Published in final edited form as:

Curr Biol. 2024 April 08; 34(7): 1519–1531.e4. doi:10.1016/j.cub.2024.03.003.

Learning leaves a memory trace in motor cortex

Darby M. Losey[1,2,3], **Jay A. Hennig**[1,2,3,13], **Emily R. Oby**[2,7,13], **Matthew D. Golub**[2,4,5,12], **Patrick T. Sadtler**[2,7], **Kristin M. Quick**[2,7], **Stephen I. Ryu**[5,8], **Elizabeth C. Tyler-Kabara**[2,9,10,11], **Aaron P. Batista**[2,7,13,*], **Byron M. Yu**[1,2,4,6,13,*], **Steven M. Chase**[1,2,6,13,14,*]

[1]Neuroscience Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA

[2]Center for the Neural Basis of Cognition, Pittsburgh, PA 15213, USA

[3]Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA

[4]Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

[5]Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA

[6]Department of Biomedical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

[7]Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA 15213, USA

[8]Department of Neurosurgery, Palo Alto Medical Foundation, Palo Alto, CA 94301, USA

[9]Department of Physical Medicine and Rehabilitation, University of Pittsburgh, Pittsburgh, PA 15213, USA

[10]Department of Neurological Surgery, University of Pittsburgh, Pittsburgh, PA 15213, USA

[11]Department of Neurosurgery, Dell Medical School, University of Texas at Austin, Austin, TX 78712, USA

[12]Paul G. Allen School of Computer Science & Engineering, University of Washington, Seattle, WA 98195, USA

[13]These authors contributed equally

[14]Lead contact

SUMMARY

[*]Correspondence: aaron.batista@pitt.edu (A.P.B.), byronyu@cmu.edu (B.M.Y.), schase@andrew.cmu.edu (S.M.C.).

AUTHOR CONTRIBUTIONS

Conceptualization & validation: D.M.L., J.A.H., E.R.O., M.D.G., A.P.B., B.M.Y., and S.M.C.; data curation & investigation: E.R.O., P.T.S., and K.M.Q.; formal analysis & methodology: D.M.L. and J.A.H.; funding acquisition, project administration, and supervision: A.P.B., B.M.Y., and S.M.C.; resources (surgical assistance): E.R.O., S.I.R., and E.C.T.-K.; resources (software): M.D.G.; writing – original draft: D.M.L.; writing – reviewing and editing: D.M.L., J.A.H., E.R.O., M.D.G., P.T.S., K.M.Q., S.I.R., E.C.T.-K., A.P.B., B.M.Y., and S.M.C.

SUPPLEMENTAL INFORMATION

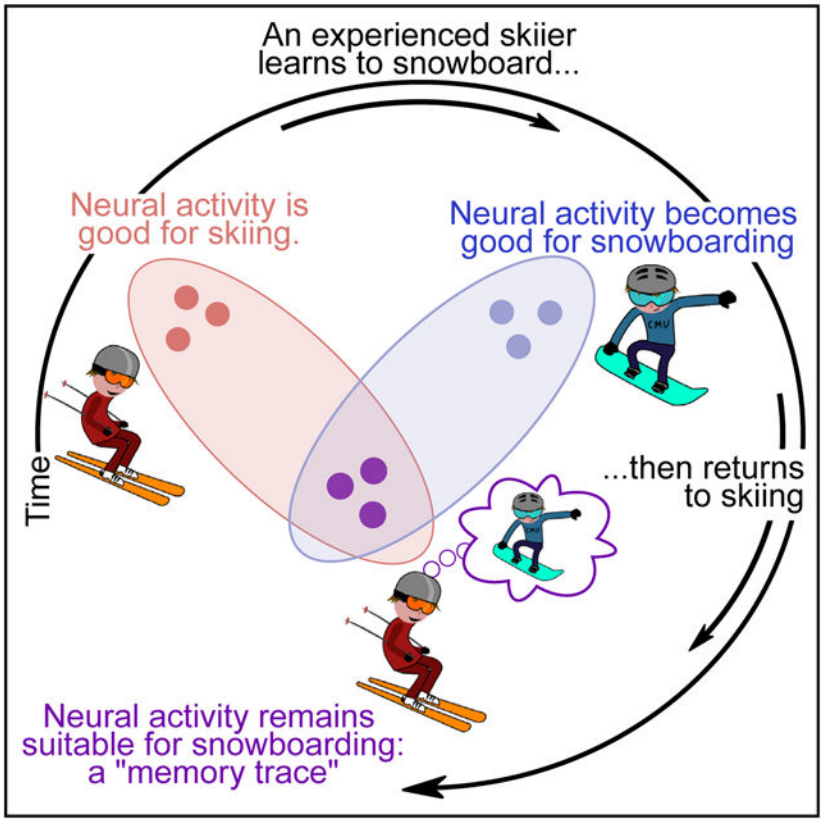
Supplemental information can be found online at https://doi.org/10.1016/j.cub.2024.03.003.

DECLARATION OF INTERESTS

The authors declare no competing interests.

How are we able to learn new behaviors without disrupting previously learned ones? To understand how the brain achieves this, we used a brain-computer interface (BCI) learning paradigm, which enables us to detect the presence of a memory of one behavior while performing another. We found that learning to use a new BCI map altered the neural activity that monkeys produced when they returned to using a familiar BCI map in a way that was specific to the learning experience. That is, learning left a “memory trace” in the primary motor cortex. This memory trace coexisted with proficient performance under the familiar map, primarily by altering neural activity in dimensions that did not impact behavior. Forming memory traces might be how the brain is able to provide for the joint learning of multiple behaviors without interference.

Graphical Abstract



In brief

How new tasks can be learned without interfering with old knowledge is unclear. Using a brain-computer interface, Losey et al. find that learning something new alters the neural activity used to perform a familiar task, such that neural activity remains appropriate for the new task but does not impede performance on the familiar task.

INTRODUCTION

How can the brain store multiple memories without interference? For example, suppose an experienced skier learns to snowboard. Skiing and snowboarding require different sets of

muscle activations, driven by different neural population activity patterns, to achieve the same goal of getting down the mountain without falling. How is knowledge about how to snowboard incorporated without overwriting the ability to ski? An intriguing possibility is that the memory of the recently learned task leaves a "memory trace": an alteration of neural activity that allows the brain to simultaneously support the memory of the newly learned task and the performance of the familiar task. Here, we distinguish a memory trace from the memory itself, in that the memory trace is specifically observable in the altered firing of neural populations and may be observed in more areas than those in which the memory is stored.

Motor learning alters neural activity in multiple brain regions, including the cerebellum,[1] hippocampus,[2] spinal cord,[3] basal ganglia,[4] and motor cortical areas.[5–10] New learning can be retained for hours,[11] days,[12] and even decades,[13] without major disruption from the performance of other skills. Motor learning retention has been observed across various behaviors, including visuomotor adaptation,[14] finger dexterity,[15] non-intuitive sensory-motor mappings,[16–18] and rhythmic tasks.[13] Beyond motor learning, the brain integrates new learning alongside existing memories during rule learning,[19–23] perceptual learning,[24–26] fear conditioning,[27,28] and more.

How might the integration of new learning proceed without impacting the performance of already-familiar behaviors? Consider how neural population activity might change when our experienced skier goes skiing, then learns to snowboard, and then returns to skiing (Figure 1A). One possibility is that the neural activity used for skiing remains unchanged after learning to snowboard. In this scenario, the new neural activity for snowboarding would only be recalled when snowboarding again. Such context-dependent recall has been observed in certain learning settings, such as the remapping of hippocampal place-fields between environments,[29] and has been proposed as a potential mechanism for motor memory storage.[30] A second possibility is that the neural activity used for skiing is actually altered by the recently acquired ability to snowboard. Several studies have suggested this to be the case, as neural tuning has been observed to change after motor adaptation.[5,31–34] What is unclear from the previous studies is whether these changes relate directly to the learned behavior. An intriguing possibility is that they constitute a memory of the learning experience. That is, learning could lead to a memory trace, which we define as an alteration of the population activity patterns used to perform familiar tasks in a manner that renders them also appropriate for a newly learned task (Figure 1B). However, it is also possible that the neural changes that accompany learning could be due to any one of the many task-agnostic factors that can influence neural activity in the motor system, such as changes in arousal,[35,36] motivation,[37] posture,[38] altered arm dynamics,[32,33] or learning-related changes that do not constitute a memory.

To assess if any changes in neural activity after learning are directly related to the newly learned ability, we need a way to determine the suitability of neural population activity for a given behavioral task. This is challenging in experiments using arm movements because the causal relationship between neural activity and behavior is not typically known. To address this challenge, we utilized a brain-computer interface (BCI) paradigm (Figure 1C;[10,39–45]). In a BCI, we specify the causal mapping between the recorded neural activity and behavior

(in this case, the movement of a computer cursor). We can then assess the suitability of neural activity for a BCI task that is not being performed, which enables us to detect a memory trace, if it is present. We used two different BCI maps in each experimental session (Figure 1D). Much like the example of an experienced skier learning to snowboard, a monkey first controlled a computer cursor using a familiar map (familiar task 1) and then learned how to use a new map (new task). Following learning, we reinstated the familiar map (familiar task 2). It is during familiar task 2 that we can assess whether or not a memory trace is formed by examining whether that neural activity is suitable for the recently learned new map.

We found that learning leaves a memory trace in the primary motor cortex (M1). That is, after the monkey learned to control the cursor using the new map, the neural activity that the monkey produced to control the cursor under the familiar map reflected the learned experience by becoming more suitable for the new map than it was prior to learning. Furthermore, we found that this memory trace coexisted alongside proficient familiar map performance by altering neural activity in a manner that did not interfere with the subsequent behavior. We speculate that the formation of memory traces may allow for the learning of multiple motor skills without interference and enable the rapid relearning of motor skills that characterizes motor savings.[46]

RESULTS

Here, we assess the formation of a memory trace by studying how learning to perform a new task affects the neural activity produced for a familiar task. We trained three monkeys to perform an eight-target center-out task using a BCI. The monkey's goal on each trial was to guide a computer cursor to an instructed target by modulating its neural activity (see STAR Methods). Notably, the arm of the monkey did not move throughout the experiment.[43] Only the recorded neural activity dictated cursor movement. Our BCI maps involved a two-step procedure. First, at each time step (45 ms), the activity of 90 neural units in primary M1 was projected into a ten-dimensional (10D) space that captured the majority of the variance shared among the neural units. Second, the activity of these 10 dimensions was used to drive a Kalman filter that determined the cursor's 2D velocity. Each experiment utilized two different BCI maps, the "familiar map" and the "new map", presented across three blocks of trials.

During the first block ("familiar task 1"), the monkey used the familiar map, which allowed for proficient cursor control without any learning (i.e., it was intuitive for the monkey to use; see Figure 1D). For the second block ("new task"), we changed the BCI map to the new map, which the monkey had never used before (see STAR Methods). This resulted in an initial decrement in the monkey's performance, which improved over the course of several hundred trials as he learned to control the cursor. The new map was selected to be a "within-manifold perturbation" (WMP), which we previously showed was well-learned within a single-day session.[43,47] A WMP changes how each of the 10 latent dimensions of neural activity influences the cursor velocity (see Figure S1 for details). Notably, the perturbation is applied in the 10D latent space rather than the animal's 2D workspace and thus can result in different behavioral impacts for each of the 8 targets in the workspace. In

the third block ("familiar task 2"), we reinstated the familiar map. This typically resulted in the well-known aftereffect that follows a bout of motor learning, after which performance returned to a level comparable to that of familiar task 1.[48] Data from the periods called familiar task 1 and new task have been examined in our prior work.[36,43,47,49] In this study, we now focus on the neural activity recorded during familiar task 2 to look for a memory trace.

Although the monkey uses the familiar map to control the computer cursor, we can evaluate how appropriate that same neural activity is for the new map (Figure 2A). This is the key advantage of a BCI that enables us to probe for the existence of a memory trace. Many different population activity patterns can be equally suitable for the familiar map (Figure 2B;[49,50]), due to neural redundancy (see Figure S2). Among these sets of redundant population activity patterns, some may be better for the new map than others. We will leverage the ability to evaluate neural activity through the offline map to test for the presence of a memory trace.

