Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in
Neurobiology

Representational drift: Emerging theories for continual learning and experimental future directions

Laura N. Driscoll[1], Lea Duncker[2] and Christopher D. Harvey[3]

Abstract

Recent work has revealed that the neural activity patterns correlated with sensation, cognition, and action often are not stable and instead undergo large scale changes over days and weeks—a phenomenon called representational drift. Here, we highlight recent observations of drift, how drift is unlikely to be explained by experimental confounds, and how the brain can likely compensate for drift to allow stable computation. We propose that drift might have important roles in neural computation to allow continual learning, both for separating and relating memories that occur at distinct times. Finally, we present an outlook on future experimental directions that are needed to further characterize drift and to test emerging theories for drift's role in computation.

Addresses

[1] Department of Electrical Engineering, Stanford University, Stanford, CA, USA
[2] Howard Hughes Medical Institute, Stanford University, Stanford, CA, USA
[3] Department of Neurobiology, Harvard Medical School, Boston, MA, USA

Corresponding authors: Harvey, Christopher D. (harvey@hms.harvard.edu) Driscoll, Laura N. (lndrisco@stanford.edu); Duncker, Lea (lduncker@stanford.edu)
(Driscoll L.N.), (Duncker L.)

Current Opinion in Neurobiology 2022, **76:**102609

This review comes from a themed issue on **Systems Neuroscience**

Edited by **Joshua Johansen** and **Laura L Colgin**

For a complete overview see the Issue and the Editorial

Available online xxx

https://doi.org/10.1016/j.conb.2022.102609

Recent advances in experimental recording techniques have made it possible to record large populations of neurons with high spatial resolution [1,2,3]. This has facilitated improvements in methods for tracking activity of the same neurons across days and weeks, providing insight into how population activity may change over timescales exceeding the typical length of a recording session [4,5]. Data from both hippocampus and neocortex have revealed that, on the timescale of several days, many neurons maintain consistent activity patterns [6,7]. However, neural activity correlations with sensory and behavioral variables such as presented stimuli, environmental cues, or actions can drastically change over a period of weeks [8,9,10,11,12,13]. A small number of neurons make mostly discrete changes from one day to the next. This means that correlations of neural activity with environmental and behavioral variables are mostly stable over this short time-window. Over the course of weeks, however, a slow and gradual accumulation of change at the population level results in changing correlations between the activity of the recorded population of neurons and these variables.

Cells whose activity was previously correlated with environmental and behavioral variables are most frequently no longer active in response to the same variables weeks later. At the same time, a mostly new pool of neurons develops activity patterns correlated with these variables. Less commonly, cells previously correlated with certain variables become correlated with new variables. In one example of these findings, as mice move through an environment, a subpopulation of hippocampal place cells are informative about the animal's spatial position. Over time, the pool of place cells that make up this spatial code changes, despite the environment staying the same [13]. These features of neural data have been reported even in the context of an animal stably performing a learned behavioral task equally well over weeks, in a brain region shown to be necessary for the task [8]. While few neurons maintain their activity patterns over weeks, population level statistics such as the fraction of task-correlated neurons and the overall activity levels in the population remain largely unchanged [8,14,15,10,11,12,13].

This phenomenon has been observed using both calcium imaging and electrophysiology in relation to a diverse set of stimuli and behaviors in the hippocampus [9,11,13] (although see Ref. [16]), posterior parietal cortex [8], anterior piriform cortex [12], and even primary visual cortex [17,10]. Drift has generally not been observed in motor areas [18,19,20,21] (although see Refs. [22,23]). In addition to these experimental results, theoretical work has focused on how these changes may arise, how downstream areas may overcome nonstationarities upstream, as well as potential computational benefits [24,25,26,23,27,28].

For consistency with previous work, we refer to this phenomenon as *representational drift* [27]. We note, however, that the terms "representation" and "drift" both make assumptions about the nature of changes in neural activity and potentially limit possible explanations for this phenomenon. Representational drift suggests that the brain has an internal map, or representation, of the external world that changes over time. However, brain activity evolves in time to perform computations on internal variables that extend beyond creating representations of the external world [29,30]. For example, if an animal is trained to navigate to the right in response to a given stimulus, the brain does not necessarily need to represent the stimulus. Rather, some evolution of neural activity must result in an eventual right turn. While representations can be useful for flexibility and generalization and may be appropriate descriptions of neural activity in some brain areas, it is unlikely that a snapshot of neural activity in all brain areas is best considered as a representation of a stimulus [31,32]. Therefore, although the term "representation" is convenient for naming this phenomenon, the implications limit the scope of potential explanations. We argue that reframing neural activity as interacting with the world, rather than representing it, may be critical in revealing how gradual changes in the relationship between neural activity and the environment reflect a continuously evolving dynamical system. Additionally, while drift has specific meanings in physics, including distinctions from diffusion, we propose that the exact form of the changes needs further characterization.

In this review, we discuss recently proposed alternative explanations of representational drift, review how the brain might overcome nonstationarities, and propose a computational role for drift in continual learning. We end with an outlook on future work to move toward a better understanding of the properties and functions of drift.

Alternate explanations for representational drift

There have been several proposals for how experimental confounds could explain the observed changes in neural activity patterns across time. In this section, we review these proposals but argue that they are unlikely to fully account for representational drift.

First, long term recordings damage neural tissue due to the implantation of electrophysiology probes [33], cranial window surgeries [34], or long-term expression of calcium indicators [35]. However, electrical and optical recording methods have revealed drift with similar rates and statistics, even though their effects on tissue damage might be expected to occur at different rates and scales [17]. The drift rate is often constant over weeks, suggesting it is a steady-state phenomenon instead of a response to a specific event. Further, and perhaps most convincingly, drift rates in the same animal can differ across contexts [12].

