

## Neural Population Dynamics Underlying Motor Learning Transfer

### Highlights

- Covert learning via a brain-machine interface transfers to overt reaching behavior
- Covert learning systematically changes motor cortical preparatory activity
- Covert and overt movements share preparatory neural states and facilitate transfer
- Covert and overt movements engage a similar neural dynamical system

### Authors

Saurabh Vyas, Nir Even-Chen, Sergey D. Stavisky, Stephen I. Ryu, Paul Nuyujukian, Krishna V. Shenoy

### Correspondence

smvyas@stanford.edu

### In Brief

Vyas et al. ask whether learning “covertly,” without physical movements, can transfer to overt behavior. By using visuomotor perturbations, they show that covert and overt movements derive from a common neural substrate consisting of motor cortical preparatory activity that facilitates transfer of learning.



# Neural Population Dynamics Underlying Motor Learning Transfer

Saurabh Vyas,<sup>1,9,\*</sup> Nir Even-Chen,<sup>2,5</sup> Sergey D. Stavisky,<sup>2,3</sup> Stephen I. Ryu,<sup>2,8</sup> Paul Nuyujukian,<sup>1,2,3,5,6</sup> and Krishna V. Shenoy<sup>1,2,4,5,6,7</sup>

<sup>1</sup>Department of Bioengineering, Stanford University, Stanford, CA 94305, USA

<sup>2</sup>Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA

<sup>3</sup>Department of Neurosurgery, Stanford University, Stanford, CA 94305, USA

<sup>4</sup>Department of Neurobiology, Stanford University, Stanford, CA 94305, USA

<sup>5</sup>Bio-X Program, Stanford University, Stanford, CA 94305, USA

<sup>6</sup>Stanford Neurosciences Institute, Stanford University, Stanford, CA 94305, USA

<sup>7</sup>Howard Hughes Medical Institute, Stanford University, Stanford, CA 94305, USA

<sup>8</sup>Palo Alto Medical Foundation, Palo Alto, CA 94301, USA

<sup>9</sup>Lead Contact

\*Correspondence: [smvyas@stanford.edu](mailto:smvyas@stanford.edu)

<https://doi.org/10.1016/j.neuron.2018.01.040>

## SUMMARY

Covert motor learning can sometimes transfer to overt behavior. We investigated the neural mechanism underlying transfer by constructing a two-context paradigm. Subjects performed cursor movements either overtly using arm movements, or covertly via a brain-machine interface that moves the cursor based on motor cortical activity (in lieu of arm movement). These tasks helped evaluate whether and how cortical changes resulting from “covert rehearsal” affect overt performance. We found that covert learning indeed transfers to overt performance and is accompanied by systematic population-level changes in motor preparatory activity. Current models of motor cortical function ascribe motor preparation to achieving initial conditions favorable for subsequent movement-period neural dynamics. We found that covert and overt contexts share these initial conditions, and covert rehearsal manipulates them in a manner that persists across context changes, thus facilitating overt motor learning. This transfer learning mechanism might provide new insights into other covert processes like mental rehearsal.

## INTRODUCTION

Understanding motor-related covert mental processes, such as imagined or intended movements, and mental rehearsal is tantalizing as these internal behaviors have been shown to exhibit varying degrees of motor learning transfer (Denis, 1985; Papaxanthis et al., 2002). Decades of human behavioral studies have shown that mental rehearsal can improve motor skills such as throwing darts or making free throws (Feltz and Landers, 1983), and mental rehearsal has also been shown to sometimes

aid in rehabilitation (Warner and McNeill, 1988; Buch et al., 2008; Saposnik et al., 2010; Silvoni et al., 2011). Working theories posit that motor learning transfer is a result of covert learning engaging neural population activity similar to that employed during overt practice. In support of this, “mirror neurons” in ventral premotor cortex have been shown to discharge both when actions are overtly performed and when they are observed (Rizzolatti et al., 2001). These results, however, are still debated (Hickok, 2009) and do not propose mechanistic hypotheses about why neural similarity is helpful for learning transfer.