Our central question is: how does learning the new map affect the neural activity produced while using the familiar map? We consider three possibilities for what neural activity might look like after behavior stabilizes during familiar task 2. One possibility is that, after learning, the population activity patterns produced during familiar task 2 are similar to those produced during familiar task 1 (reversion, Figure 2C). Reversion has been observed in various contexts, such as reaching tasks,[32,33] reaching in conjunction with BCI learning,[51] BCI tasks in visual cortex,[45] and in the remapping of hippocampal place-fields.[29] This would indicate that the neural activity we observed in M1 during performance of a task can remain unaffected by an intervening learning experience.

A second possibility is that neural activity changes in a manner agnostic to the learning experience (representational drift, Figure 2D). Representational drift[52–56] could occur alongside proficient task performance due to many different activity patterns corresponding to the same behavioral output.[49,50] This drift could be attributed to any number of uncontrolled factors, such as arousal[35] or engagement.[36]

A third possibility is that neural activity differs between familiar task 2 and familiar task 1 in a manner that is directly related to having learned the new task ("memory trace," Figure 2E). We consider the possibility that neural activity changes to support the memory of the learned task while simultaneously supporting accurate cursor movement during familiar task 2. In this case, the neural activity produced during familiar task 2 would be more appropriate for the new map than that produced during familiar task 1.

We start by considering the reversion hypothesis. We asked whether the same population activity patterns were used during both familiar task 1 and familiar task 2. We observed that, for many targets, neural activity during familiar task 1 and familiar task 2 occupied different regions within the neural population space (Figure 3A), in contradiction to Figure 2C. Similarly, we observed changes in neuronal tuning between familiar task 1 and familiar task 2 (Figure S3), as seen in previous arm movement studies[5,31] (however, see other studies[33,57]). Thus, our data are not consistent with the reversion hypothesis (Figure 3B).

Next, we attempt to distinguish the memory trace hypothesis from the representational drift hypothesis. To do so, we must evaluate how the observed changes in neural activity relate to the previously learned behavior. Our BCI approach makes this possible. To illustrate this analysis, we compare neural activity from a single trial during each of familiar task 1 and familiar task 2 corresponding to the same target (Figure 4A, top). For the neural activity produced at each timestep, we can evaluate its "progress" through the familiar map as the extent to which it moves the cursor toward the target (see STAR Methods). Progress has units of cursor velocity (mm/s) and thus gives us a behavioral readout at the time resolution of a single timestep (45 ms). During both familiar task 1 and familiar task 2, the familiar map determines cursor velocity, and the monkeys showed proficient control of the cursor (i.e., large progress values) during both tasks (Figures 4B and S2E). Note that, during the new task, the new map determines cursor velocity, and learning can be observed by the straightening of cursor trajectories and the increase of progress values with practice (Figures S4A–S4C). These changes were not due to different neural subpopulations being employed for each map (Figure S4D), but rather they emerged when the entire neural population worked together in a new way.

Since we are using a BCI, progress can also be calculated for the new map even when the animal is using the familiar map to control the cursor. Progress under the new map measures the extent to which a given population activity pattern *would have moved* the cursor toward the target had the new map been instantiated. During familiar task 1, the monkeys exhibited low progress through the new map, as the velocities through the new map were small and haphazardly oriented relative to the target (Figure 4A, bottom, familiar task 1). This is expected because the monkey had not yet experienced the new map, and the new map was selected to be difficult to control using the familiar map's neural activity.[43] By contrast, during familiar task 2, the velocities through the new map were higher and more directed toward the target than they were during familiar task 1 (Figure 4A, bottom, familiar task 2); that is, they show higher progress (Figure 4C). This occurred despite the fact that the new map had no influence on behavior during familiar task 2, and thus the monkeys had no external incentive while performing familiar task 2 to maintain high progress through the new map. We define the memory trace as the average increase in the progress toward a given target when passing the recorded neural activity through the new map during familiar task 2, relative to familiar task 1. We found that the progress through the new map tended to be larger during familiar task 2 than familiar task 1, yielding a positive memory trace (Figures 4D and 4E; see also Figure S4F). This finding supports the memory trace hypothesis (Figure 2E), but not the representational drift hypothesis (Figure 2D), which would predict more variable new map progress across sessions rather than consistently higher new map progress.

In a minority of targets, the animal exhibited a memory trace value below zero, indicating that neural activity during familiar task 2 was less suitable for the new map than it was during familiar task 1. To understand why this might occur, we found that targets with a negative memory trace exhibited substantially less learning compared with those with a positive memory trace (monkey J, $P = 2.22 \times 10^{-10}$, two-sided unpaired Wilcoxon sign-rank test; monkey N, $P = 0{:}0048$; monkey L, $P = 0{:}00051$). When more learning occurred, the memory trace tended to be larger (Figure 4F). As monkeys J and N generally showed more

learning than monkey L, this could explain why the memory traces for monkeys J and N tended to be larger than that of monkey L (Figure 4F).

We considered whether recording instabilities could lead to an apparent memory trace. This is unlikely to be the case. As a first-order approximation, recording instabilities are likely to result in a global shift in neural activity. Even if this were to sometimes benefit performance on a particular target, due to the linearity of the BCI mappings, we would expect to see a decline in performance on the opposing target. Instead, we see positive memory traces when averaging across targets within each session (Figure 4E), indicating the changes in neural activity were learning specific.

We next assessed two properties that would make memories useful. The first property is that a memory should persist, meaning that it is present in neural activity without dissipating as time passes. To test whether this was true of the memory trace, we took the sessions with the most trials of familiar task 2 and examined whether the memory trace was still present at the end of the familiar task 2 period (Figure 5A). We found the memory trace was consistently present at the end of these longer familiar task 2 periods (Figure 5B). Furthermore, the memory trace remained positive for the duration of most experimental sessions (Figure S5A).

The second property of a memory is that it should coexist alongside proficient performance of other tasks. We define “coexist” here to indicate that the monkey is able to produce neural activity that simultaneously achieves high performance through both maps. To assess this, we examined whether the size of the memory trace was contingent on how proficient the behavior was during familiar task 2 (Figure 5C). Behavior during familiar task 2 was variable and typically a little worse than during familiar task 1 (Figure S2E). This is not surprising, considering these long experiments might lead to both mental fatigue and also satiation. We can harness this variability to assess whether there might be a trade-off between the strength of the memory trace and familiar task performance. If the instances with worse behavioral performance during familiar task 2 had the largest memory trace, it could suggest that the memory trace arises due to a trade-off in performance through the two BCI maps. Alternatively, if the memory trace were present even when behavioral performance during familiar task 2 returned to the levels seen during familiar task 1, it would suggest that the memory trace can coexist without hindering the monkey’s ability to perform the familiar task. We found that targets with the best behavioral performance during familiar task 2 showed a memory trace that was as strong as (or stronger than) that of targets with worse behavioral performance during familiar task 2 (Figures 5D, S5B, and S5C). These results indicate the memory trace coexists alongside proficient behavioral performance of the familiar task and does not represent a compromise between the two learned behaviors.

How can a memory trace coexist without degrading behavior during the re-performance of a familiar task? To understand this, we considered how the changes in neural activity induced by learning the new map relate to the familiar map. Because there are more dimensions of neural activity than there are of cursor movement, not all changes in neural activity affect cursor movement. We refer to changes in neural activity that affect cursor

movement as “output-potent” with respect to that map, and changes that do not as “output-null.[50]“ Because the familiar map and the new map do not share the same output-potent space, it is possible to have neural changes that affect cursor movement through one map without impacting cursor movements through the other map.

We examined whether the memory trace of the new map (Figure 6A) resides in the output-potent or output-null space of the familiar map (Figure 6B) by decomposing the memory trace into its output-potent and output-null components (Figure 6C). We found that the memory trace resides predominantly in the output-null space of the familiar map (Figures 6D and 6E). This means the memory trace resides primarily in dimensions that do not influence task performance under the familiar map (Figure S6A). Furthermore, the size of the memory trace could not be explained by the angle between the familiar map and the new map (Figure S6B).

Lastly, we asked, how does the monkey arrive at the familiar task 2 solution? There are two possibilities. The first possibility is that there is a partial “unwinding” of the learning that occurred during the new task. This would suggest that the solution used during familiar task 2 is not novel and was employed sometime during the learning experience. If this were true, we would expect that the path neural activity takes from the end of the new task to the end of the familiar task 2 (i.e., “the path of washout,” red arrow in Figure 7A) would retrace the path that neural activity takes from the end of familiar task 1 to the end of the new task (i.e., “the path of learning,” blue arrow in Figure 7A). The other possibility is that the path of washout is distinct from the path of learning (Figure 7B). This would imply that the solution the monkey uses during familiar task 2 is novel. To differentiate between these possibilities, we calculated the angle formed between the path of learning and the path of washout (see STAR Methods). We found that the path of washout is distinct from the path of learning (Figure 7C). Thus, relearning the familiar task is not simply the forgetting of learning the new task. In other words, the neural activity does not unwind during washout, but rather a novel solution to the task is found.

DISCUSSION

We studied how the brain can retain a memory of a newly learned motor task without compromising the performance of familiar tasks. We considered that learning may leave a memory trace observable in M1 population activity, such that neural activity remains appropriate for the learned task after the animal resumes performing a familiar task. A BCI enables new insight into the longstanding question of the joint consolidation of multiple skills. This is because using a BCI allows us to assess the extent to which the same neural activity is suitable for a task that is currently being performed and another task that is not actively being performed. We found that the neural activity produced while using a familiar map after learning a new map was better for the new map compared with before the learning experience. This memory trace of the learned map resided primarily in dimensions of population activity space which were output-null to the familiar map. In this way, neural activity simultaneously supported memory of the recently learned map without compromising behavioral performance through the familiar map.

What is the utility of maintaining a memory trace in neural population activity? A memory trace could enable proficient performance to be reached more quickly upon re-exposure to the learned task. This phenomenon, known as savings, has been frequently observed in motor learning behavior and is often taken as evidence that a memory formed.[30,58] Savings could be observed in two ways: performance could either start off immediately better upon re-exposure, or performance could show an improved rate of improvement on re-exposure. Our finding that neural activity remains more appropriate for the new map during familiar task 2 could enable savings through the first mechanism, allowing performance to start from a better position. Our results do not speak to whether there would also be savings in the form of an increased rate of improvement during re-exposure.

It is not a given that learning would leave a trace detectable in M1. For example, one could imagine that motor memories might be stored through synaptic weight changes in any number of brain areas, and only when a pattern of errors is experienced might they lead to a context-dependent recall of the appropriate action.[30,45] Instead, we found that a memory trace was evident in M1 during the performance of other actions. Our results do not rule out that context-dependent recall might be operating in parallel with the memory trace we observe. Similarly, our results do not speak to whether a memory trace would also be observable in brain areas outside of M1.

Motor memory consolidation refers to, among other things,[59] the process through which motor memories can become less susceptible to interference over time.[12,58] This consolidation process can take several hours to complete,[11] may necessitate continued practice,[60,61] and has been suggested to involve M1.[62–64] How might the brain bridge from the short-timescale retention of a memory trace that we studied here to the longer-timescale consolidation of a motor memory?[11,44,65] Our results focused on the short-term inception of a motor memory within an hour or so of the learned experience. Three possibilities would be consistent with our results. First, a long-term consolidated memory might resemble the memory trace we observed here. Second, it might be that the memory trace we observed is only a short-term phenomenon in M1, dissipating after consolidation. Finally, it could be that with further practice with both maps over many days, the neural activity evolves[66,67] to lead to even greater coexistence between the two behaviors.[68,69]

Many neurons contributed to the memory trace (Figure S4D). This coding scheme contrasts with the hippocampus, where a sparse subset of neurons can encode the memory.[27] We observed that the memory trace was primarily due to changes in neural activity orthogonal (i.e., output-null) to the familiar task. Notably, the utilization of different subsets of neurons in hippocampus to encode memories is a special case of orthogonal representations in population activity space.[29] Recent studies have proposed that the hippocampus,[70,71] auditory cortex,[72] and prefrontal cortex[73,74] also use orthogonal subspaces to incorporate multiple memories without interference. Avoiding interference may be harder in the spinal cord, where there are fewer neurons than in cortex. As fewer neurons likely lead to a more constrained encoding space, a “negotiated equilibrium” between multiple learned behaviors may be required in the spinal cord.[75]

Skill acquisition and adaptation are distinct aspects of motor learning. Skill acquisition focuses on gaining a new ability by developing and refining complex motor actions and techniques, often incorporating multiple sensory inputs and motor outputs. It generally necessitates a longer time for learning, consolidation, and retention.[15] Motor adaptation, by contrast, involves quickly adjusting motor actions to compensate for environmental changes, such as recalibrating sensory-motor mappings[76] or adapting to altered dynamical environments.[48] Both types of learning lead to memory formation, though they may have different mechanisms of doing so. Learning WMPs, as examined in this study, may harness neural mechanisms that are more similar to those used during adaptation learning than skill learning,[10,43] though precisely how BCI learning relates to the learning of reaching movements is still an open question.[77]

Our finding of a memory trace in BCI learning may reveal a general phenomenon that is also present for motor learning with the arm and hands. In classic studies,[5,31] the activity of individual neurons in M1 was shown to change with motor learning. These changes were hypothesized to be consistent with a mechanism akin to the memory trace mechanism we have shown here.[68] More recently, Sun and colleagues also observed systematic changes in neural activity related to the learning experience.[34] Although learning an arm-reaching task in a curl force field, animals exhibited a “uniform shift” in preparatory neural activity that persisted after the force field was removed. The authors conjecture that this shift indexes motor memories.[78] It is conceivable that the memory trace shares similar properties to the uniform shift observed in Sun et al.,[34] as both are shifts that occur during learning that remain in population activity during washout. One key advance of our findings is that we show that this shift is beneficial for performing the learned task. It is intriguing to speculate whether the uniform shifts observed by Sun et al. modify preparatory activity to become more appropriate for the learned task and could, in that sense, constitute a memory trace.