Second, identifying the same cells across days is challenging. Cells are identified in electrical recordings based on the action potential waveform, which can have similar shapes for different cells. Even continuous recordings suffer from movement of the brain relative to the electrode. Optical imaging relies on matching cells across days based on their spatial fluorescence profile, which can be difficult if indicator expression is unstable or low. Optical tracking of cells is especially difficult if supra-cellular resolution imaging methods are used. However, in some cases, researchers have gone to great lengths to ensure the reliable identification of neurons, through painstaking visual inspection of each individual neuron in imaging data [8] and extensive electrophysiological validation [18,12,3]. Furthermore, drift rates vary depending on the familiarity of sensory stimuli, suggesting that drift is related to factors other than cell identity assignments [12].

Third, it is possible that while neural activity is correlated with measured environmental and behavioral variables, this activity is only computationally relevant for variables that are not measured [36]. In this case, neural activity may be stable and only appear to drift due to erroneous inference of the neural activity—behavior relationship. To mitigate these confounding factors, some experiments have recorded a large number of external variables and modeled each neuron's activity as a function of these variables [8]. Even in this case, drift was still identified. In these models, time was more predictive of a neuron's activity levels over weeks than were behavioral variables. In addition, perturbations of the recorded region impact behavioral performance in a manner consistent with causal influence of neural activity on measured behavior, lending support to the validity of the correlations studied. Behavioral confounds are a difficult problem to rule out completely. Great efforts must be taken to record and model as many features of behavior as possible to better understand the evolving relationship between neuronal activity and behavior [16,37,38,39,40,41].

Does representational drift create challenges for stable behavior?

The brain is not made of silicon [42]. Synaptic connections in many parts of the brain are not stable over long time periods and instead turn over at high rates, even with complete turnover in a month in the hippocampus [43]. However, a subset of cortical spines nonetheless appears to be stable over extended periods

[44]. Most biological elements exhibit turnover, including the proteins that make up synapses and define the strength and signaling of synapses [45,46,47]. Biological systems therefore have homeostatic mechanisms to take advantage of under-constrained variables in the system in order to maintain consistency in crucial features of cells and networks [48]. Representational drift could arise for this reason and thus be a biological inevitability. In this case, drift might be a nuisance that the brain needs to overcome, instead of a feature for computation. However, it is possible that drift could reflect both challenges and benefits for computation, since computational strategies have evolved in the presence of these biological constraints.

Recent work has examined if drift creates challenges for the nervous system to encode and decode behaviorally relevant information. In high dimensional systems, such as a large population of neurons, random changes in neural activity patterns might be expected due to turnover of biological elements [49]. These changes are likely to occur in dimensions separate from those relevant to low-dimensional cognitive computations [50,51,23,52]. This is because most dimensions in a high-dimensional space are nearly orthogonal to one another. However, populations of neurons likely code for many different features [53,40], making it unclear if drift would be mostly orthogonal to all these coding dimensions. Experimental evidence indicates that drift impacts coding of environmental and behavioral variables and thus is not strictly confined to “null” dimensions [8,28]. If these changes go uncorrected, then decoding in a neural population appears to degrade over the time course of several weeks, indicating that the nervous system must have methods to counteract drift. Theoretical work demonstrates that drift could be compensated by synaptic plasticity mechanisms that only require small weight changes over several days to maintain a stable readout of information by a downstream network [28,54]. However, the regulatory mechanisms that could coordinate these changes are unknown.

Theoretical proposals of mechanisms underlying representational drift

Since the discovery of representational drift in neural data, theoretical works have put forward proposals for network level mechanisms that could result in representational drift. Most of this work focuses on noisy synaptic weight updates that evolve to minimize an overall objective function of the network under Hebbian or anti-Hebbian plasticity [55,56,57,58,26,59]. Stochasticity prevents the network from ever fully converging to a stable solution, which can lead to a diffusion around a local optimum in weight space [56] or exploration of multiple local optima relating to redundant solutions of the overall objective function [59]. It has been proposed that this could have a role in generalization performance and probabilistic computation [55,58]. Characterization of dynamical systems in the context of changes in synaptic efficacy reveals that time varying attractors are easier to maintain than fixed points [60]. Recently, it has also been proposed that drift could have computational benefits during learning as a regularization strategy similar to drop-out in artificial neural networks [61,62]. It will be of interest to explore how drift may be related to continual learning under changing objective functions with varying task demands.

Drift as a useful feature for neural computation in the context of continual learning

It is currently unknown whether representational drift is merely a nuisance that the brain needs to overcome or if it may also serve a specific purpose with potential computational benefits. Theories of learning in network models have long considered a plasticity-stability tradeoff for memory capacity [63,64,25]. It has been proposed that memories are actively or passively reorganized over time for optimal storage [65]. Modeling this process of reorganization has been the subject of several theoretical studies [66,65,67,68]. Similar considerations about plasticity and stability may also play an important role in studying representational drift as a useful feature for learning. Recent work on memory allocation and consolidation may present a way to link this body of theoretical work to observations about representational drift.