This debate stems primarily from the fact that mental rehearsal, and covert processes in general, are difficult to define and even more challenging to experimentally study. They are open-loop hidden processes, where experimenters cannot directly observe the internal process or the trial-by-trial progression of learning. In this study, we present a covert process that enables a direct and real-time probe into this evolution, by “closing the loop.” We use a brain-machine interface (BMI), which takes as input neural activity from dorsal premotor and primary motor cortex. This neural activity is mapped through a fixed mathematical function, i.e., “decoder,” to produce a two-dimensional cursor movement. This defines a closed-loop system by which subjects receive visual feedback of the on-screen cursor, and the experimenters observe both the behavior and the evolving neural activity on a trial-by-trial basis. The BMI context elicits internal motor processes that share an end-goal with overt processes because subjects use the decoder (i.e., neural activity without overt movements) to make the same cursor movements as they will perform subsequently using arm movements. We constructed the decoder by associating the kinematics of automated cursor movements with neural activity recorded while subjects observed these movements (Gilja et al., 2012). This was done in contrast to using neural activity measured during overt movements. Previous findings have shown that neural signals involved in watching cursor movements are engaged in mental rehearsal and involve many of the same cells as when generating movement (Cisek and Kalaska, 2004).



The BMI paradigm is a powerful tool for studying learning because the decoder establishes a causal link between behavior (i.e., cursor movements), and all the neurons directly responsible for producing that behavior. Thus, the only way to modify the behavior is to causally modify the input neurons to the decoder. Several groups have used BMIs to dissect a variety of sensorimotor and learning processes (recent germane studies include Jarosiewicz et al., 2008; Ganguly and Carmena, 2009; Ganguly et al., 2011; Orsborn et al., 2014; Sadtler et al., 2014; Golub et al., 2015, 2018; Athalye et al., 2017; Prsa et al., 2017). Here, we use BMIs to establish a “covert rehearsal” paradigm whereby subjects can “practice” or in a sense “rehearse” a motor task directly using their neural activity without movement (i.e., covertly). We can then evaluate the degree of learning transfer by having the subjects repeat the same task using overt movements.

In this non-human primate study, we cannot definitively equate covert rehearsal to imagined movements or mental rehearsal (though this may well be what the monkeys are doing). Critically, covert rehearsal differs from mental rehearsal in that it provides the monkeys with real-time visual feedback of the on-screen cursor. This design, however, is intentional as it provides a first-ever direct probe into studying the single-trial neural (and behavioral) correlates of covert learning. Thus, the goal of this study is to use the covert rehearsal paradigm to evaluate two key scientific questions underlying most covert processes: (1) can covert processes (which covert rehearsal is a type of) facilitate overt motor learning, and (2) if so, what neural mechanism mediates this transfer? If covert processes can facilitate overt learning, we expect to observe that neural changes resulting from learning in one context would result in behavioral changes in the other. This would suggest that covert learning does transfer, and the corresponding neural activity would provide a glimpse at its mechanism. We note, however, that in this study learning is measured through the lens of motor adaptation; monkeys learn to adapt to a visuomotor rotation. Adaptation is used here, as is commonly done in the literature, as one simple subset of motor learning (Jarosiewicz et al., 2008; Huang et al., 2011; Chase et al., 2012; Ranganathan et al., 2014; Mathis et al., 2017). Thus, a conservative interpretation of our claims concerning motor learning and its transfer, including any relation to mental rehearsal, should be restricted to the transfer of adaptation. Nonetheless, recent evidence (e.g., Churchland et al., 2012) suggests that the dynamical systems-based mechanism that we describe in this study could generally be at work for other more complex motor behaviors. The present experiments thus set the stage for future studies of motor learning transfer in skilled motor tasks.