There is mounting evidence that learning and control of a BCI employ similar underlying neural mechanisms as during arm movements.[77,79] For instance, the formation of internal (forward) models, which involve predicting and compensating for the sensory feedback of a motor command in subsequent commands, has been observed in both BCI experiments[80] and arm movement tasks.[81] Furthermore, studies have found that learning in BCI contexts can facilitate learning in arm movement tasks,[51] suggesting that the two tasks share common neural substrates. BCI learning has also been shown to engage subcortical areas, such as the striatum,[41] which is known to be involved in motor learning and control. Our results using BCI provide empirical evidence supporting theoretical results of sensorimotor learning[68] and align with similar principles in the spinal cord.[75]

Human and animal learners distinguish themselves from current artificial learning systems in that they can learn to perform a large number of different behaviors and flexibly switch among them. It is a notoriously challenging problem for artificial agents to learn new tasks without overwriting the ability to perform previously learned tasks, an effect termed “catastrophic forgetting.”[82–85] Our findings suggest that artificial learning systems could overcome catastrophic forgetting by implementing some of the same learning principles employed by biological learning systems.[86,87] A sufficiently high dimensional activity space, utilized effectively for the storage of multiple memories without interference, may be

important not only in the brain but also for artificial agents learning multiple tasks without interference.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Steven Chase (schase@cmu.edu).

Materials availability—This study did not generate new unique reagents.

Data and code availability—The data that support the findings of this study are available from the lead contact upon reasonable request. All original code is publicly available as of the date of publication. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Three male Rhesus macaques (Macaca mulatta, ages 7, 7 and 8 for monkeys J, N and L, respectively) were implanted with 96 electrode arrays (Blackrock Microsystems) in the proximal arm region of the primary motor cortex. All animal care and handling procedures conformed to the NIH Guidelines for the Care And Use of Laboratory Animals and were approved by the University of Pittsburgh's Institutional Animal Care and Use Committee.

METHOD DETAILS

Experimental procedures—Experimental methods are detailed in our previous work.[43,47] Briefly, we recorded neural activity (RZ2 system, TDT, Inc.) from three male Rhesus macaques (Macaca mulatta, ages 7, 7 and 8 for monkeys J, N and L, respectively) using 96 electrode arrays (Blackrock Microsystems) implanted in the proximal arm region of the primary motor cortex. The monkeys performed an eight-target center-out BCI task. In the BCI, a monkey guided a computer cursor by modulating its neural activity. The recorded neural activity was translated into movements of the computer cursor according to a BCI map (see Translating neural activity to cursor movement). Each session was split into three task periods, "Familiar Task 1", "New Task", and "Familiar Task 2". The animals performed the same center-out BCI task in all three task periods. The only difference between task periods was the BCI map instantiated by the experimenter. During Familiar Task 1, the monkey used the Familiar Map, which was selected to be intuitive for the monkey to use from the outset. This map was found through a calibration period at the beginning of the day to identify the natural covariation between neural activity and intended cursor velocity. Empirically, we find that this relationship changes little from day-to-day, and any changes appear to stem from neural recording instabilities.[88] The monkey controlled the cursor during Familiar Task 1 for 318.8 ± 95.4 (mean ± s.d.) trials.

Uncued to the monkey, we then switched to the New Map for the second task period (New Task). The monkey had never seen the New Map before and it was selected in order to initially be difficult for the monkey to use to control the cursor. The monkey was given

696.7 ± 219.4 (mean ± s.d.) trials to learn to control the cursor with the New Map. Finally, again uncued, the Familiar Map was reinstated (Familiar Task 2). The Familiar Task 2 period lasted the remainder of the experiment, 318.2 ± 153.9 (mean ± s.d.) trials.

Trial flow—At the start of each trial, the cursor appeared at the center of the monkey's workspace. Target locations were selected pseudo-randomly from a set of eight uniformly spaced locations around a circle (radius, Monkey J: 150 mm; Monkeys L and N: 125 mm). The target appeared on the screen at the beginning of the trial. For the first 300 ms, the cursor's velocity was fixed at zero. After this, the velocity of the cursor was controlled by the monkey through the BCI map corresponding to the task period of the experiment. If the monkey was able to acquire the target within 7.5s after the start of the trial, a water reward was dispersed. If the monkey failed to acquire the target within the allotted time, there was a 1.5s timeout prior to the start of the next trial.

Identifying latent dimensions of neural activity—Experiments began with a calibration period in order to define the Familiar Map. Monkey J's calibration employed either passive cursor observation or closed-loop BCI control using the previous day's BCI map. For monkeys L and N, we used a calibration procedure that gradually stepped from passive observation to closed-loop control. We then applied factor analysis (see below) to identify the 10D linear subspace (the "intrinsic manifold") that captured the dimensions of greatest shared variance in the neural population. Ten dimensions was selected using cross-validation, as described in prior work.[43]

Spike counts (i.e. threshold crossings) were taken in nonoverlapping 45 ms time windows. We denote the spike counts at timestep t as $\mathbf{u}_t \in \mathbb{R}^{q \times 1}$, where q is the number of neural units. Factor analysis describes this high-dimensional population activity, $\mathbf{u}_t$, in terms of a low-dimensional set of factors, $\mathbf{z}_t \in \mathbb{R}^{10 \times 1}$. Latent factors, $\mathbf{z}_t$, are distributed as:

$$\mathbf{z}_t \sim N(0, I) \quad \text{(Equation 1)}$$

where I is the identity matrix. Spike counts, $\mathbf{u}_t$, are related to the factors by:

$$\mathbf{u}_t | \mathbf{z}_t \sim N(L\mathbf{z}_t + \boldsymbol{\mu}, \Psi) \quad \text{(Equation 2)}$$

where parameters $L \in \mathbb{R}^{q \times 10}$ (termed the loading matrix), $\boldsymbol{\mu} \in \mathbb{R}^{q \times 1}$, and $\Psi \in \mathbb{R}^{q \times q}$ (a diagonal matrix of variances independent to each neuron) are estimated using the expectation-maximization algorithm. The latent factors at timestep t are estimated as the posterior expectation given the spike counts:

$$\hat{\mathbf{z}}_t = L^T (LL^T + \Psi)^{-1} (\mathbf{u}_t - \boldsymbol{\mu}) \quad \text{(Equation 3)}$$

For all analyses, we orthonormalized $\hat{\mathbf{z}}_t$ so that it had units of spike counts per timestep to facilitate the interpretability of the factor activity. As the majority of the shared variance of the neural population is captured in these latent dimensions, and neural activity cannot be readily produced outside this low-dimensional subspace during short-term learning,[10,43] we focus our analyses on this factor activity, referred to as "population activity patterns" throughout.

Translating neural activity to cursor movement—At each 45 ms timestep t, neural activity drove the computer cursor according to the BCI map for that task period. Specifically, the cursor velocity was determined using a Kalman filter:

$$\mathbf{v}_t = A\mathbf{v}_{t-1} + B\hat{\mathbf{z}}_t + \mathbf{c} \quad \text{(Equation 4)}$$

The parameters $A \in \mathbb{R}^{2\times 2}$, $B \in \mathbb{R}^{2\times 10}$ and $\mathbf{c} \in \mathbb{R}^{2\times 1}$ are determined during the calibration period,[43] and $\mathbf{v}_t \in \mathbb{R}^{2\times 1}$ comprises the horizontal and vertical cursor velocities. The two BCI maps differ only in the B term. For the Familiar Map, $B = B^{\text{familiar}}$, which is found during the calibration period. For the New Map, $B = B^{\text{new}}$ was a permutation applied to the columns of B^{familiar}, equivalent to permuting the elements of $\hat{\mathbf{z}}_t$ before applying Equation 4. This means that the New Map remained within the intrinsic manifold (a "within-manifold perturbation"). Thus the New Map changed the relationship between the factor activity and cursor velocity.

Full details of how the New Map was selected can be found in our previous work.[43,47] In short, there are 10! = 3,628,800 unique permutations that could be applied to a the Familiar Map to yield a the New Map, of which we selected just one per experiment. Our aim was to select a the New Map that was difficult enough to induce learning, but not so difficult as to discourage the animal from participating in the task. To inform this selection, we predicted the cursor velocities the animal would produce under each of the candidate the New Map's during the first 200 trials of Familiar Task 1 using:

$$\mathbf{v}_t^{(pred)} = B^{\text{new}}\hat{\mathbf{z}}_t + \mathbf{c} \quad \text{(Equation 5)}$$

where $\mathbf{v}_t^{(pred)}$ is the predicted velocity, and B^{new} and $\mathbf{c}$ are the candidate parameters of the New Map. We then compared this to the velocities produced by the same neural activity under the Familiar Map (see Figure S1). We used the difference in angle and speeds to eliminate candidate the New Maps that are deemed too difficult or not difficult enough.[47] In a typical experiment, approximately 50 candidate mappings satisfied all the requirements, and one was randomly selected for use in the experiment.

QUANTIFICATION AND STATISTICAL ANALYSIS

The data analyzed in this study was part of a larger study that included both within-manifold perturbations (WMPs) and outside-manifold perturbations (OMPs).[43] As we have previously

found that WMPs show stronger learning than OMPs, we only considered sessions that used WMPs. Data from the Familiar Task 1 and New Task periods of these WMP sessions were analyzed in prior work.[36,47,49] Here we focused on neural activity recorded during Familiar Task 2, which has not been previously studied. To ensure an adequate amount of Familiar Task 2 data to analyze per session, we only considered sessions that included at least 100 Familiar Task 2 trials. This yielded a total of 43 sessions (Monkey J, 22 sessions, 362.6 ± 170.2 Familiar Task 2 trials; Monkey N, 12 sessions, 333.3 ± 107.3 Familiar Task 2 trials; Monkey L, 9 sessions, 171.0 ± 49.7 Familiar Task 2 trials; all values mean +/− s.d.).

Selecting experiments and trials for analysis—As our central question focuses on neural activity during proficient Familiar Task 2 performance, we restricted analyses of Familiar Task 2 trials to after behavior had stabilized. To do this, we examined trials after at least 50 trials of Familiar Task 2 had elapsed (see Figure 5). Unless stated otherwise, the remaining Familiar Task 2 trials are referred to as Familiar Task 2 throughout the manuscript. Additionally, we only analyzed successful trials, as it is otherwise difficult to determine whether the monkey was engaged in the task. Across all trials in all blocks, the success rates for Monkeys J, N and L were 79.6% +/− 11.5%, 84.3% +/− 15.2% and 76.1% +/− 16.9% (mean ± s.d.), respectively.

On each trial, we discarded the first 90 ms (2 timesteps while the cursor's velocity was fixed at zero) as the activity in M1 would not yet reflect the target due to sensory processing delays.[80] Additionally, because we report trial-averaged and target-averaged quantities, we wanted to ensure neural activity came from instances in which the monkey needed to push the cursor in the same direction. Thus, we only analyzed timesteps in which the angle between the cursor and the target was no greater than 22:5° away from the target direction for that trial. Performing our analyses without this exclusion criterion did not change our results qualitatively.