The memory allocation hypothesis states that neurons with high excitability are more likely to be recruited for memory encoding [69,70]. The pool of active neurons that participate in computation changes over time, in part due to changes in the excitability of neurons [71]. Thus, at any given time, the pool of “excitable” neurons is the one that will learn new associations [72]. This framework implies that drift might be relevant for continual learning in order to avoid catastrophic forgetting that might occur if new memories are incorporated into indiscriminate sets of neurons and synapses. Drift could play a role in continual learning by allowing for learning without interference with previously learned associations [73]. By continuously updating the pool of active neurons, new information could be incorporated into distinct computational resources that do not interfere with previously learned information [74,75,11].

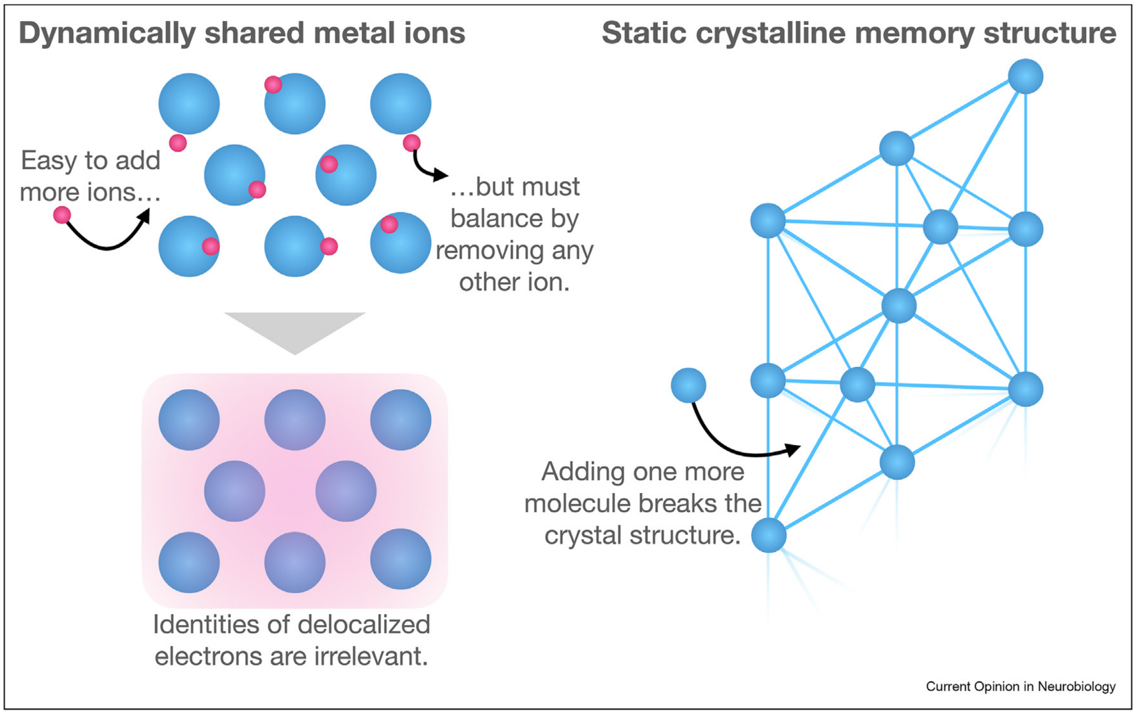
Another aspect of the memory allocation hypothesis is that drift could be a key feature of consolidation, linking memories that are separated in time [76,77,78,79]. Artificial neural networks trained to perform related and interfering computations utilize shared resources for similar computations [74]. It might be beneficial for related associations to share similar neural resources, even if they were learned at different times. This would suggest that new information should eventually be stored in a similar reference frame to previously learned associations. Drift enforces a baseline malleability, constantly updating the structure and organization of computation (see Fig. 1). Thus, as memories are rehearsed or revisited, either actively through experience or inactively during replay events, memories could be transferred to new sets of neurons [25]. This continuous update in the neurons that perform computation could result in the merging of new and old associations when appropriate. Both associations would merge into the "excitable" pool of neurons. Indeed, as mice continued to practice a navigation decision task, the population of neurons in the posterior parietal cortex that coded for new cues were overlapping with neurons most recently active on familiar cue trials [8]. In this framework, by continuously rewriting memories as they are revisited, new information ultimately leads to refined and improved memories and internal models. A prediction of this hypothesis would be that neurons important for both previously learned and novel computations are those that are newly active (Fig. 2). Thus, in some cases, it might be beneficial to maintain new and old memories in separate sets of neurons in order to ensure separability of these memories. Alternatively, by continually rewriting memories as they are revisited, new learning could refine internal models and establish a compatible reference frame for relating new and old information. Drift could play a role in both these functions for continual learning.

In addition, the brain does not need to maintain a perfect memory of every learned association. Drift could add noise so that associations that do not continue to be useful are not updated and therefore are forgotten. Drift could thus provide a mechanism for forgetting information that is no longer useful to guide future behaviors.

