is the case Section V identified as the open one. For each recorded session, the same reanalysis question can be asked in reverse: as the population representation reorganizes across weeks, does the animal's ability to recover from mid-trial perturbations or task-irrelevant distraction track preservation of decodable task variables, preservation of population covariance structure, or preservation of estimated continuation sets? Rule et al.'s null-space account predicts decodability should track recovery. The continuation account predicts something that could diverge from decodability specifically on trials where the decodable variable is intact but the surrounding repertoire of nearby, task-

relevant trajectories has narrowed or shifted — a pattern decodability analysis is not built to detect because it only asks about the moment of readout, not about what happens next.

**The new experiment.** The reanalysis can weaken or strengthen the proposal but cannot fully decide it, because $\Gamma_C(x)$ as defined in Section VI has never been directly estimated from data collected for this purpose; every dataset above was collected to test something else. The decisive experiment holds population trajectory geometry approximately fixed while independently manipulating oscillatory phase organization — the manipulation Miller et al.'s own framework treats as functionally significant, since Spatial Computing assigns explanatory weight specifically to alpha/beta phase patterns and cross-frequency relationships, and Chen et al. (2026) report that phase-pattern mismatch predicts behavioral errors even when, presumably, the underlying population geometry is not obviously disrupted.

The feasibility of perfectly separating trajectory geometry from phase organization remains an open technical challenge; no current method achieves this cleanly, and closed-loop or optogenetic approaches to phase manipulation may perturb trajectory geometry as a side effect rather than leaving it untouched. The experiment should therefore be understood as defining the discriminating evidence required in principle — the standard any future methodology would need to meet to adjudicate the question — rather than as a protocol presupposing techniques that already exist. This does not weaken the claim that such a separation is what the question requires; it only acknowledges that meeting that requirement is itself an open problem, and possibly a productive one for methods development independent of the theoretical debate.

This design produces outcomes that discriminate between the accounts under consideration, but the discrimination is sharper than a first pass suggests, and stating it precisely matters. The relevant question is not simply whether phase manipulation changes routing or behavior — it should be expected to, on almost any account, since phase is not causally inert. The question is what kind of change it produces.

If phase manipulation alters *which* admissible continuation is realized — the population still follows some trajectory in $\Gamma_C(x)$, task-relevant recovery remains possible along the paths $\Gamma_C(x)$ still contains, but a different one is taken than would have been taken otherwise — this is not evidence against the continuation account. It is compatible with, and in fact predicted by, the two-tier structure argued for in Section VIII: continuations set the envelope of what remains reachable, and waves are one plausible mechanism for selecting a path within that envelope. Under this outcome, waves remain the selector, and the essay's account of what they select from is unchanged. This case should not be mistaken for disconfirmation merely because behavior changes.

The account is genuinely threatened only by a different result: if phase manipulation changes $\Gamma_C(x)$ itself — narrows or shifts the set of task-successful trajectories still reachable from a state, such that recovery paths available before the manipulation are no longer available after it, holding population trajectory geometry and task constraints C fixed — then phase is not merely selecting among continuations but altering the continuation structure directly, in a way the framework as stated treats as a property of the state and its dynamics, not of an independent physical coordinate layered on top. This is the result that would show the continuation description has omitted a causally relevant variable, since $\Gamma_C(x)$ as defined in

Section VI is a function of x, F, and C alone, with no phase term.

Conversely, if systems with different phase arrangements — including, in the limit, a recurrent network with no oscillatory or ephaptic mechanism at all — preserve the same $\Gamma_C(x)$, the same perturbation-response structure, and the same recovery pathways as the biological system, that is direct support for the claim that continuation structure is what is conserved and phase is one contingent implementation of it among others, whether or not it also plays a selection role within that structure.