Even after learning to use the New Map, the monkeys generally exhibited lower performance with the New Map than the Familiar Map (see Figure 1D). Thus, New Task trials tended to be longer than the Familiar Task 1 and Familiar Task 2 trials. To compare the New Task trials to the Familiar Task 1 and Familiar Task 2 trials, we only utilized the first 25 timesteps from each trial. This number was selected because it is approximately equal to the average Familiar Task 1 acquisition time across all monkeys.

Testing the reversion hypothesis—To measure tuning changes between task periods (Figure S3), we fit cosine tuning curves for each neural unit using ordinary least squares regression:

$$\gamma(\theta) = r_0 + (r_{\max} - r_0)\cos(\theta - \theta_{pd})$$

where $\lambda(\theta)$ is the estimated firing rate for a given cursor-target direction θ. The parameters θ_{pd}, r_0 and $r_{\max}$ can be interpreted as the preferred direction, the average firing rate, and the maximum firing rate of the unit, respectively. For each neural unit, we fit a separate tuning curve for each task period of the experiment.

We compared the preferred direction θ_{pd} for each neural unit between Familiar Task 1 and Familiar Task 2 by computing the average absolute change in preferred direction (Figure S3C). To calculate the control distribution, for each neural unit, we randomly permuted the task labels for each timestep during Familiar Task 1 and Familiar Task 2. The difference in preferred direction between Familiar Task 1 and Familiar Task 2 was then recalculated using these new task labels.

To visualize how neural activity changes in the 10D latent space, we applied linear discriminant analysis to $\hat{\mathbf{z}}_t$, taken in 45ms timesteps, in to order to find the 2D plane that best separates the activity from the three task periods (Figure 3A). Web applied a QR decomposition in order to orthonormalize the basis vectors found by LDA, then projected the neural activity onto this orthonormal basis.

To quantify the changes in population activity between Familiar Task 1 and Familiar Task 2, we calculated the Mahalanobis distance on a per-target basis between the population activity means across $\hat{\mathbf{z}}_t$, taken in 45ms timesteps, for each task period (Figure 3B). This distance was computed in the 10D latent space, using the covariance of the Familiar Task 1 neural activity for that target. To calculate the control distribution, for each target, we randomly permuted the task labels for each timestep during Familiar Task 1 and Familiar Task 2. The Mahalanobis distance between the mean activity for each target was recalculated using the new task labels.

Defining the memory trace—Progress quantifies the appropriateness of a particular population activity pattern for a particular BCI map, i.e., the extent to which that population activity pattern drives the cursor towards the target, and is computed as follows. First, we determine the neural push of this activity pattern, $\hat{\mathbf{z}}_t$, through a particular map, B, as $B\hat{\mathbf{z}}_t$. In Equation 4, A and $\mathbf{c}$ do not rely on the instantaneous neural activity, and so we do not consider the contributions from these terms. Next, we compute the component of this neural push in the direction of the target. More specifically, for each timestep t, we define a unit vector, $\mathbf{e}_t \in \mathbb{R}^{2\times 1}$, pointing from the current location of the cursor to the target. Thus, the progress at timestep t is evaluated as:

$$p_t = \mathbf{e}_t^T B\hat{\mathbf{z}}_t \quad \text{(Equation 6)}$$

We sought to determine how much more appropriate neural activity is for the New Map during Familiar Task 2 than it is during Familiar Task 1. We call this change in appropriateness a “memory trace” because it measures the lasting alteration of neural activity used during a familiar task (the Familiar Map) after a learning experience (the New Map). Specifically, we define the memory trace as the difference in progress when neural activity is passed through the New Map during Familiar Task 2 minus that during Familiar Task 1. For each target, we average the progress per timestep across all trials. We obtain similar results if we first average within a trial, then average across all trials to the given target.

Defining learning—We defined *learning* as how well the monkey performed with the New Map after learning, relative to how well it would have performed with the New Map if it continued producing the same neural activity as it did during Familiar Task 1 (i.e., if there was no learning). Thus, we defined learning as the difference in the New Map progress (see Defining the memory trace for how progress is computed) of the last 10 trials to a given target during the New Task minus the average the New Map progress of trials to that same target during Familiar Task 1.

Testing how Familiar Task 2 duration affects the memory trace—We sought to determine whether the memory trace persisted over time (Figures 5A and 5B). We considered the sessions in which the Familiar Task 2 period was at least as long as the median length across all sessions (300 trials). This resulted in 22 sessions (14/22 sessions from Monkey J, average length of 464.36 ± 119.76 Familiar Task 2 trials; 8/12 sessions from Monkey N, 400.00 ± 53.45, 0/9 sessions from Monkey L; all values are mean ± s.d.). In order to focus on trials where the monkey had longer exposure to Familiar Task 2, we excluded the first 200 trials when calculating the memory trace, leaving at least 100 trials of Familiar Task 2 for analysis.

Testing how Familiar Task 2 behavior affects the memory trace—We additionally sought to determine whether the memory trace differed as a function of performance through the Familiar Map (Figures 5C and 5D). To address this, we separated targets into two groups. Targets with acquisition times during Familiar Task 2 that were at least as good as Familiar Task 1 were placed in the “better behavior group” (see Figure S2E). There were 48 targets in this group, with an average of 75.0 ± 57.7 ms (mean ± s.d.) faster target acquisition in Familiar Task 2 relative to Familiar Task 1. Targets which had acquisition times during Familiar Task 2 that were worse than Familiar Task 1 were placed in the “worse behavior group”. There were 296 targets in this group, with an average of 241.3 ms ± 210.2 ms (mean ± s.d.) slower target acquisition in Familiar Task 2 relative to Familiar Task 1.

Decomposing the memory trace into output-potent and output-null components—In order to determine how the memory trace can coexist without degrading behavioral performance during Familiar Task 2, we wanted to determine how changes in neural activity between Familiar Task 1 and Familiar Task 2 relate to the Familiar Map. To address this question, we decomposed neural activity into a component that is output-potent to the Familiar Map and a component that is output-null to the Familiar Map (Figure 6). This decomposition was done by applying the singular value decomposition (SVD) to the Familiar Map:

$$B^{\text{familiar}} = UDV^T \quad \text{(Equation 7)}$$

where $U \in \mathbb{R}^{2\times 10} D \in \mathbb{R}^{10\times 10}$, and $V \in \mathbb{R}^{10\times 10}$. D is a diagonal matrix, whose diagonal elements are the singular values of B^{familiar}. As B^{familiar} is a matrix of rank two, only the first two diagonal entries of D are non-zero. This means that the first two columns of V form an orthonormal basis for the output-potent space of B^{familiar}. We denote this basis as

$R \in \mathbb{R}^{10\times 2}$.The last 8 columns of V form an orthonormal basis of the output-null space of B^{familiar}. We denote this basis as $N \in \mathbb{R}^{10\times 8}$.

We can find the component of neural activity potent to the Familiar Map as $\mathbf{z}_t^{\text{pot}} = RR^T\hat{\mathbf{z}}_t$. Similarly, the null component is found as $\mathbf{z}_t^{\text{null}} = NN^T\hat{\mathbf{z}}_t$. Both $\mathbf{z}_t^{\text{pot}}$ and $\mathbf{z}_t^{\text{null}}$ are 10×1 vectors, and have the property that $\hat{\mathbf{z}}_t = \mathbf{z}_t^{\text{pot}} + \mathbf{z}_t^{\text{null}}$. We calculate the potent and null component of the memory trace as before, except utilizing $\mathbf{z}_t^{\text{pot}}$ and $\mathbf{z}_t^{\text{null}}$ for $\hat{\mathbf{z}}_t$ respectively in Equation 4. This decomposition is utilized in Figures 6 and S6. Note that this decomposition is performed with respect to the Familiar Map and not with respect to the New Map. This is because, by definition, the memory trace must be in output-potent dimensions of the New Map, as those are the only dimensions that determine the cursor velocity through the New Map.

Path of learning and washout—To distinguish whether the path of washout retraces the path of learning (Figure 7), we first define the path of learning as the vector in 10D latent space from the mean activity during Familiar Task 1 to the mean activity during the late New Task period (see Selecting experiments and trials for analysis). We similarly define the path of washout as the 10D vector between the mean neural activity during late New Task and the mean activity during Familiar Task 2. We then compared the paths of learning and washout by finding the the angle between these two vectors. To obtain a control distribution, for each target, we randomly permuted the task labels for each timestep during Familiar Task 1 and Familiar Task 2. This mimics a situation in which Familiar Task 1 and Familiar Task 2 activity patterns come from the same distribution. As task labels for New Task were not shuffled, the paths of learning and washout would thus be equal and opposite on average under this construction. The angle between the paths for each target was recalculated using the new task labels.

Statistics—Unless otherwise noted, to test for statistical significance, we used nonparametric tests (for example, Wilcoxon signed-rank test or ranked-sum test), which do not assume normality. All *P*-values less than 10^{-10} were reported as $P < 10^{-10}$, regardless of how small the *P*-value was.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Caleb Kemere for helpful discussions. We thank Katrina P. Nguyen for providing artwork used in the graphical abstract. This work was supported by the National Science Foundation Graduate Research Fellowship DGE1745016 and DGE2140739 (to D.M.L.), the Richard King Mellon Presidential Fellowship (to J.A.H.), the Carnegie Prize Fellowship in Mind and Brain Sciences (to J.A.H.), NIH R01 HD071686 (to A.P.B., B.M.Y., and S.M.C.), NIH R01 NS129584 (to A.P.B., S.M.C., and B.M.Y.), NSF NCS BCS1533672 (to S.M.C., B.M.Y., and A.P.B.), NSF NCS DRL 2124066 and 2123911 (to B.M.Y., S.M.C., and A.P.B.), NSF CAREER award IOS1553252 (to S.M.C.), NIH CRCNS R01 NS105318 (to B.M.Y. and A.P.B.), NSF NCS BCS1734916 (to B.M.Y.), NIH CRCNS R01 MH118929 (to B.M.Y.), NIH R01 EB026953 (to B.M.Y.), Simons Foundation 543065 (to B.M.Y.), NIH R01 NS120579 (to A.P.B.), and NIH R01 HD090125 (to A.P.B.).