Experimental challenges for the future

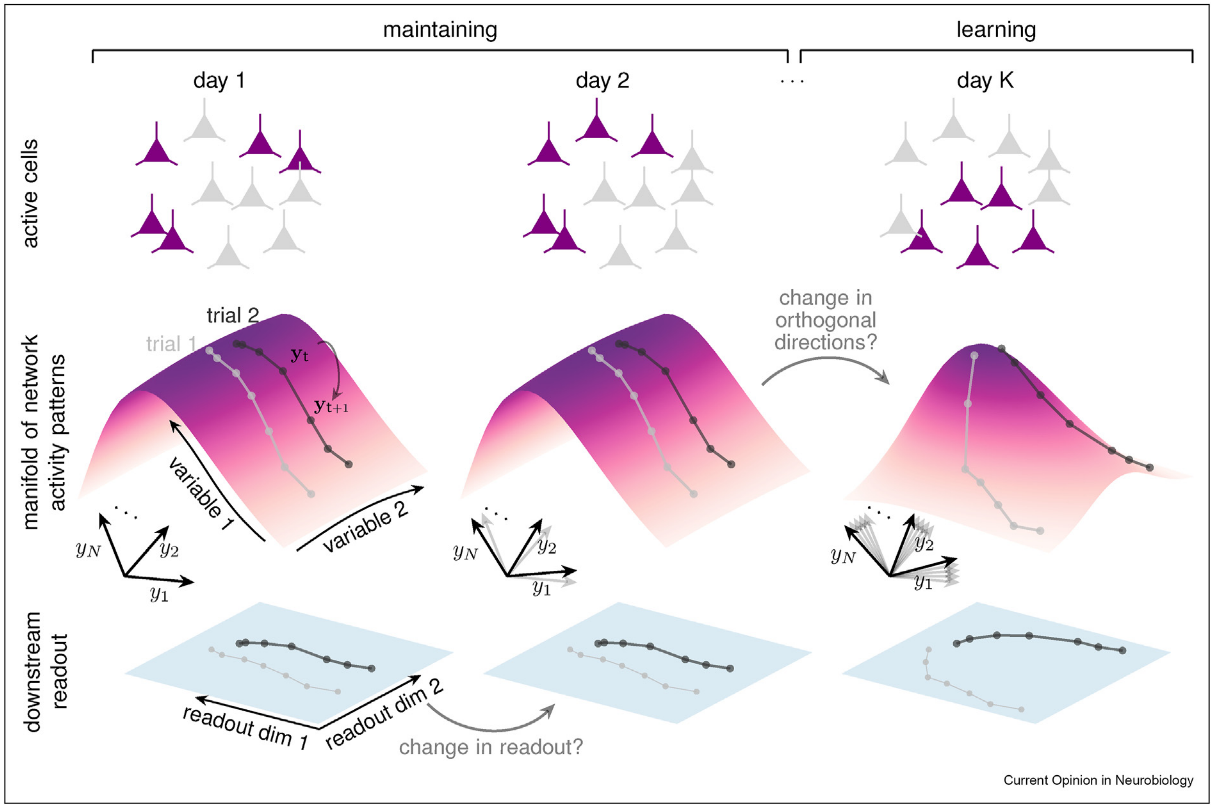
Although many observations of drift have been reported, much more experimental work is needed to systematically and quantitatively characterize drift, including comparisons across brain regions, cell types, and learning conditions. Furthermore, while theoretical frameworks have been developed to better understand the

Figure 1



Dynamically reconfigured computations may be implemented by different neurons over time. This idea is somewhat analogous to shared electrons in a lattice of positively charged metal ions. In either case, the identities of the neurons and the electrons are both irrelevant. Rather, emergent properties of the system such as population dynamics, or the overall charge of negatively and positively charged metal ions are crucial features of the system that must be maintained over time. It is easier to imagine how new tasks and memories might be 'mixed' into a continuously evolving system compared to one that maintains a rigid and fixed relationship between inputs and outputs. This alternative rigid structure is analogous to a crystal lattice, where the relationships between molecules must be broken down in order to make any changes to the system.

Figure 2



As the pool of active cells changes on each day, the manifold of network activity patterns shifts and rotates in neural state space (day 1 to day 2). We propose that shifts and rotations in the position of the manifold occur during baseline maintenance of previously learned computations. New learning results in changes in the geometry of the neural manifold (day K). To prevent manifold changes from interfering with previously learned computation, new features might be added in orthogonal dimensions. When a new pool of excitable cells develops activity correlations with the new features that are learned, manifold changes are orthogonal to previously learned dimensions that were encoded by previously excitable cells. This orthogonal change would allow for continual learning without changes to previously learned dimensions of continuously drifting manifolds.

computational implications of drift, if possible, drift must be manipulated to test these ideas.

A comprehensive study of the drift rate across brain regions in the context of the same task and during tasks with different levels of complexity could be highly informative about the cause and function of drift. Relating drift rates across regions is challenging when comparing across studies, stimuli, and tasks. Therefore, experiments are needed to measure drift rates in multiple brain regions in the same animal, with the same measurement and analysis techniques, and during the identical stimuli and tasks. If drift is similar across areas, then we might interpret this finding as drift being an inevitability of biological networks. Instead, if drift is slower in areas more closely related to the external world, including areas involved in processing sensory stimuli or generating motor outputs, and highest in association areas, then theories relating drift to continual association learning might be favored. Other causes of different drift rates might include different connectivity patterns, cell types, or input statistics. It has not been examined if different cell types drift at different rates, including cells with specific projection targets or molecular profiles.

Drift rates might differ for different kinds of behavioral variables. While we know that different areas of the brain preferentially encode different types of variables, recent work has shown that each cortical area often encodes a diversity of sensory, cognitive, and motor variables [53,40]. Rather than considering drift as a property of a brain region, it will be interesting to determine if some variables are encoded in a more stable way than others in the same region. In this case, we might predict that drift rates differ for neurons or synapses depending on what they encode. For example, in V1, we might expect that feedforward inputs for building sensory representations might have a high degree of stability, whereas recurrent intracortical inputs to V1 that mediate visual learning might have faster drift rates [10].