Chen et al.'s finding that alpha/beta mismatch predicts behavioral errors sits, at present, ambiguously among these possibilities, and this essay has no basis for resolving it in advance. It is compatible with phase carrying genuinely irreducible information that alters $\Gamma_C(x)$ itself when disrupted, in which case the threatening outcome above is closer to correct. It is also compatible with phase mismatch being a reliable *correlate* of continuation-set disruption without being the causal variable — the population's reachable-future structure could be disrupted by whatever also disrupts phase coordination, without phase itself being the operative quantity. And it is compatible with a third possibility this section's revised framing now makes visible: that phase mismatch predicts errors because it governs *which* admissible continuation is selected, and errors arise when the wrong one is chosen from an otherwise-intact set, in which case Chen et al.'s result is fully consistent with the continuation account and simply documents one way waves perform selection. Only the manipulation described above, holding trajectory geometry and $\Gamma_C(x)$ fixed while varying phase independently, distinguishes these three possibilities. Correlational evidence, however well-replicated, cannot.

This is the sense in which the essay's argument, taken as a whole, makes a real bet rather than relocating a mystery from one vocabulary to another. Section III showed that waves are not necessary for the phenomena that originally motivated them. Section IV showed that manifold theory already explains a serious class of stability phenomena on its own terms. Section V showed that a further class of phenomena leaves open a question manifold theory has not yet answered. Section VI proposed a minimal object adequate to state that question as a testable prediction. This section specifies the experiment whose outcome the essay is prepared to be wrong about.

## VIII. Two Kinds of "Wave"

The anesthesia argument in Miller et al.'s paper is the essay's last remaining piece of unexamined evidence, and it deserves the same treatment given to mixed selectivity and subspace coding: not dismissal, but a precise account of what it does and does not establish.

The paper's version of the argument is qualitative. Different anesthetic agents, acting through different molecular targets, converge on a common systems-level signature — desynchronized, high-power slow-delta activity replacing the mixed, faster rhythms of wakefulness — and consciousness is lost when this happens. Cabral, Fernandes, and Shemesh (2023) supply a considerably more precise version of the same correlation. Using ultrafast fMRI in sedated and anesthetized rats, they show that spontaneous cortical activity organizes into a discrete

repertoire of standing-wave modes with fixed phase relationships across distant regions, and that both the peak frequency and the resonance quality (Q-factor) of these modes decline systematically and measurably with anesthetic depth, collapsing to near-aperiodic fluctuation under deep anesthesia. If the essay is going to grant the anesthesia argument any weight, this is the citation that earns it — a dose-dependent, quantitatively tracked degradation, not a binary before/after comparison.

But the mechanism Cabral et al. propose for these modes is structural, not computational. The modes are hypothesized to be eigenmodes of brain geometry itself — a repertoire set by anatomy and expressed spontaneously, at rest, independent of task content. This is a claim about the fixed architecture within which cortical activity operates, not a claim about how that architecture is used to route a particular decision or bind a particular memory on a particular trial. Spatial Computing, by contrast, needs waves to do a second and much more specific job: select and route specific task content, moment to moment, fast enough to produce the mixed selectivity and context-dependent routing examined in Sections III through V. These are different theoretical burdens, resting on different kinds of evidence, and Miller et al.'s paper does not keep them apart.

In the vocabulary introduced in Section VI, the distinction can be stated exactly. The Cabral et al. data are evidence for waves as one candidate physical realization of the fixed geometry of $\Gamma_C(x)$ itself — the anatomically constrained envelope of what is reachable at all, prior to any particular task or context. Nothing in Sections III through VII disputes that cortex has such an envelope, or that its physical realization plausibly involves oscillatory dynamics; the anesthesia data are good evidence for exactly this, and better evidence than anything the original paper cites in its own anesthesia section. What remains disputed is whether waves are additionally the mechanism that selects a specific trajectory *within* that envelope on a given trial — the claim Section III showed does not follow from mixed selectivity, subspace coding, or context-dependent routing, all of which recurrent dynamics without any wave mechanism already produce.

This is where Miller et al.'s argument borrows force it has not earned. Evidence that waves help constitute the boundary of what is reachable gets folded, without the distinction being drawn, into support for the separate claim that waves determine what gets selected inside that boundary. The anesthesia argument, however well quantified, bears only on the first claim. Nothing in the paper, and nothing in Cabral et al., speaks to the second.