REFERENCES

1. Morton SM, and Bastian AJ (2006). Cerebellar contributions to loco-motor adaptations during splitbelt treadmill walking. J. Neurosci 26, 9107–9116. 10.1523/JNEUROSCI.2622-06.2006. [PubMed: 16957067]
2. Wise SP, and Murray EA (1999). Role of the hippocampal system in conditional motor learning: mapping antecedents to action. Hippocampus 9, 101–117. https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1098-1063(1999)9:2%3C101::AID-HIPO3%3E3.0.CO;2-L. [PubMed: 10226772]
3. Vahdat S, Lungu O, Cohen-Adad J, Marchand-Pauvert V, Benali H, and Doyon J (2015). Simultaneous brain–cervical cord fMRI reveals intrinsic spinal cord plasticity during motor sequence learning. PLOS Biol. 13, e1002186. 10.1371/journal.pbio.1002186. [PubMed: 26125597]
4. Graybiel AM, Aosaki T, Flaherty AW, and Kimura M (1994). The basal ganglia and adaptive motor control. Science 265, 1826–1831. 10.1126/science.8091209. [PubMed: 8091209]
5. Li CSR, Padoa-Schioppa C, and Bizzi E (2001). Neuronal correlates of motor performance and motor learning in the primary motor cortex of monkeys adapting to an external force field. Neuron 30, 593–607. 10.1016/s0896-6273(01)00301-4. [PubMed: 11395017]
6. Padoa-Schioppa C, Li CSR, and Bizzi E (2004). Neuronal activity in the supplementary motor area of monkeys adapting to a new dynamic environment. J. Neurophysiol 91, 449–473. 10.1152/jn.00876.2002. [PubMed: 12968016]
7. Komiyama T, Sato TR, O'Connor DH, Zhang YX, Huber D, Hooks BM, Gabitto M, and Svoboda K (2010). Learning-related fine-scale specificity imaged in motor cortex circuits of behaving mice. Nature 464, 1182–1186. 10.1038/nature08897. [PubMed: 20376005]
8. Mandelblat-Cerf Y, Novick I, Paz R, Link Y, Freeman S, and Vaadia E (2011). The neuronal basis of long-term sensorimotor learning. J. Neurosci 31, 300–313. 10.1523/JNEUROSCI.4055-10.2011. [PubMed: 21209216]
9. Perich MG, Gallego JA, and Miller LE (2018). A neural population mechanism for rapid learning. Neuron 100, 964–976.e7. 10.1016/j.neuron.2018.09.030. [PubMed: 30344047]
10. Oby ER, Golub MD, Hennig JA, Degenhart AD, Tyler-Kabara EC, Yu BM, Chase SM, and Batista AP (2019). New neural activity patterns emerge with long-term learning. Proc. Natl. Acad. Sci. USA 116, 15210–15215. 10.1073/pnas.1820296116. [PubMed: 31182595]
11. Shadmehr R, and Holcomb HH (1997). Neural correlates of motor memory consolidation. Science 277, 821–825. 10.1126/science.277.5327.821. [PubMed: 9242612]
12. Walker MP, Brakefield T, Hobson JA, and Stickgold R (2003). Dissociable stages of human memory consolidation and reconsolidation. Nature 425, 616–620. 10.1038/nature01930. [PubMed: 14534587]
13. Park SW, Dijkstra TM, and Sternad D (2013). Learning to never forget-time scales and specificity of long-term memory of a motor skill. Front. Comp. Neurosci 7, 111. 10.3389/fncom.2013.00111.
14. Shadmehr R, Smith MA, and Krakauer JW (2010). Error correction, sensory prediction, and adaptation in motor control. Annu. Rev. Neurosci 33, 89–108. 10.1146/annurev-neuro-060909-153135. [PubMed: 20367317]
15. Karni A, Meyer G, Jezzard P, Adams MM, Turner R, and Ungerleider LG (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. Nature 377, 155–158. 10.1038/377155a0. [PubMed: 7675082]
16. Mosier KM, Scheidt RA, Acosta S, and Mussa-Ivaldi FA (2005). Remapping hand movements in a novel geometrical environment. J. Neurophysiol 94, 4362–4372. 10.1152/jn.00380.2005. [PubMed: 16148276]
17. Berger DJ, Gentner R, Edmunds T, Pai DK, and d'Avella A (2013). Differences in adaptation rates after virtual surgeries provide direct evidence for modularity. J. Neurosci 33, 12384–12394. 10.1523/JNEUROSCI.0122-13.2013. [PubMed: 23884944]
18. Huberdeau DM, Krakauer JW, and Haith AM (2015). Dual-process decomposition in human sensorimotor adaptation. Curr. Opin. Neurobiol 33, 71–77. 10.1016/j.conb.2015.03.003. [PubMed: 25827272]
19. Wallis JD, Anderson KC, and Miller EK (2001). Single neurons in prefrontal cortex encode abstract rules. Nature 411, 953–956. 10.1038/35082081. [PubMed: 11418860]

20. Peyrache A, Khamassi M, Benchenane K, Wiener SI, and Battaglia FP (2009). Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. Nat. Neurosci 12, 919–926. 10.1038/nn.2337. [PubMed: 19483687]
21. Durstewitz D, Vittoz NM, Floresco SB, and Seamans JK (2010). Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. Neuron 66, 438–448. 10.1016/j.neuron.2010.03.029. [PubMed: 20471356]
22. Jeanne JM, Sharpee TO, and Gentner TQ (2013). Associative learning enhances population coding by inverting interneuronal correlation patterns. Neuron 78, 352–363. 10.1016/j.neuron.2013.02.023. [PubMed: 23622067]
23. Bartolo R, Saunders RC, Mitz AR, and Averbeck BB (2020). Dimensionality, information and learning in prefrontal cortex. PLoS Comp. Biol 16, e1007514. 10.1371/journal.pcbi.1007514.
24. Schoups A, Vogels R, Qian N, and Orban G (2001). Practising orientation identification improves orientation coding in V1 neurons. Nature 412, 549–553. 10.1038/35087601. [PubMed: 11484056]
25. Gu Y, Liu S, Fetsch CR, Yang Y, Fok S, Sunkara A, DeAngelis GC, and Angelaki DE (2011). Perceptual learning reduces interneuronal correlations in macaque visual cortex. Neuron 71, 750–761. 10.1016/j.neuron.2011.06.015. [PubMed: 21867889]
26. Poort J, Khan AG, Pachitariu M, Nemri A, Orsolic I, Krupic J, Bauza M, Sahani M, Keller GB, Mrsic-Flogel TD, and Hofer SB (2015). Learning enhances sensory and multiple non-sensory representations in primary visual cortex. Neuron 86, 1478–1490. 10.1016/j.neuron.2015.05.037. [PubMed: 26051421]
27. Josselyn SA, and Tonegawa S (2020). Memory engrams: Recalling the past and imagining the future. Science 367, eaaw4325. 10.1126/science.aaw4325. [PubMed: 31896692]
28. Wu CT, Haggerty D, Kemere C, and Ji D (2017). Hippocampal awake replay in fear memory retrieval. Nat. Neurosci 20, 571–580. 10.1038/nn.4507. [PubMed: 28218916]
29. Alme CB, Miao C, Jezek K, Treves A, Moser EI, and Moser MB (2014). Place cells in the hippocampus: eleven maps for eleven rooms. Proc. Natl. Acad. Sci. USA 111, 18428–18435. 10.1073/pnas.1421056111. [PubMed: 25489089]
30. Herzfeld DJ, Vaswani PA, Marko MK, and Shadmehr R (2014). A memory of errors in sensorimotor learning. Science 345, 1349–1353. 10.1126/science.1253138. [PubMed: 25123484]
31. Arce F, Novick I, Mandelblat-Cerf Y, and Vaadia E (2010). Neuronal correlates of memory formation in motor cortex after adaptation to force field. J. Neurosci 30, 9189–9198. 10.1523/JNEUROSCI.1603-10.2010. [PubMed: 20610753]
32. Cherian A, Fernandes HL, and Miller LE (2013). Primary motor cortical discharge during force field adaptation reflects muscle-like dynamics. J. Neurophysiol 110, 768–783. 10.1152/jn.00109.2012. [PubMed: 23657285]
33. Perich MG, and Miller LE (2017). Altered tuning in primary motor cortex does not account for behavioral adaptation during force field learning. Exp. Brain Res 235, 2689–2704. 10.1007/s00221-017-4997-1. [PubMed: 28589233]
34. Sun X, O'Shea DJ, Golub MD, Trautmann EM, Vyas S, Ryu SI, and Shenoy KV (2022). Cortical preparatory activity indexes learned motor memories. Nature 602, 274–279. 10.1038/s41586-021-04329-x. [PubMed: 35082444]
35. Cowley BR, Snyder AC, Acar K, Williamson RC, Yu BM, and Smith MA (2020). Slow drift of neural activity as a signature of impulsivity in macaque visual and prefrontal cortex. Neuron 108, 551–567.e8. 10.1016/j.neuron.2020.07.021. [PubMed: 32810433]
36. Hennig JA, Oby ER, Golub MD, Bahureksa LA, Sadtler PT, Quick KM, Ryu SI, Tyler-Kabara EC, Batista AP, Chase SM, and Yu BM (2021). Learning is shaped by abrupt changes in neural engagement. Nat. Neurosci 24, 727–736. 10.1038/s41593-021-00822-8. [PubMed: 33782622]
37. Roesch MR, and Olson CR (2004). Neuronal activity related to reward value and motivation in primate frontal cortex. Science 304, 307–310. 10.1126/science.1093223. [PubMed: 15073380]
38. Scott SH, and Kalaska JF (1995). Changes in motor cortex activity during reaching movements with similar hand paths but different arm postures. J. Neurophysiol 73, 2563–2567. 10.1152/jn.1995.73.6.2563. [PubMed: 7666162]

39. Jarosiewicz B, Chase SM, Fraser GW, Velliste M, Kass RE, and Schwartz AB (2008). Functional network reorganization during learning in a brain-computer interface paradigm. Proc. Natl. Acad. Sci. USA 105, 19486–19491. 10.1073/pnas.0808113105. [PubMed: 19047633]
40. Ganguly K, and Carmena JM (2009). Emergence of a stable cortical map for neuroprosthetic control. PLOS Biol. 7, e1000153. 10.1371/journal.pbio.1000153. [PubMed: 19621062]
41. Koralek AC, Jin X, Long JD, Costa RM, and Carmena JM (2012). Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. Nature 483, 331–335. 10.1038/nature10845. [PubMed: 22388818]
42. Hwang EJ, Bailey PM, and Andersen RA (2013). Volitional control of neural activity relies on the natural motor repertoire. Curr. Biol 23, 353–361. 10.1016/j.cub.2013.01.027. [PubMed: 23416098]
43. Sadtler PT, Quick KM, Golub MD, Chase SM, Ryu SI, Tyler-Kabara EC, Yu BM, and Batista AP (2014). Neural constraints on learning. Nature 512, 423–426. 10.1038/nature13665. [PubMed: 25164754]
44. Gulati T, Guo L, Ramanathan DS, Bodepudi A, and Ganguly K (2017). Neural reactivations during sleep determine network credit assignment. Nat. Neurosci 20, 1277–1284. 10.1038/nn.4601. [PubMed: 28692062]
45. Jeon BB, Fuchs T, Chase SM, and Kuhlman SJ (2022). Existing function in primary visual cortex is not perturbed by new skill acquisition of a non-matched sensory task. Nat. Commun 13, 3638.1. [PubMed: 35752622]
46. Krakauer JW, Hadjiosif AM, Xu J, Wong AL, and Haith AM (2019). Motor learning. Compr. Physiol 9, 613–663. 10.1002/cphy.c170043. [PubMed: 30873583]
47. Golub MD, Sadtler PT, Oby ER, Quick KM, Ryu SI, Tyler-Kabara EC, Batista AP, Chase SM, and Yu BM (2018). Learning by neural re-association. Nat. Neurosci 21, 607–616. 10.1038/s41593-018-0095-3. [PubMed: 29531364]
48. Shadmehr R, and Mussa-Ivaldi FA (1994). Adaptive representation of dynamics during learning of a motor task. J. Neurosci 14, 3208–3224. 10.1523/JNEUROSCI.14-05-03208.1994. [PubMed: 8182467]
49. Hennig JA, Golub MD, Lund PJ, Sadtler PT, Oby ER, Quick KM, Ryu SI, Tyler-Kabara EC, Batista AP, Yu BM, and Chase SM (2018). Constraints on neural redundancy. eLife 7, e36774. 10.7554/eLife.36774. [PubMed: 30109848]
50. Kaufman MT, Churchland MM, Ryu SI, and Shenoy KV (2014). Cortical activity in the null space: permitting preparation without movement. Nat. Neurosci 17, 440–448. 10.1038/nn.3643. [PubMed: 24487233]
51. Vyas S, Even-Chen N, Stavisky SD, Ryu SI, Nuyujukian P, and Shenoy KV (2018). Neural population dynamics underlying motor learning transfer. Neuron 97, 1177–1186.e3. 10.1016/j.neuron.2018.01.040. [PubMed: 29456026]
52. Druckmann S, and Chklovskii DB (2012). Neuronal circuits underlying persistent representations despite time varying activity. Curr. Biol 22, 2095–2103. 10.1016/j.cub.2012.08.058. [PubMed: 23084992]
53. Rule ME, O'Leary T, and Harvey CD (2019). Causes and consequences of representational drift. Curr. Opin. Neurobiol 58, 141–147. 10.1016/j.conb.2019.08.005. [PubMed: 31569062]
54. Mau W, Hasselmo ME, and Cai DJ (2020). The brain in motion: how ensemble fluidity drives memory-updating and flexibility. eLife 9, e63550. 10.7554/eLife.63550. [PubMed: 33372892]
55. Deitch D, Rubin A, and Ziv Y (2021). Representational drift in the mouse visual cortex. Curr. Biol 31, 4327–4339.e6. 10.1016/j.cub.2021.07.062. [PubMed: 34433077]
56. Schoonover CE, Ohashi SN, Axel R, and Fink AJP (2021). Representational drift in primary olfactory cortex. Nature 594, 541–546. 10.1038/s41586-021-03628-7. [PubMed: 34108681]
57. Stevenson IH, Cherian A, London BM, Sachs NA, Lindberg E, Reimer J, Slutzky MW, Hatsopoulos NG, Miller LE, and Kording KP (2011). Statistical assessment of the stability of neural movement representations. J. Neurophysiol 106, 764–774. 10.1152/jn.00626.2010. [PubMed: 21613593]
58. Krakauer JW, Ghez C, and Ghilardi MF (2005). Adaptation to visuomotor transformations: consolidation, interference, and forgetting. J. Neurosci 25, 473–478. 10.1523/JNEUROSCI.4218-04.2005. [PubMed: 15647491]