Beyond these factors, it is important to test whether the experience of an animal is a key factor that could contribute to drift rates in order to understand if drift plays a role in learning and consolidation. Recent work has shown that the frequency of an odor stimulus, but not its valence, affects drift rate in the piriform cortex [12]. These findings are suggestive that continuous interactions with a particular stimulus might stabilize representations as they continue to be relevant in the environment, which is consistent with previous work in the hippocampus [9,80,81]. Some of the most striking examples of stability in the relationship between neural activity and behavior have been discovered in animals producing highly stereotyped behaviors [19,20,16]. It will be of interest to determine whether the stability reported in motor areas depends on the familiarity of the stereotyped behaviors or the sensitivity of the stereotyped behavioral output to small changes in neural activity. Further experiments will be needed to understand whether drift rate varies according to the relevance of a behavioral variable in other settings. In the context of learning, it will be interesting to assess whether codes for variables that have been associated during the course of learning develop coordinated drift kinetics. Determining whether enriched versus impoverished home cages or the age of an animal relates to drift may help test a role of drift in learning [82].

Ultimately, the best test of the proposed roles for drift in computation would be experiments to manipulate the drift rate. Several ideas have emerged on how this might be done. One possibility is that drift could be stabilized in a brain-machine interface in which decoders require stable activity to function across days [83,84]. Another possibility is that drift could be altered with optogenetic or genetic manipulations. Optogenetics, in particular with cellular resolution targeting, may allow an experimenter to change the "excitability" of targeted neurons to force them into or out of the "excitable" population [85,86]. A genetic manipulation could involve altering the expression of ion channels that control cellular excitability, which studies have suggested are important for the turnover in the "excitable" pool of neurons [76,87,88]. If perturbations that decrease the drift rate result in lessened behavioral flexibility, such results would support the idea that representational drift is useful for continual learning.

Finally, experiments are necessary to connect the phenomenon of drift to other established principles in learning and memory. As just one example, memory engrams are a prominent concept for hippocampal memory [89]. Although engrams are often discussed in a manner that presupposes stability of neural activity patterns and synaptic connections over time, recent work suggests that dynamic neural ensembles contribute to engrams [90]. Is it possible to connect drift and engrams into a unified concept? As a starting point, it will be of interest to know if engram neurons, which are typically labeled by induction of the immediate early gene c-fos, have lower drift than neighboring neurons. Additionally, if there is drift in the engram population, then it will be of interest to determine if there are mechanisms to transition the memory trace across different populations of neurons. This transition would effectively bind together "old" and "new" engram cells for a single memory; in which case, an engram might be an ever-evolving population of neurons [70,90].

In summary, exciting progress has been made to demonstrate that drift is a common phenomenon, with a recent surge in studies observing drift. Inspiring ideas from theory have emerged about the potential computational implications and functions of drift, providing a rich conceptual foundation. However, the study of representational drift is still in its infancy. The next steps will require a deeper experimental characterization of drift, attempts to identify testable distinctions between different theoretical proposals, and ultimately a close iteration of the emerging experimental and theoretical investigations into drift.

Conflict of interest statement

Nothing declared.

Acknowledgments

This work was supported by NIH grant R01 NS089521, NIH Director's Pioneer Award DP1 MH125776, and the Human Frontier Science Program Award RGY0069.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Jun JJ, Steinmetz NA, Siegle JH, *et al.*: **Fully integrated silicon probes for high-density recording of neural activity**. *Nature* 2017, **551**:232–236.

2. Sofroniew NJ, Flickinger D, King J, *et al.*: **A large field of view two-photon mesoscope with subcellular resolution for in vivo imaging**. *Elife* 2016, **5**, https://doi.org/10.7554/eLife.14472.

3. Steinmetz NA, Aydin C, Lebedeva A, *et al.*: **Neuropixels 2.0: a miniaturized high-density probe for stable, long-term brain recordings**. *Science* 2021, **372**, https://doi.org/10.1126/science.abf4588.

4. Lee CR, Najafizadeh L, Margolis DJ: **Investigating learning-related neural circuitry with chronic in vivo optical imaging**. *Brain Struct Funct* 2020, **225**:467–480.

5. Luo TZ, Bondy AG, Gupta D, *et al.*: **An approach for long-term, multi-probe Neuropixels recordings in unrestrained rats**. *Elife* 2020, **9**, https://doi.org/10.7554/eLife.59716.

6. Chestek CA, Batista AP, Santhanam G, *et al.*: **Single-neuron stability during repeated reaching in macaque premotor cortex**. *J Neurosci: the official journal of the Society for Neuroscience* 2007, **27**:10742–10750.

7. Mank M, Santos AF, Direnberger S, *et al.*: **A genetically encoded calcium indicator for chronic in vivo two-photon imaging**. *Nat Methods* 2008, **5**:805–811.

8. * Driscoll LN, Pettit NL, Minderer M, *et al.*: **Dynamic reorganization of neuronal activity patterns in parietal cortex**. *Cell* 2017, **170**:986–999. e16.

As mice stably perform a navigation-based decision-making task, posterior parietal cortical neurons exhibit unstable activity correlations with measured behavioral variables over a period of weeks. Most change is related to gaining or losing correlations with measured variables. Inactivating the posterior parietal cortex resulted in a drastic reduction in task performance.

9. Kentros CG, Agnihotri NT, Streater S, *et al.*: **Increased attention to spatial context increases both place field stability and spatial memory**. *Neuron* 2004, **42**:283–295.

10. * Marks TD, Goard MJ: **Stimulus-dependent representational drift in primary visual cortex**. *Nat Commun* 2021, **12**:5169.