One question this raises is not yet answered by any existing data, and is worth naming as a further target for the reanalysis strategy of Section VII: does task-relevant continuation structure collapse in step with the Q-factor as anesthesia deepens, or does some coarse structure survive the loss of oscillatory quality even as behavior itself is lost? This is the anesthesia-domain version of the same question Section VII asked of the drift and perturbation datasets — whether what degrades under a given manipulation is the wave, the manifold, or the reachable continuation set, and whether those three things degrade together or come apart. No existing dataset distinguishes them here either.

With this distinction in place, the essay's account of the paper's evidence is complete: the necessity claim examined in Section III does not hold, the paper's strongest recent functional

result (Chen et al.) remains genuinely open pending the test proposed in Section VII, and its anesthesia evidence, once separated from the claim it is often used to support, turns out to bear on a different and less contested question than the one the essay is asking. What remains is to locate this account among the standing theories of consciousness the paper positions itself against — a comparison that can now be brief, because the substantive work has already been done.

## IX. Positioning Relative to Consciousness Theories

Miller et al. position their account alongside Global Workspace Theory, Global Neuronal Workspace Theory, Integrated Information Theory, and Higher-Order Theories, differing from each mainly in mechanism: where these theories locate consciousness in broadcast, integration, or higher-order representation, Miller et al. locate it in the organizing action of wave dynamics across cortex.

The argument this essay has made about their neuroscience applies, at one remove, to this positioning as well. Each of the theories just named identifies a property reliably associated with conscious states — global broadcast, integrated information, higher-order representation, now wave organization — and treats the presence of that property as the explanation of consciousness rather than as a correlate that itself requires explaining. Section III showed this pattern operating at the level of specific neural phenomena: a property (wave organization) correlated with a function (flexible routing) was elevated to a claim of necessity that a fixed recurrent architecture, on inspection, does not require. The same structure recurs one level up. Wave organization correlating with conscious states, even robustly and even with the dose-dependent precision Cabral et al. provide, is not yet an account of why organization of that particular kind generates subjective experience rather than merely accompanying it. Nothing in this essay resolves that question, for wave theories or for its rivals; the point is only that the essay's central methodological move — distinguish what is correlated from what is required, and ask what invariant is shared across implementations — applies to the consciousness claim with exactly the same force it applied to mixed selectivity, and the paper does not extend that scrutiny to itself at this higher level.

## X. Conclusion

The argument of this essay can be stated in one sentence, and it is worth stating in exactly the form it was earned rather than a stronger one: cortical waves may implement and enhance cognitive control, but mixed selectivity and low-dimensional contextual dynamics do not uniquely implicate them, because recurrent systems lacking any wave-based mechanism already produce these phenomena from connectivity and population dynamics alone. The deeper explanatory target these phenomena point to is the organization of reachable task continuations — what a system can still do from a given state, not merely what can currently be read out of it — and wave dynamics are one candidate biological mechanism for establishing and selecting

among those continuations, not their definition.

Each section arrived at this conclusion by declining a stronger one that was available and tempting. Section III showed waves are not necessary for context-dependent routing, without concluding they are unimportant. Section IV conceded that manifold theory already explains a serious class of stability phenomena, without conceding the whole question. Section V identified a gap between decodability and reachability, without asserting that continuation structure fills it. Section VI introduced the minimal formal object needed to state that gap as a testable claim, without claiming the claim had been tested. Section VII specified the experiment that would test it, in both directions, and named what result would count against the essay's own proposal. Section VIII extended the same discipline to the anesthesia and consciousness evidence, separating what it establishes from what it is typically used to support.

The essay is therefore not a defense of continuation geometry against wave-based control theory. It is a demonstration that the two accounts have not yet been distinguished by any existing evidence, followed by a specification of what would distinguish them. The argument developed across Sections III through VIII ultimately points toward a layered account rather than a replacement of one framework by another: manifold structure constrains the repertoire of dynamics a population can generate at all, admissible continuation structure characterizes which futures within that repertoire remain reachable from a given state under task constraints, and oscillatory processes may participate in selecting among those reachable futures without thereby defining the space of what is reachable. Waves, on this picture, are not competing with continuations for the same explanatory role; they may be operating one layer down, inside a structure continuations describe and manifolds constrain. If the phase-manipulation experiment in Section VII shows that recovery, routing, and behavioral choice change under phase manipulation even when population trajectory geometry and its estimated continuation set are held fixed, the wave carries something the continuation account has left out, and Spatial Computing will have won a mechanistic point no correlational evidence could establish on its own. That is the essay's real bet, stated plainly rather than hedged: not that continuation structure is confirmed, but that this is the specific way, and the only way currently available, that the question could be settled rather than merely restated in different vocabulary.