59. Robertson EM, Pascual-Leone A, and Miall RC (2004). Current concepts in procedural consolidation. Nat. Rev. Neurosci 5, 576–582. 10.1038/nrn1426. [PubMed: 15208699]
60. Shibata K, Sasaki Y, Bang JW, Walsh EG, Machizawa MG, Tamaki M, Chang LH, and Watanabe T (2017). Overlearning hyperstabilizes a skill by rapidly making neurochemical processing inhibitory-dominant. Nat. Neurosci 20, 470–475. 10.1038/nn.4490. [PubMed: 28135242]
61. Mooney RA, Bastian AJ, and Celnik PA (2021). Training at asymptote stabilizes motor memories by reducing intracortical excitation. Cortex 143, 47–56. 10.1016/j.cortex.2021.06.014. [PubMed: 34375797]
62. Muellbacher W, Ziemann U, Wissel J, Dang N, Kofler M, Facchini S, Boroojerdi B, Poewe W, and Hallett M (2002). Early consolidation in human primary motor cortex. Nature 415, 640–644. 10.1038/nature712. [PubMed: 11807497]
63. Kawai R, Markman T, Poddar R, Ko R, Fantana AL, Dhawale AK, Kampff AR, and Ølveczky BP (2015). Motor cortex is required for learning but not for executing a motor skill. Neuron 86, 800–812. 10.1016/j.neuron.2015.03.024. [PubMed: 25892304]
64. Rubin DB, Hosman T, Kelemen JN, Kapitonava A, Willett FR, Coughlin BF, Halgren E, Kimchi EY, Williams ZM, Simeral JD, et al. (2022). Learned motor patterns are replayed in human motor cortex during sleep. J. Neurosci 42, 5007–5020. 10.1523/JNEUROSCI.2074-21.2022. [PubMed: 35589391]
65. Kim J, Joshi A, Frank L, and Ganguly K (2023). Cortical–hippocampal coupling during manifold exploration in motor cortex. Nature 613, 103–110. 10.1038/s41586-022-05533-z. [PubMed: 36517602]
66. Nader K, and Hardt O (2009). A single standard for memory: the case for reconsolidation. Nat. Rev. Neurosci 10, 224–234. 10.1038/nrn2590. [PubMed: 19229241]
67. Gershman SJ, Monfils MH, Norman KA, and Niv Y (2017). The computational nature of memory modification. eLife 6, e23763. 10.7554/eLife.23763. [PubMed: 28294944]
68. Ajemian R, D'Ausilio A, Moorman H, and Bizzi E (2013). A theory for how sensorimotor skills are learned and retained in noisy and nonstationary neural circuits. Proc. Natl. Acad. Sci. USA 110, E5078–E5087. 10.1073/pnas.1320116110. [PubMed: 24324147]
69. Gallego JA, Perich MG, Naufel SN, Ethier C, Solla SA, and Miller LE (2018). Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. Nat. Commun 9, 4233. 10.1038/s41467-018-06560-z. [PubMed: 30315158]
70. Gava GP, McHugh SB, Lefèvre L, Lopes-Dos-Santos V, Trouche S, El-Gaby M, Schultz SR, and Dupret D (2021). Integrating new memories into the hippocampal network activity space. Nat. Neurosci 24, 326–330. 10.1038/s41593-021-00804-w. [PubMed: 33603228]
71. Nieh EH, Schottdorf M, Freeman NW, Low RJ, Lewallen S, Koay SA, Pinto L, Gauthier JL, Brody CD, and Tank DW (2021). Geometry of abstract learned knowledge in the hippocampus. Nature 595, 80–84. 10.1038/s41586-021-03652-7. [PubMed: 34135512]
72. Libby A, and Buschman TJ (2021). Rotational dynamics reduce interference between sensory and memory representations. Nat. Neurosci 24, 715–726. 10.1038/s41593-021-00821-9. [PubMed: 33821001]
73. Tang C, Herikstad R, Parthasarathy A, Libedinsky C, and Yen SC (2020). Minimally dependent activity subspaces for working memory and motor preparation in the lateral prefrontal cortex. eLife 9, e58154. 10.7554/eLife.58154. [PubMed: 32902383]
74. Xie Y, Hu P, Li J, Chen J, Song W, Wang XJ, Yang T, Dehaene S, Tang S, Min B, and Wang L (2022). Geometry of sequence working memory in macaque prefrontal cortex. Science 375, 632–639. 10.1126/science.abm0204. [PubMed: 35143322]
75. Wolpaw JR, and Kamesar A (2022). Heksor: the central nervous system substrate of an adaptive behaviour. J. Physiol 600, 3423–3452. 10.1113/JP283291. [PubMed: 35771667]
76. Krakauer JW (2009). Motor learning and consolidation: the case of visuomotor rotation. Adv Exp Med Biol. 629, 405–421. 10.1007/978-0-387-77064-2_21. [PubMed: 19227512]
77. Golub MD, Chase SM, Batista AP, and Yu BM (2016). Brain–computer interfaces for dissecting cognitive processes underlying sensorimotor control. Curr. Opin. Neurobiol 37, 53–58. 10.1016/j.conb.2015.12.005. [PubMed: 26796293]

78. Sheahan HR, Franklin DW, and Wolpert DM (2016). Motor planning, not execution, separates motor memories. Neuron 92, 773–779. 10.1016/j.neuron.2016.10.017. [PubMed: 27817979]

79. Orsborn AL, and Pesaran B (2017). Parsing learning in networks using brain–machine interfaces. Curr. Opin. Neurobiol 46, 76–83. 10.1016/j.conb.2017.08.002. [PubMed: 28843838]

80. Golub MD, Yu BM, and Chase SM (2015). Internal models for interpreting neural population activity during sensorimotor control. eLife 4, e10015. 10.7554/eLife.10015. [PubMed: 26646183]

81. Shadmehr R, and Krakauer JW (2008). A computational neuroanatomy for motor control. Exp. Brain Res 185, 359–381. 10.1007/s00221-008-1280-5. [PubMed: 18251019]

82. Masse NY, Grant GD, and Freedman DJ (2018). Alleviating catastrophic forgetting using context-dependent gating and synaptic stabilization. Proc. Natl. Acad. Sci. USA 115, E10467–E10475. 10.1073/pnas.1803839115. [PubMed: 30315147]

83. Parisi GI, Kemker R, Part JL, Kanan C, and Wermter S (2019). Continual lifelong learning with neural networks: A review. Neural Netw. 113, 54–71. 10.1016/j.neunet.2019.01.012. [PubMed: 30780045]

84. Yang GR, Joglekar MR, Song HF, Newsome WT, and Wang XJ (2019). Task representations in neural networks trained to perform many cognitive tasks. Nat. Neurosci 22, 297–306. 10.1038/s41593-018-0310-2. [PubMed: 30643294]

85. Kao TC, Jensen K, van de Ven G, Bernacchia A, and Hennequin G (2021). Natural continual learning: success is a journey, not (just) a destination. Adv. Neural Inf. Process. Syst 34, 28067–28079.

86. Duncker L, Driscoll L, Shenoy KV, Sahani M, and Sussillo D (2020). Organizing recurrent network dynamics by task-computation to enable continual learning. Adv. Neural Inf. Process. Syst 33, 14387–14397.

87. Hennig JA, Oby ER, Losey DM, Batista AP, Yu BM, and Chase SM (2021). How learning unfolds in the brain: toward an optimization view. Neuron 109, 3720–3735. 10.1016/j.neuron.2021.09.005. [PubMed: 34648749]

88. Degenhart AD, Bishop WE, Oby ER, Tyler-Kabara EC, Chase SM, Batista AP, and Yu BM (2020). Stabilization of a brain–computer interface via the alignment of low-dimensional spaces of neural activity. Nat. Biomed. Eng 4, 672–685. 10.1038/s41551-020-0542-9. [PubMed: 32313100]

Highlights

- Learning can change the neural activity that is used to perform familiar tasks
- After learning, neural activity alters to be more appropriate for the learned task
- These changes can occur without interfering with performance of the familiar task
- This “memory trace” of new learning can be seen using a brain-computer interface

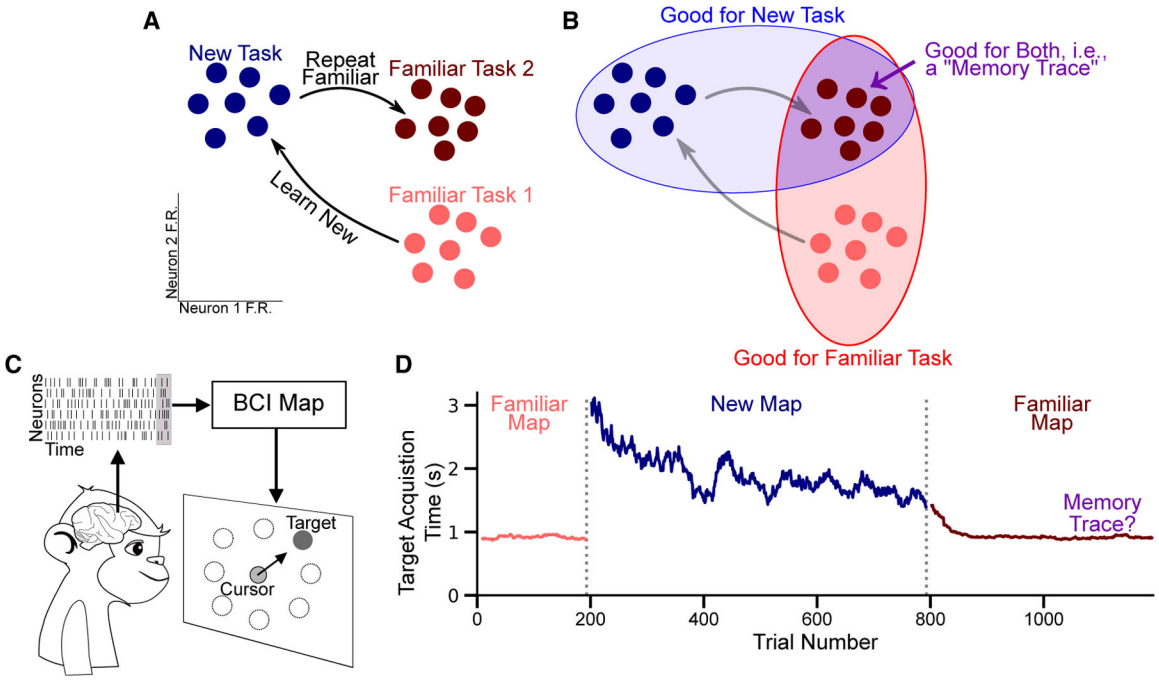


Figure 1. How learning could leave a memory trace in neural population activity

(A) Schematic of how neural activity (colored dots) may change when performing different tasks. Performing the familiar task for the first time (light red; familiar task 1), then the new task (blue), then the familiar task again (dark red; familiar task 2) may each yield distinct population activity patterns.

(B) There are many different population activity patterns that can be appropriate for the same task (red oval for the familiar task, blue oval for the new task). We consider the possibility that the activity patterns when reperforming the familiar task after learning the new task are appropriate for both the familiar task and the new task (purple intersection). We refer to this as a memory trace.

(C) The activity of ~90 neural units recorded in the primary motor cortex (M1) was translated into movements of a computer cursor using a BCI. A BCI directly relates neural activity to behavior (the horizontal and vertical velocities of the cursor) using a map specified by the experimenters. Specifically, only the 10 dimensions of greatest shared variance in the neural population dictate how the cursor moves (see STAR Methods).

(D) Target acquisition times for a representative session (N20160714). The monkey first used the familiar map to control the cursor proficiently (light red, familiar task 1). We then switched to the new map, which the monkey had not used before, and target acquisition times abruptly increased (dark blue, new task). The monkey learned to use the new map through trial and error (dark blue, decreasing target acquisition times). Finally, the familiar map was reinstated (dark red, familiar task 2). We focus on this second familiar map period for identifying whether there exists a memory trace. For visualization, acquisition times were smoothed with a causal 25-trial moving window and are not shown for the first 8 trials of each task. Success rates were 100% for all three tasks of this example session.