In the present study, we provide key evidence that covert learning does indeed transfer to overt performance. Concomitantly, we propose a dynamical systems mechanism for motor learning transfer. In particular, our analyses will reveal that overt and covert movements derive from a common neural substrate, which consists of motor preparatory activity. Our analyses will also reveal that learning is consistent with manipulating this preparatory activity, and the common substrate enables persistence of these changes, hence facilitating learning transfer. Furthermore, we find that this substrate is also common to neural

activity recorded during contexts previously shown to be engaged in mental rehearsal. While this does not prove that our covert rehearsal paradigm is the same as mental rehearsal, the neural and behavioral results suggest that our proposed neural mechanism could generally be at work in other covert processes. We will also argue that our covert rehearsal paradigm can serve as a valuable tool for motor learning (and potentially rehabilitation) in and of itself. Lastly, our results add to a growing body of evidence that suggest a dynamical systems interpretation of motor cortex (Gallego et al., 2017), where in particular, we now argue for a more fundamental role of motor preparation in learning.

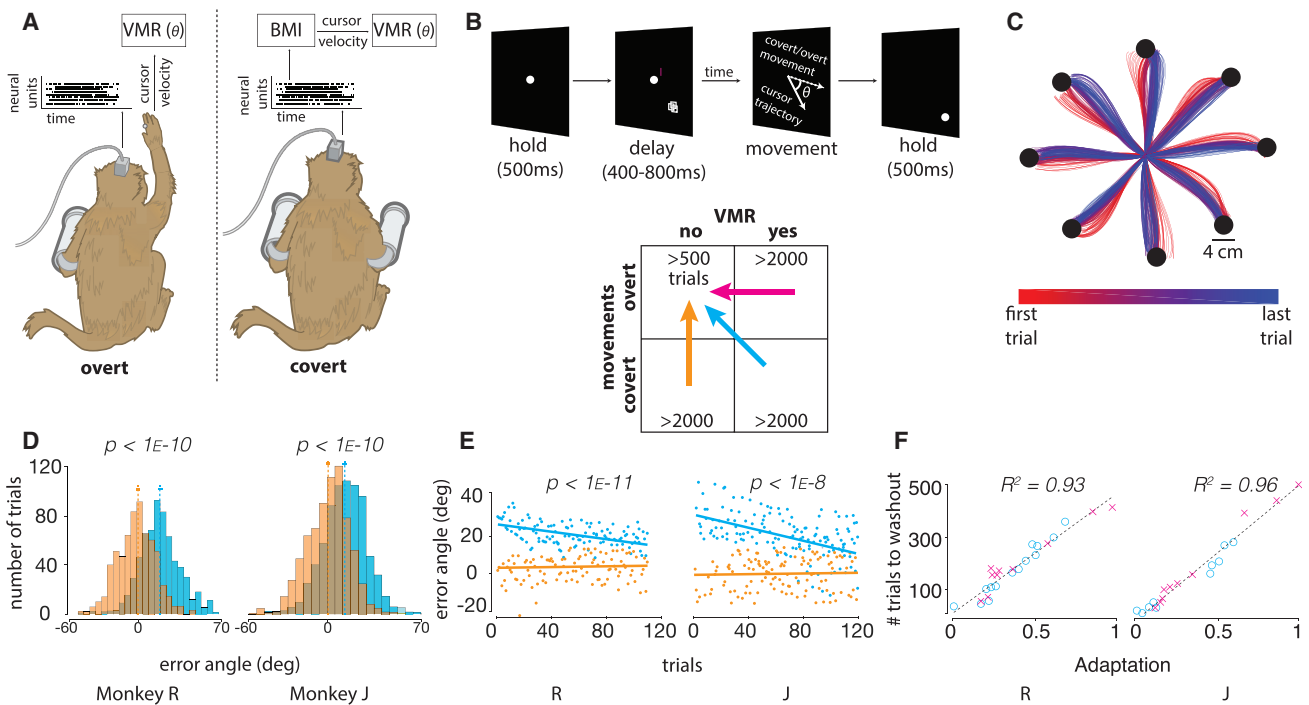
## RESULTS

Two monkeys were trained to move a computer cursor from the center of a virtual reality workspace to one of eight radially arranged targets. Monkeys performed this task either overtly using arm movements or covertly using a BMI (Figure 1A; Figure S1). The task appeared visually similar in both contexts as the virtual reality setup occluded the monkey’s view of his reaching arm. To directly test whether covert rehearsal can affect overt performance, we used the degree of transfer of motor adaptation to visuomotor rotations (VMRs) as an assay to study the relationship between covert and overt movements (Krakauer et al., 2000; Tanaka et al., 2009; Krakauer and Mazzoni, 2011).

We first measured VMR adaptation behavior in the overt context (i.e., “overt-to-overt transfer”) and then separately in the covert context (i.e., “covert-to-covert transfer”) to confirm that our experiments replicated expected VMR adaptation. We focus on VMR after-effects because this will subsequently be a key measure of across-context learning transfer. When VMRs were applied in the overt context, monkeys learned to move their arm at an angle (45°, 60°, or 90° relative to the non-rotated condition) in order for the cursor to directly reach the target (Figure 1B, task-flow). After adaptation, we removed the VMR and measured cursor movement errors (i.e., after-effects). We observed the well-known post-VMR adaptation after-effect: monkeys