## XI. Future Directions

Two kinds of follow-up work are worth distinguishing, because conflating them would misrepresent what this essay has and has not done.

The near-term work is empirical and already specified in Section VII: an estimation procedure for $\Gamma_C(x)$ from recorded neural trajectories, applied first as a reanalysis of existing datasets — Sadtler and Golub's perturbation data, Driscoll's drift recordings — before any new experiment is attempted. Nothing in the present essay constitutes that estimator; Section VII specifies what such an analysis would need to compare, not how to compute $\Gamma_C(x)$ from finite, noisy population recordings in practice, which is a nontrivial statistical problem in its own right and the most immediate open task the argument here creates.

A second, more mathematical line of extension was deliberately kept out of this essay and belongs, if pursued, in a separate treatment. Appendix B's continuation-equivalence relation invites formalization as a quotient construction, identifying states whose admissible continuation sets coincide or nearly coincide; distances between continuation sets could be defined via Hausdorff distance or a symmetric-difference measure, yielding a continuation metric rather than the coarser equivalence used here; the cardinality or measure of $\Gamma_C(x)$ suggests a continuation entropy, $H_C(x) = \log |\Gamma_C(x)|$, quantifying the number of admissible futures available from a state rather than merely their existence; and the equivalence and metric structures together suggest a category-theoretic treatment, with states as objects and admissibility-preserving transitions as morphisms, connecting the present proposal to a broader continuation-geometric framework. None of this machinery is needed to state or test the empirical claim this essay makes, and adding it here would obscure rather than strengthen that claim. It is noted only to mark where the formal object introduced in Section VI could go if the reachability hypothesis survives the empirical test proposed in Section VII — a question that should be settled before, not instead of, building further mathematics on top of it.

# Appendices

## A Admissible Continuations and Viability Theory

Section VI's formal object was introduced without reference to the mathematical literature it most resembles, on the grounds that the essay's empirical claim does not depend on that resemblance. This appendix makes the resemblance precise, because doing so preempts an objection a mathematically literate reader will otherwise raise unprompted, and because the precise relationship turns out to clarify rather than undermine the proposal.

Let a population's dynamics be given by $\dot{x} = F(x)$ on a state space $X$, and let $K \subseteq X$ denote the instantaneous task-admissible region: the set of states satisfying the pointwise constraint $C$ introduced in Section VI. In this notation, the admissible continuation set from Section VI can be rewritten as

$$\Gamma_C(x) = \{\, \gamma : [0, T] \to X \mid \gamma(0) = x,\ \dot{\gamma}(t) = F(\gamma(t)),\ \gamma(t) \in K \text{ for all } t \in [0, T] \,\}.$$

This is, on its face, close to the central object of Aubin's viability theory [1]. Recall that the viability kernel of $K$ under $F$, written $\mathrm{Viab}(K)$, is the largest subset of $K$ from which at least one trajectory of $F$ remains in $K$ for all future time. The classical viability theorem states that $\mathrm{Viab}(K)$ is well-defined and that a trajectory persisting in $K$ indefinitely exists from $x$ if and only if $x \in \mathrm{Viab}(K)$.

**Proposition.** If $K = \mathrm{Viab}(K)$ — that is, if $K$ is already a viability kernel in Aubin's sense, closed under the dynamics — then for any finite horizon $T$, $\Gamma_C(x) \neq \varnothing$ if and only if $x \in K$.

The proof is immediate from the definition: if $K$ is itself a viability kernel, every point of $K$ admits at least one trajectory remaining in $K$ indefinitely, and a fortiori for any finite $T$; conversely, no trajectory from a point outside $K$ can satisfy the constraint at $t = 0$, so $\Gamma_C(x)$ is empty by definition.