Chronic 2-photon imaging of mouse V1 populations reveal representational drift in response to naturalistic movies and stable responses to classical gratings stimuli. Cells that are responsive to both movies and gratings have different drift rates in relation to either stimulus. These results suggest representational stability in V1 is stimulus-dependent.

11. Rubin A, Geva N, Sheintuch L, *et al.*: **Hippocampal ensemble dynamics timestamp events in long-term memory**. *Elife* 2015, **4**, https://doi.org/10.7554/eLife.12247.

12. * * Schoonover CE, Ohashi SN, Axel R, *et al.*: **Representational drift in primary olfactory cortex**. *Nature* 2021, **594**:541–546.

Electrophysiological recordings of single units maintained over weeks reveal representational drift of odor-evoked responses in mouse piriform cortex. Authors report different rates of drift within the same population of cells for familiar and unfamiliar odors. Daily exposure to the same odorant slows the rate of drift, but when exposure is halted the rate increases again.

13. Ziv Y, Burns LD, Cocker ED, *et al.*: **Long-term dynamics of CA1 hippocampal place codes**. *Nat Neurosci* 2013, **16**:264–266.

14. Gallego JA, Perich MG, Naufel SN, *et al.*: **Cortical population activity within a preserved neural manifold underlies multiple motor behaviors**. *Nat Commun* 2018, **9**:4233.

15. Gallego JA, Perich MG, Chowdhury RH, *et al.*: **Long-term stability of cortical population dynamics underlying consistent behavior**. *Nat Neurosci* 2020, **23**:260–270.

16. Liberti WA, Schmid TA, Forli A, *et al.*: **A stable hippocampal code in freely flying bats**. *Nature* 2022:1–6. Nature Publishing Group.

17. * Deitch D, Rubin A, Ziv Y: **Representational drift in the mouse visual cortex**. *Curr Biol: CB* 2021, **31**:4327–4339. e6.

Authors analyze publicly available calcium imaging and electrophysiology datasets from multiple brain regions as mice view natural movies. The findings contradict one theory that sensory areas should be less susceptible to drift than association areas.

18. Dhawale AK, Poddar R, Wolff SB, *et al.*: **Automated long-term recording and analysis of neural activity in behaving animals**. *Elife* 2017, **6**, https://doi.org/10.7554/eLife.27702.

19. Jensen KT, Harpaz NK, Dhawale AK, *et al.*: **Long-term stability of neural activity in the motor system**. *bioRxiv* 2021, https://doi.org/10.1101/2021.10.27.465945.

20. * Katlowitz KA, Picardo MA, Long MA: **Stable sequential activity underlying the maintenance of a precisely executed skilled behavior**. *Neuron* 2018, **98**:1133–1140. e3.

HVC neural activity correlations with birdsong are in major contrast to reports of drift. Rather, neurons exhibit a highly stable relationship with behavior over weeks to months. These results call into question whether measured drift is a feature of subtle instabilities of the behaviors studied in other contexts.

21. Stevenson IH, Cherian A, London BM, *et al.*: **Statistical assessment of the stability of neural movement representations**. *J Neurophysiol* 2011, **106**:764–774.

22. Liberti 3rd WA, Markowitz JE, Perkins LN, *et al.*: **Unstable neurons underlie a stable learned behavior**. *Nat Neurosci* 2016, **19**:1665–1671.

23. Rokni U, Richardson AG, Bizzi E, *et al.*: **Motor learning with unstable neural representations**. *Neuron* 2007, **54**:653–666.

24. Ajemian R, D'Ausilio A, Moorman H, *et al.*: **A theory for how sensorimotor skills are learned and retained in noisy and nonstationary neural circuits**. *Proc Natl Acad Sci USA* 2013, **110**:E5078–E5087.

25. Káli S, Dayan P: **Off-line replay maintains declarative memories in a model of hippocampal-neocortical interactions**. *Nat Neurosci* 2004, **7**:286–294.

26. Masset P, Qin S, Zavatone-Veth JA: **Drifting neuronal representations: bug or feature?** *Biol Cybern* 2022, https://doi.org/10.1007/s00422-021-00916-3.

27. Rule ME, O'Leary T, Harvey CD: **Causes and consequences of representational drift**. *Curr Opin Neurobiol* 2019, **58**:141–147.

28. * * Rule ME, Loback AR, Raman DV, *et al.*: **Stable task information from an unstable neural population**. *Elife* 2020, **9**, https://doi.org/10.7554/eLife.51121.

Analysis of data from Driscoll et al. 2017 reveals that drift occurs mostly in coding null dimensions, although a significant component of drift continually degrades a fixed readout. A biologically plausible local learning rule can compensate for drift, accurately decoding behavior over many days.

29. Duncker L, Sahani M: **Dynamics on the manifold: identifying computational dynamical activity from neural population recordings**. *Curr Opin Neurobiol* 2021, **70**:163–170.

30. Vyas S, Golub MD, Sussillo D, *et al.*: **Computation through neural population dynamics**. *Annu Rev Neurosci* 2020, **43**:249–275.

31. Brette R: **Is coding a relevant metaphor for the brain?** *Behav Brain Sci* 2018, **42**:e215.

32. Freeman III WJ, Skarda CA: ***Representations: who needs them?***. Available at: https://escholarship.org/uc/item/0gq5z3sh. Accessed 14 May 2022. accessed.

33. Welle CG, Gao Y-R, Ye M, *et al.*: **Longitudinal neural and vascular structural dynamics produced by chronic microelectrode implantation**. *Biomaterials* 2020, **238**, 119831.