See also Figure S1.

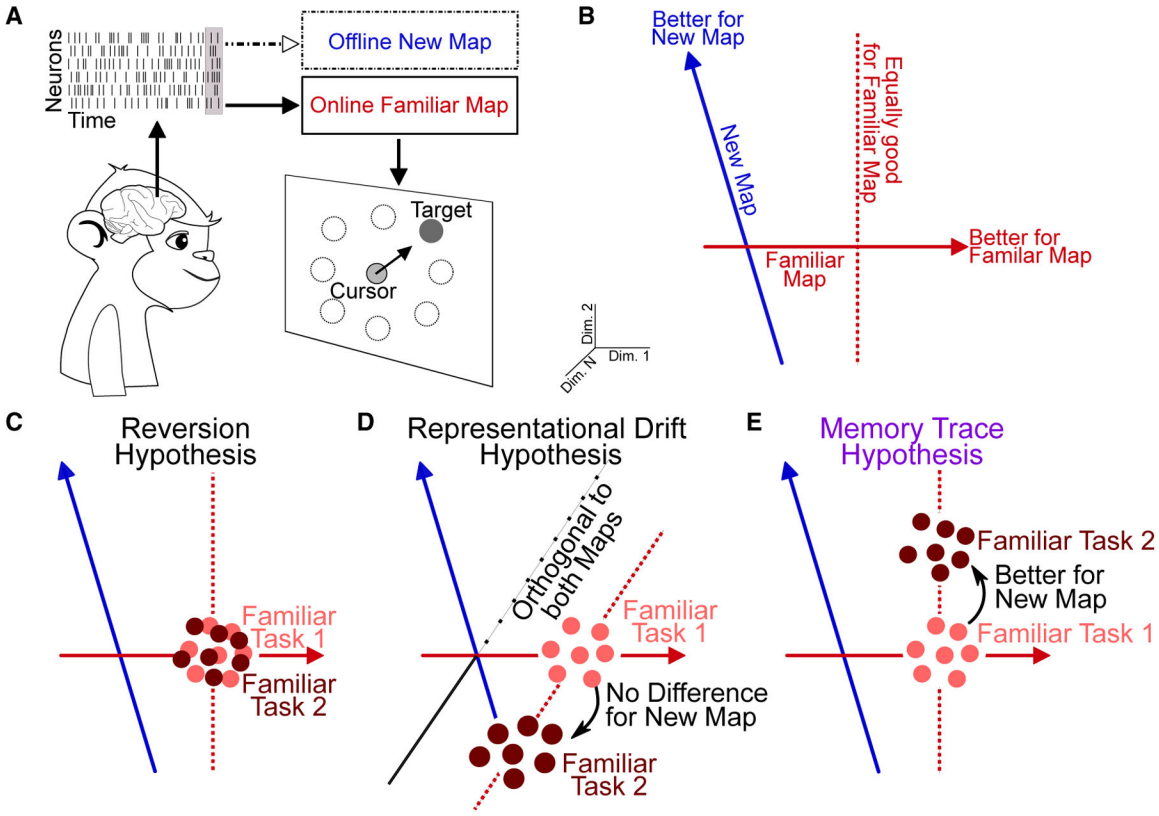


Figure 2. Leveraging a BCI to probe the existence of a memory trace

(A) The online BCI map dictates cursor movement. The same neural activity can also be interpreted with respect to an offline BCI map that did not determine cursor movement. During familiar task 1 and familiar task 2, the online map is the familiar map.

(B) Schematic of BCI maps in population activity space. The projection of neural activity onto a given BCI map (online or offline) determines how appropriate the activity is for that map. For example, many different population activity patterns (vertical dotted line) are projected to the same point and are thus equally good, for the familiar map. However, those same activity patterns are not all equally good for the new map; those near the top of the dashed line are better for the new map than those near the bottom. For illustrative purposes, we show a 2D neural space mapped to a 1D cursor velocity. In the actual experiments, the neural space was ~90D, which was mapped to a 2D cursor velocity.

(C–E) We explore three possibilities for where neural activity might reside during familiar task 2. (C) Reversion hypothesis: familiar task 2 neural activity is similar to that used during familiar task 1. (D) Representational drift hypothesis: familiar task 2 neural activity is different from that used during familiar task 1, but not in a manner that influences performance through the new map in a systematic way. We show a stylized three-dimensional (3D) space, with an axis (black line) coming out of the page to illustrate how neural activity could change along a dimension orthogonal to both the familiar map and the new map. (E) Memory trace hypothesis: familiar task 2 neural activity contains a memory trace, whereby neural activity is more appropriate for the new map than it was during familiar task 1.

See also Figure S2.

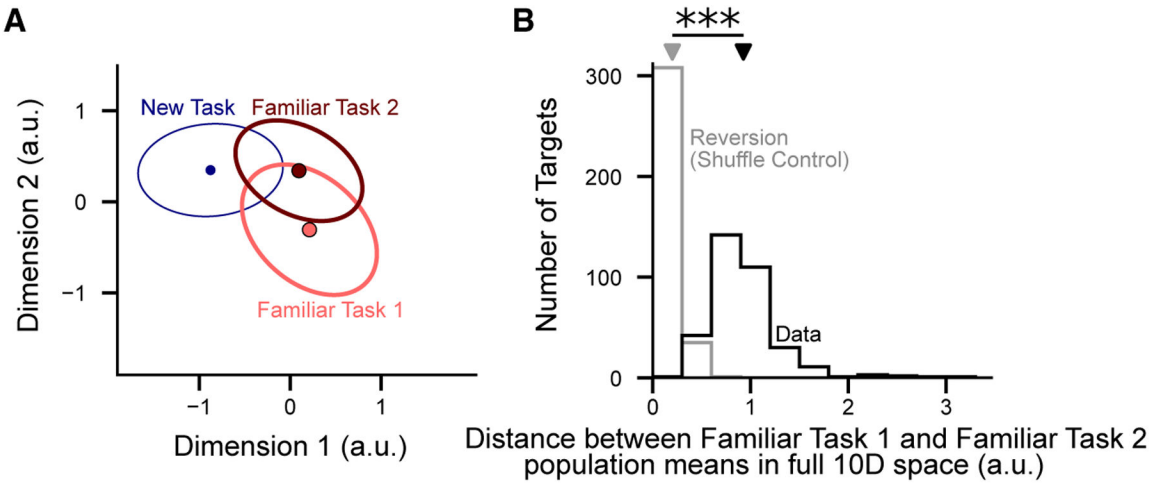


Figure 3. Learning a new task changes the neural representation of a familiar task

(A) A view of the population neural activity for one example target (J20120601; target 270°) across all three task periods. We applied linear discriminant analysis (LDA) to find the plane that best separates the neural activity from the three task periods. Activity is projected onto that plane, with mean and covariances across timesteps shown. Examining the dimensions of highest shared variance provides similar results (Figure S3A).

(B) Population activity was different between familiar task 1 and familiar task 2 ($P < 10^{-10}$, two-sided paired Wilcoxon signed-rank test, $n = 344$ targets). Black shows the Mahalanobis distance between the familiar task 1 and familiar task 2 population activity means in the 10D latent space. This distance was computed separately for each of the eight targets in the experiment, aggregated over all 43 experiments. Gray indicates the prediction of the reversion hypothesis, obtained using a shuffle control (see STAR Methods).

See also Figure S3.

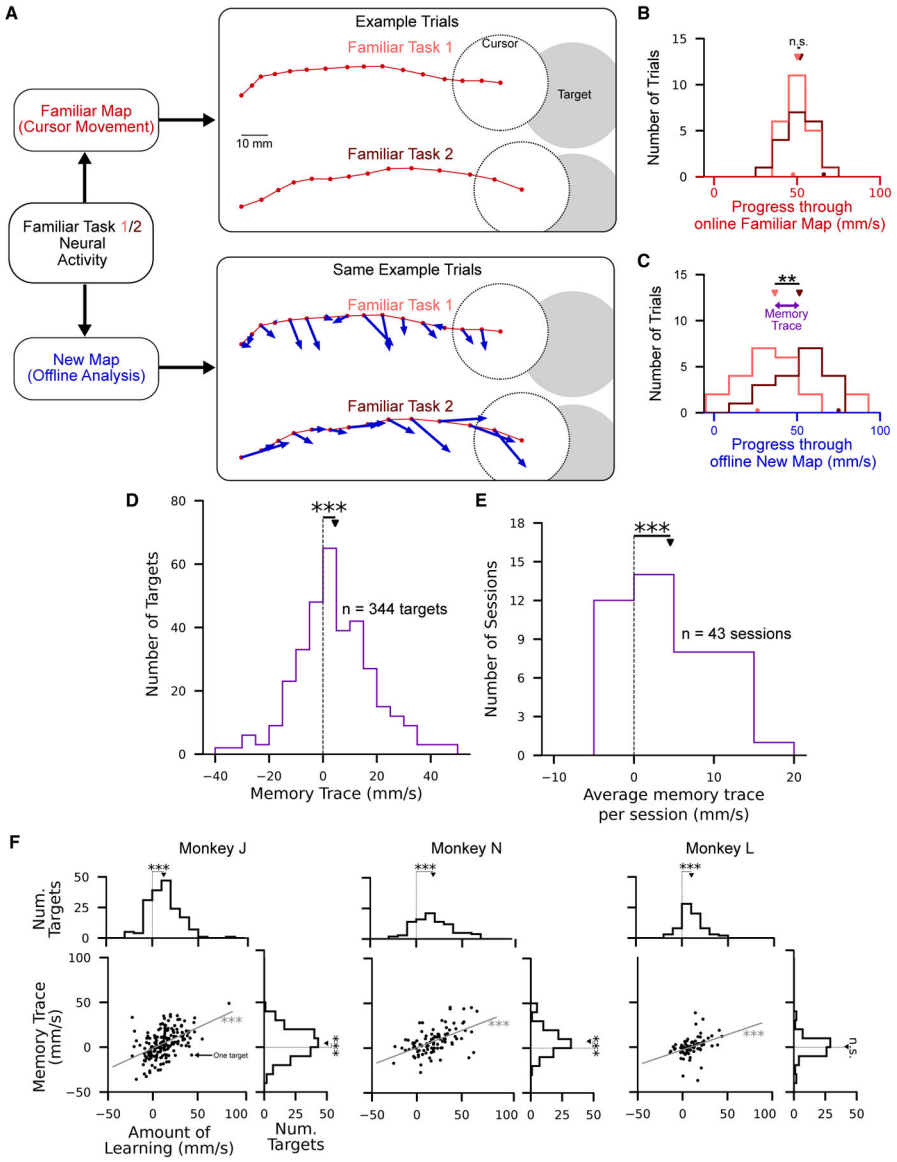


Figure 4. Learning leaves a memory trace

(A) During familiar task 1 and familiar task 2, neural activity drives the cursor through the familiar map (red trajectories, with dots denoting cursor positions at each timestep). The same neural activity can also be projected through the new map in an offline analysis (blue arrows indicate cursor velocity at each time step). Both trials come from the 225^{+} target from session N20160329. For visualization purposes, the target directions were rotated to orient at 0°.

(B) Average progress through the online familiar map for each trial to the target shown in (A). Task performance, measured by progress (see STAR Methods), was not different between familiar task 1 and familiar task 2 (P = 0:43, two-sided unpaired Wilcoxon rank-sum test). Dots on the horizontal axis denote the average progress for the trials shown in (A). Triangles above the histograms denote the mean of each distribution.

(C) Average progress through the offline new map for each trial to the target shown in (A). The difference in average progress defines the memory trace for that target. For this target, there was higher progress through the new map during familiar task 2 than there was during familiar task 1, yielding a memory trace of 14.49 mm/s ($P = 0{:}0077$, two-sided unpaired Wilcoxon rank-sum test). (D) The memory trace for each of the eight targets per session, aggregated across all sessions and all three monkeys. On average, the memory trace was positive ($P = 2{:}53 \times 10^{-8}$, $n = 344$ targets, two-sided paired Wilcoxon signed-rank test). Note that the memory trace can be negative, which indicates that progress through the new map is worse during familiar task 2 than familiar task 1.