Two disanalogies are worth stating precisely, because they are what keep $\Gamma_C(x)$ from being simply a restatement of $\mathrm{Viab}(K)$ under a new name. First, classical viability theory is typically posed over an infinite or unspecified horizon, and asks a binary question — does a viable trajectory exist at all. Section VI's object is explicitly finite-horizon, tied to $T$, because neural tasks have finite-horizon structure (a trial, a delay period, a decision window), and the interesting empirical comparisons in Section VII concern how $\Gamma_C(x)$ changes across finite, task-relevant windows, not whether it is eventually empty. Second, and more importantly, viability theory in its standard form is posed for controlled systems, where $F$ depends on a control input the theory is implicitly asking whether some admissible control policy exists; $\Gamma_C(x)$ as used in this essay presupposes no control input and no policy — $F$ is the population's autonomous or externally-driven dynamics as recorded, not a system being optimized over. The empirical claim in Section VII is about which existing trajectory a population follows and what set of nearby trajectories remain compatible with $C$, not about the existence of some optimal controller.

Given these two disanalogies, the appropriate framing is not that $\Gamma_C(x)$ is a novel construction, but that it is a finite-horizon, uncontrolled, trajectory-level refinement of viability-style admissibility — asking not only whether $K$ can be maintained, but how many ways it can be maintained, and from exactly which states, over exactly the timescales a given task imposes. The continuation framework may therefore be interpreted as sitting inside the viability literature rather than beside it, which is a stronger and more easily defended position than the claim of conceptual originality the main text was careful never to make.

# B Reachability Preservation versus Decodability Preservation

The argument of Sections V through VII depends on the logical possibility that a population can preserve a decodable readout while its admissible continuation set changes. This appendix demonstrates that possibility directly with a minimal worked example, rather than leaving it as an assertion about what the existing literature has not yet ruled out.

Let $X = \mathbb{R}^2$, with state $x = (a, b)$. Let the decoder be the projection onto the first coordinate, $D(a, b) = a$, a natural analogue of a task-relevant readout dimension in the sense of Kaufman et al.'s potent space: $a$ is the decodable variable, and $b$ is a null-space coordinate the decoder does not read. Let the dynamics be linear and autonomous,

$$\dot{a} = 0, \qquad \dot{b} = b,$$

so that $a$ is conserved exactly and $b$ grows exponentially: $b(t) = b(0)\, e^t$. Let the task constraint be a bound on the null coordinate, $K = \{(a, b) : |b| \leq 1\}$, representing a physiological or task-structural bottleneck — for instance, a bound on some latent variable that must not saturate for the trial to remain admissible, of the kind introduced operationally in Section VI. Since the horizon $T$ is fixed throughout Section VI but will vary explicitly in what follows, write $\Gamma_C^{(T)}(x)$ for the admissible continuation set defined over $[0, T]$, making the horizon dependence that was previously left implicit visible in the notation.

Consider two states with identical decodable readout: $x = (a_0, \varepsilon)$ and $y = (a_0, 1 - \varepsilon)$ for some small $\varepsilon > 0$. Since $D$ depends only on the first coordinate, $D(x) = D(y) = a_0$: the two states are indistinguishable to any readout built from $D$, at $t = 0$ and, since $a$ is conserved under the given dynamics, at every subsequent time as well. Decodability is preserved between $x$ and $y$ for the entire trajectory, trivially.

Their admissible continuation sets are not equal, for a suitable choice of horizon. Under the given dynamics, $b(t) = \varepsilon\, e^t$ from $x$, which remains within $[-1, 1]$ for all $t \leq \ln(1/\varepsilon)$; choosing $\varepsilon$ small makes this threshold arbitrarily large, so for any fixed task horizon $T$, $\varepsilon$ can be chosen small enough that $\Gamma_C^{(T)}(x) \neq \varnothing$. From $y$, $b(t) = (1 - \varepsilon)\, e^t$ exceeds 1 almost immediately — at $t = \ln(1/(1 - \varepsilon)) \approx \varepsilon$ for small $\varepsilon$ — so once $T$ exceeds this small threshold, $\Gamma_C^{(T)}(y) = \varnothing$: no trajectory from $y$ satisfies the constraint for the full horizon $[0, T]$, regardless of what happens along the way. The horizon dependence is essential and not a technicality to be suppressed: for $T$ below the threshold, $\Gamma_C^{(T)}(y)$ is also nonempty, since $y$ has not yet left $K$. What the example shows is that a fixed, shared, task-relevant horizon $T$ — the kind Section VI's operational

constraints are built to specify in advance — is enough to separate the two sets even though no finite-time distinction exists at every horizon simultaneously.