34. Hammer DX, Lozzi A, Abliz E, *et al.*: **Longitudinal vascular dynamics following cranial window and electrode implantation measured with speckle variance optical coherence angiography**. *Biomed Opt Express* 2014, **5**(8):2823–2836.

35. Yang Y, Liu N, He Y, *et al.*: **Improved calcium sensor GCaMP-X overcomes the calcium channel perturbations induced by the calmodulin in GCaMP**. *Nat Commun* 2018, **9**:1504.

36. Sadeh S, Clopath C: **Contribution of behavioural variability to representational drift**. *bioRxiv* 2022, https://doi.org/10.1101/2022.01.02.474731.

37. Musall S, Kaufman MT, Juavinett AL, *et al.*: **Single-trial neural dynamics are dominated by richly varied movements**. *Nat Neurosci* 2019, **22**:1677–1686.

38. Parker PRL, Brown MA, Smear MC, *et al.*: **Movement-related signals in sensory areas: roles in natural behavior**. *Trends Neurosci* 2020, **43**:581–595.

39. Pereira TD, Shaevitz JW, Murthy M: **Quantifying behavior to understand the brain**. *Nat Neurosci* 2020, **23**:1537–1549.

40. Steinmetz NA, Zatka-Haas P, Carandini M, *et al.*: **Distributed coding of choice, action and engagement across the mouse brain**. *Nature* 2019, **576**:266–273.

41. Stringer C, Pachitariu M, Steinmetz N, *et al.*: **Spontaneous behaviors drive multidimensional, brainwide activity**. *Science* 2019, **364**:255.

42. Goaillard J-M, Marder E: **Ion Channel degeneracy, variability, and covariation in neuron and circuit resilience**. *Annu Rev Neurosci* 2021, **44**:335–357.
43. Attardo A, Fitzgerald JE, Schnitzer MJ: **Impermanence of dendritic spines in live adult CA1 hippocampus**. *Nature* 2015, **523**: 592–596.
44. Holtmaat A, Svoboda K: **Experience-dependent structural synaptic plasticity in the mammalian brain**. *Nat Rev Neurosci* 2009, **10**:647–658.
45. Dörrbaum AR, Kochen L, Langer JD, *et al.*: **Local and global influences on protein turnover in neurons and glia**. *Elife* 2018, **7**, https://doi.org/10.7554/eLife.34202.
46. Gray NW, Weimer RM, Bureau I, *et al.*: **Rapid redistribution of synaptic PSD-95 in the neocortex in vivo**. *PLoS Biol* 2006, **4**, e370.
47. Spalding KL, Bergmann O, Alkass K, *et al.*: **Dynamics of hippocampal neurogenesis in adult humans**. *Cell* 2013, **153**: 1219–1227.
48. Marder E, Goaillard J-M: **Variability, compensation and homeostasis in neuron and network function**. *Nat Rev Neurosci* 2006, **7**:563–574.
49. Shimizu G, Yoshida K, Kasai H, *et al.*: **Computational roles of intrinsic synaptic dynamics**. *Curr Opin Neurobiol* 2021, **70**: 34–42.
50. Gao P, Ganguli S: **On simplicity and complexity in the brave new world of large-scale neuroscience**. *Curr Opin Neurobiol* 2015, **32**:148–155.
51. Mastrogiuseppe F, Ostojic S: **Linking connectivity, dynamics, and computations in low-rank recurrent neural networks**. *Neuron* 2018, **99**:609–623. e29.
52. Xia J, Marks TD, Goard MJ, *et al.*: **Stable representation of a naturalistic movie emerges from episodic activity with gain variability**. *Nat Commun* 2021, **12**:5170.
53. Minderer M, Brown KD, Harvey CD: **The spatial structure of neural encoding in mouse posterior cortex during navigation**. *Neuron* 2019, **102**:232–248. e11.
54. Rule ME, O'Leary T: **Self-healing codes: how stable neural populations can track continually reconfiguring neural representations**. *Proc Natl Acad Sci USA* 2022, **119**, https://doi.org/10.1073/pnas.2106692119.
55. Aitchison L, Jegminat J, Menendez JA, *et al.*: **Synaptic plasticity as Bayesian inference**. *Nat Neurosci* 2021, **24**:565–571.
56. Kappel D, Habenschuss S, Legenstein R, *et al.*: **Network plasticity as Bayesian inference**. *PLoS Comput Biol* 2015, **11**, e1004485.
57. Kappel D, Legenstein R, Habenschuss S, *et al.*: **A dynamic connectome supports the emergence of stable computational function of neural circuits through reward-based learning**. *eNeuro* 2018, **5**, https://doi.org/10.1523/ENEURO.0301-17.2018.
58. Llera-Montero M, Sacramento J, Costa RP: **Computational roles of plastic probabilistic synapses**. *Curr Opin Neurobiol* 2019, **54**:90–97.
59. * Qin S, Farashahi S, Lipshutz D, *et al.*: **Coordinated drift of receptive fields during noisy representation learning**. *bioRxiv* 2021, https://doi.org/10.1101/2021.08.30.458264.

A biologically plausible Hebbian/anti-Hebbian network model is implemented to examine the hypothesis that noisy synaptic updates drive the network to explore a degenerate solution space, causing representational drift. Consistent with neural data, similarity of population codes are stable over time.