(E) The average memory trace per session was positive ($P = 4{:}25 \times 10^{-5}$, $n = 43$ sessions, two-sided paired Wilcoxon signed-rank test).

(F) The size of the memory trace was correlated with the size of learning (monkey J, $R^2 = 0{:}25; P < 10^{-10}$, one-sided F test, $n = 176$ targets; monkey N, $R^2 = 0{:}29; P = 1{:}23 \times 10^{-8}; n = 96$; monkey L, $R^2 = 0{:}13; P = 0{:}0017; n = 72$; see Figure S4E for same plot colored by experimental session). We then considered the marginal distribution of each quantity. The amount of learning was positive for all three monkeys (top marginal histograms; monkey J, $P < 10^{-10}$, two-sided paired Wilcoxon sign-rank test, $n = 176$ targets; monkey N, $P < 10^{-10}$, $n = 96$; $P = 3{:}21 \times 10^{-9}$, $n = 72$ targets). The memory trace per target was positive for monkeys J and N, but not significantly different from zero for monkey L (right marginal histograms; monkey J, $P = 1{:}57 \times 10^{-4}$, two-sided paired Wilcoxon sign-rank test, $n = 176$ targets; monkey N, $P = 1{:}99 \times 10^{-6}$, $n = 96$; monkey L, $P = 0{:}61$, $n = 72$).

See also Figure S4.

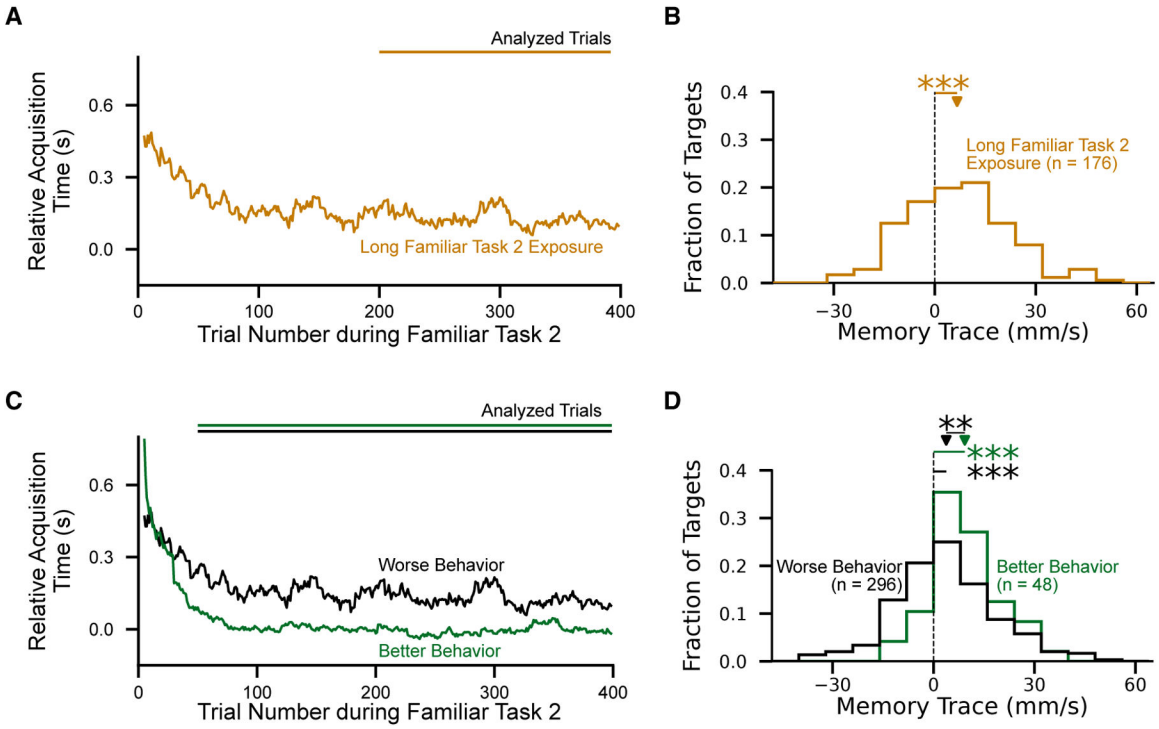


Figure 5. The memory trace persists over time and coexists alongside proficient task performance

(A) To study the persistence of the memory trace, we considered the end of familiar task 2 (gold bar) for the sessions that had the most trials of familiar task 2. Zero relative acquisition time represents the average acquisition time for that target during familiar task 1.

(B) The memory trace persisted, remaining positive after extended exposure to familiar task 2 ($P = 9{:}65 \times 10^{-8}$, two-sided paired Wilcoxon sign-rank test).

(C) Behavioral performance during familiar task 2 from two example sessions, one session with behavior as good or better than during familiar task 1 (faster acquisition time; green) and the other with worse behavior than during familiar task 1 (slower acquisition times; black).

(D) To evaluate the influence that behavioral performance during familiar task 2 has on the size of the memory trace, we split targets into two groups. The first group contained targets where the mean target acquisition time during familiar task 2 was less than the mean target acquisition time during familiar task 1 (better behavior, green). The second group contained targets where the mean target acquisition time during familiar task 2 was greater than during familiar task 1 (worse behavior, black). The size of the memory trace was larger for the better behavior group than the worse behavior group ($P = 0{:}0025$, two-sided unpaired Wilcoxon sign-rank test). Considered separately, the memory trace was positive for each group of sessions (better behavior, $P = 8{:}08 \times 10^{-8}$, two-sided paired Wilcoxon rank-sum test; worse behavior, $P = 5{:}73 \times 10^{-5}$).

See also Figure S5.

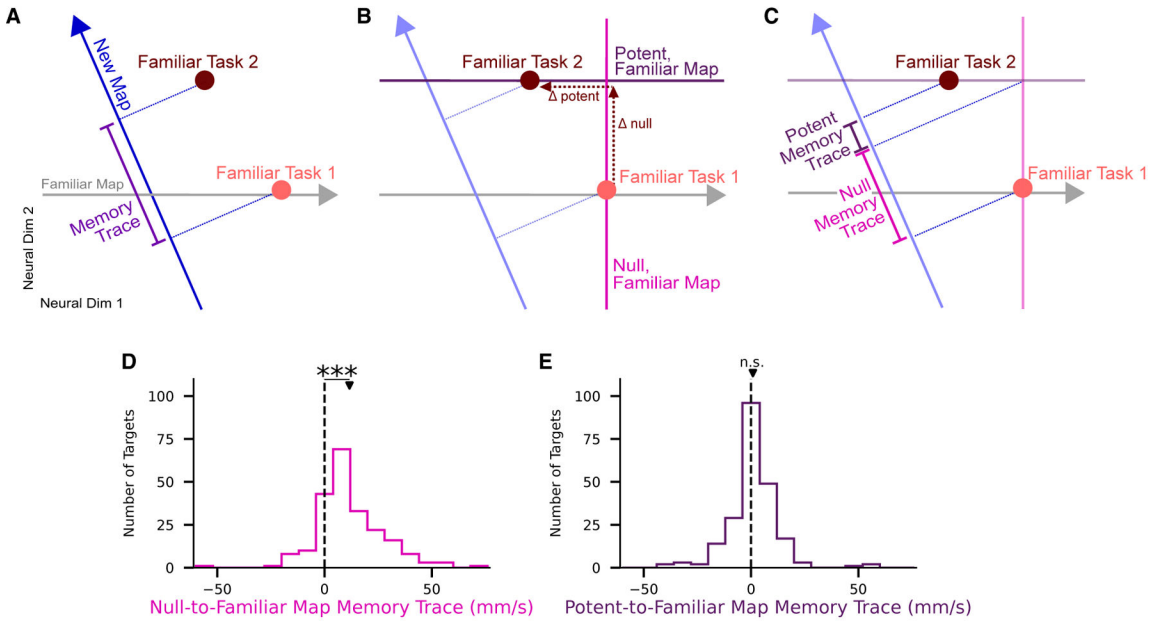


Figure 6. The memory trace is predominantly in the null space of the familiar map

(A) Memory trace depicted in same space as Figures 2C–2E.

(B and C) During familiar task 1 (light red dot) and familiar task 2 (dark red dot), the cursor is controlled using the familiar map (gray arrow). Familiar task 2 activity is further along the new map (blue arrow) than familiar task 1 activity, indicating higher progress along the new map. The memory trace is defined as difference in the projection onto the new map. (B) The change in neural activity from (A) can be decomposed into a component that is output-potent to the familiar map (Δ potent) and a component that is output-null to the familiar map (Δ null). (C) We can correspondingly decompose the memory trace into output-potent and output-null components.

(D) Of the targets with a positive memory trace (218 out of 344 targets), the memory trace consistently resided in dimensions null to the familiar map ($P < 10^{-10}$, two-sided paired Wilcoxon signed-rank test, $n = 218$ targets across all monkeys).

(E) The contributions from the potent space are not significantly different from zero ($P =$ 0:17, two-sided paired Wilcoxon signed-rank test, $n = 218$ targets across all monkeys).

See also Figure S6.

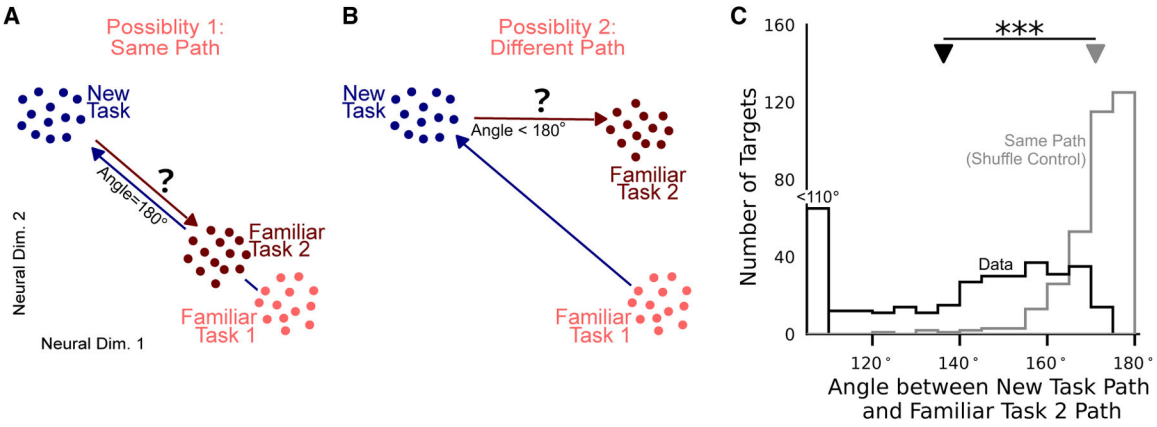


Figure 7. The path of washout does not retrace the path of learning

(A and B) We consider two possibilities for how the memory trace arises during familiar task 2. (A) The first possibility is that the path of washout (i.e., from the end of the new task to familiar task 2) retraces the path of learning (i.e., from familiar task 1 to the end of the new task). (B) The second possibility is that these two paths are distinct, implying that the washout is not simply "unlearning."

(C) To distinguish between these two possibilities, we measured the angle between these two paths in the 10-dimensional latent space of neural activity. This angle (black histogram) was smaller than the angles that would be obtained under possibility 1 (gray histogram; see STAR Methods; $P < 10^{-10}$, two-sided paired Wilcoxon signed-rank test, $n = 344$ targets across monkeys). This implies the paths of learning and washout are distinct (possibility 2). The targets that exhibited near 180° angles between the learning and washout paths did not all come from the same sessions, meaning there was no single session in which learning was undone through the process of unwinding. Among the 14 targets where the difference between the paths exceeded 170°, the maximum number from a single session was 3 (consistent with selecting 14 of the targets at random, $P = 0{:}16$).

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms / strains		
Rhesus macaques (Macaca mulatta)	Alpha Genesis	N/A
Software and algorithms		
Python 3.11	Python Software Foundation	https://www.python.org/
Pandas 2.03	PyPI	https://pandas.pydata.org/
NumPy 1.24.3	PyPI	https://numpy.org/
Matplotlib 3.7.2	PyPI	https://matlabplotlib.org/
SciPy 1.11.1	PyPI	https://www.scipy.org/
Analysis code	Darby Losey	https://github.com/loseydm/MemoryTrace
Other		
96 electrode arrays	Blackrock Microsystems	N/A