**Example (formal statement).** For the system above, for any fixed task horizon $T > 0$, there exist $x, y \in X$, depending on $T$, such that $D(x) = D(y)$ at every $t \in [0, T]$, while

$$\Gamma_C^{(T)}(x) \neq \varnothing \qquad \text{and} \qquad \Gamma_C^{(T)}(y) = \varnothing.$$

This is a minimal example, not a claim about what actual neural populations do; nothing here asserts that cortical null-space dynamics behave like an unstable linear system, or that real task constraints take this particular form. What the example establishes is only the logical point the main text needs and does not independently prove: decodability-preservation and reachability-preservation are not logically equivalent claims, even in the simplest systems capable of expressing the distinction, so the empirical question raised in Section V — whether real neural populations behave more like $x$-and-$y$-are-equivalent or $x$-and-$y$-are-not — is a genuine, non-vacuous question rather than one settled by definition.

## References


[1] Aubin, J.-P. (1991). *Viability Theory*. Birkhäuser.

[2] Cabral, J., Fernandes, F. F., & Shemesh, N. (2023). Intrinsic macroscale oscillatory modes driving long range functional connectivity in female rat brains detected by ultrafast fMRI. *Nature Communications*, 14, 375. https://doi.org/10.1038/s41467-023-36025-x

[3] Chen, Z., Brincat, S. L., Lundqvist, M., Loonis, R. F., Warden, M. R., & Miller, E. K. (2026). Oscillatory control of cortical space as a computational dimension. *Current Biology*, 36, 402–414.e5. https://doi.org/10.1016/j.cub.2025.11.072

[4] Driscoll, L. N., Pettit, N. L., Minderer, M., Chettih, S. N., & Harvey, C. D. (2017). Dynamic reorganization of neuronal activity patterns in parietal cortex. *Cell*, 170(5), 986–999.e16. https://doi.org/10.1016/j.cell.2017.07.021

[5] Golub, M. D., Sadtler, P. T., Oby, E. R., Quick, K. M., Ryu, S. I., Tyler-Kabara, E. C., Batista, A. P., Chase, S. M., & Yu, B. M. (2018). Learning by neural reassociation. *Nature Neuroscience*, 21(4), 607–616. https://doi.org/10.1038/s41593-018-0095-3

[6] Kaufman, M. T., Churchland, M. M., Ryu, S. I., & Shenoy, K. V. (2014). Cortical activity in the null space: permitting preparation without movement. *Nature Neuroscience*, 17, 440–448. https://doi.org/10.1038/nn.3643

[7] Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, 503(7474), 78–84. https://doi.org/10.1038/nature12742

[8] Miller, E. K., Brincat, S. L., & Roy, J. E. (2026). *Analog cognition and consciousness* [Unpublished manuscript]. Picower Institute for Learning & Memory and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology.

[9] Rule, M. E., Loback, A. R., Raman, D. V., Driscoll, L. N., Harvey, C. D., & O'Leary, T. (2020). Stable task information from an unstable neural population. *eLife*, 9, e51121. https://doi.org/10.7554/eLife.51121

[10] Sadtler, P. T., Quick, K. M., Golub, M. D., Chase, S. M., Ryu, S. I., Tyler-Kabara, E. C., Yu, B. M., & Batista, A. P. (2014). Neural constraints on learning. *Nature*, 512(7515), 423–426. https://doi.org/10.1038/nature13665

[11] Vyas, S., Golub, M. D., Sussillo, D., & Shenoy, K. V. (2020). Computation through neural population dynamics. *Annual Review of Neuroscience*, 43, 249–275. https://doi.org/10.1146/annurev-neuro-092619-094115