60. Susman L, Brenner N, Barak O: **Stable memory with unstable synapses**. *Nat Commun* 2019, **10**:4441.
61. Aitken K, Garrett M, Olsen S, *et al.*: **The geometry of representational drift in natural and artificial neural networks**. *bioRxiv* 2021, https://doi.org/10.1101/2021.12.13.472494.
62. Srivastava N, Hinton G, Krizhevsky A, *et al.*: **Dropout: a simple way to prevent neural networks from overfitting**. *J Mach Learn Res: JMLR* 2014, **15**:1929–1958.
63. Carpenter GA, Grossberg S. *Pattern recognition by self-organizing neural networks*. (MIT Press). A Bradford Book; 1991.
64. Fusi S, Abbott LF: **Limits on the memory storage capacity of bounded synapses**. *Nat Neurosci* 2007, **10**:485–493.
65. Marr D: **Simple memory: a theory for archicortex**. *Phil Trans Roy Soc Lond B Biol Sci* 1971, **262**:23–81.
66. Gluck MA, Myers CE: **Hippocampal mediation of stimulus representation: a computational theory**. *Hippocampus* 1993, **3**: 491–516.
67. McClelland JL, Goddard NH: **Considerations arising from a complementary learning systems perspective on hippocampus and neocortex**. *Hippocampus* 1996, **6**:654–665.
68. O'Reilly RC, McClelland JL: **Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off**. *Hippocampus* 1994, **4**:661–682.
69. Chen L, Cummings KA, Mau W, *et al.*: **The role of intrinsic excitability in the evolution of memory: significance in memory allocation, consolidation, and updating**. *Neurobiol Learn Mem* 2020, **173**, 107266.
70. * Mau W, Hasselmo ME, Cai DJ: **The brain in motion: how ensemble fluidity drives memory-updating and flexibility**. *Elife* 2020, **9**, https://doi.org/10.7554/eLife.63550.

A recent and thorough review on the fluidity of memory engrams.

71. Rogerson T, Cai DJ, Frank A, *et al.*: **Synaptic tagging during memory allocation**. *Nat Rev Neurosci* 2014, **15**:157–169.
72. Zhou Y, Won J, Karlsson MG, *et al.*: **CREB regulates excitability and the allocation of memory to subsets of neurons in the amygdala**. *Nat Neurosci* 2009, **12**:1438–1443.
73. McCloskey M, Cohen NJ: **Catastrophic interference in connectionist networks: the sequential learning problem**. In *Psychology of learning and motivation*. Edited by Bower GH, Academic Press; 1989:109–165.
74. Duncker L, Driscoll L, Shenoy KV, *et al.*: **Organizing recurrent network dynamics by task-computation to enable continual learning**. *Adv Neural Inf Process Syst* 2020, **33**:14387–14397.
75. Masse NY, Grant GD, Freedman DJ: **Alleviating catastrophic forgetting using context-dependent gating and synaptic stabilization**. *Proc Natl Acad Sci USA* 2018, **115**:E10467–E10475.
76. Cai DJ, Aharoni D, Shuman T, *et al.*: **A shared neural ensemble links distinct contextual memories encoded close in time**. *Nature* 2016, **534**:115–118.
77. Josselyn SA, Frankland PW: **Memory allocation: mechanisms and function**. *Annu Rev Neurosci* 2018, **41**:389–413.
78. Rashid AJ, Yan C, Mercaldo V, *et al.*: **Competition between engrams influences fear memory formation and recall**. *Science* 2016, **353**:383–387.
79. Yokose J, Okubo-Suzuki R, Nomoto M, *et al.*: **Overlapping memory trace indispensable for linking, but not recalling, individual memories**. *Science* 2017, **355**:398–403.
80. Kinsky NR, Mau W, Sullivan DW, *et al.*: **Trajectory-modulated hippocampal neurons persist throughout memory-guided navigation**. *Nat Commun* 2020, **11**:2443.
81. Muzzio IA, Levita L, Kulkarni J, *et al.*: **Attention enhances the retrieval and stability of visuospatial and olfactory representations in the dorsal hippocampus**. *PLoS Biol* 2009, **7**, e1000140.
82. Garthe A, Roeder I, Kempermann G: **Mice in an enriched environment learn more flexibly because of adult hippocampal neurogenesis**. *Hippocampus* 2016, **26**:261–271.
83. Golub MD, Sadtler PT, Oby ER, *et al.*: **Learning by neural reassociation**. *Nat Neurosci* 2018, **21**:607–616.

84. Orsborn AL, Moorman HG, Overduin SA, *et al.*: **Closed-loop decoder adaptation shapes neural plasticity for skillful neuroprosthetic control**. *Neuron* 2014, **82**:1380–1393.

85. Rickgauer JP, Deisseroth K, Tank DW: **Simultaneous cellular-resolution optical perturbation and imaging of place cell firing fields**. *Nat Neurosci* 2014, **17**:1816–1824.

86. Rickgauer JP, Tank DW: **Two-photon excitation of channelrhodopsin-2 at saturation**. *Proc Natl Acad Sci USA* 2009, **106**:15025–15030.

87. Diamantaki M, Coletta S, Nasr K, *et al.*: **Manipulating hippocampal place cell activity by single-cell stimulation in freely moving mice**. *Cell Rep* 2018, **23**:32–38.

88. Lee D, Lin B-J, Lee AK: **Hippocampal place fields emerge upon single-cell manipulation of excitability during behavior**. *Science* 2012, **337**:849–853.

89. Josselyn SA, Tonegawa S: **Memory engrams: recalling the past and imagining the future**. *Science* 2020, **367**, https://doi.org/10.1126/science.aaw4325.

90. Sweis BM, Mau W, Rabinowitz S, *et al.*: **Dynamic and heterogeneous neural ensembles contribute to a memory engram**. *Curr Opin Neurobiol* 2021, **67**:199–206.