initially reached in the opposite direction of the VMR, took longer to reach the target, and had large angular errors with respect to the straight-line direction to the target (Figures S2A–S2D). For VMRs introduced during the covert context, monkeys had to modulate neural activity to generate velocity commands at an angle in order to move the cursor directly toward the target. In this context, we ensured that no physical movements were made by ensuring that the correlation between any measured movement and the cursor movement (i.e., the task) was negligible, i.e.,  $\rho < \pm 0.1$  (Figure S1). Results during this covert context were also consistent with VMR adaptation (Figures S2E–S2H). Thus, both overt and covert VMRs exhibit well-established motor adaptation after-effects.

### Motor Adaptation Transfers between Covert and Overt Contexts

Next, we tested whether motor adaptation would transfer across contexts by applying a VMR in the covert context, and after adaptation, switching to the overt context without rotation (i.e., “covert-to-overt transfer”; Figure 1B, bottom, cyan).



**Figure 1. Motor Adaptation Transfers between Covert and Overt Movement Contexts**

(A) Monkeys performed a cursor movement task in one of two contexts: “overt,” where cursor velocity tracked hand velocity, or “covert,” where cursor velocity was controlled by neural activity via a BMI. When a visuomotor rotation (VMR) was applied, the cursor’s movements were offset by the corresponding angle ( $\theta$ ). (B) Top: instructed delay task flow; bottom: experimental flow. (C) Raw arm-controlled cursor trajectories (no VMR applied) immediately following covert VMR movements. Initial trajectories (red traces) exhibit curvature, which indicates transfer of adaptation to the VMR that was applied during the preceding covert context. (D) Histograms (first 100 trials of each session) of the error angle between the arm’s velocity (measured at the half-way radius toward the target) and the vector from workspace center to the target. Colors represent conditions from (B). Vertical dashed lines show means of distributions, and horizontal solid lines mean  $\pm$  SEM.  $p$  values obtained from two-tailed Student’s  $t$  tests. (E) Error angle over time for the same conditions as (D). Solid lines show mean. (F) Number of trials needed to wash out VMR adaptation (i.e., reach 90% of control) during no-VMR overt (magenta) or covert (cyan) blocks is strongly correlated with the degree of adaptation during the preceding covert block (VMR applied). Adaptation is defined as number of targets acquired per second in the 50 trials at the end of the VMR covert block (less the initial performance), normalized by the maximum performance across all blocks. Each point corresponds to one post-VMR block.

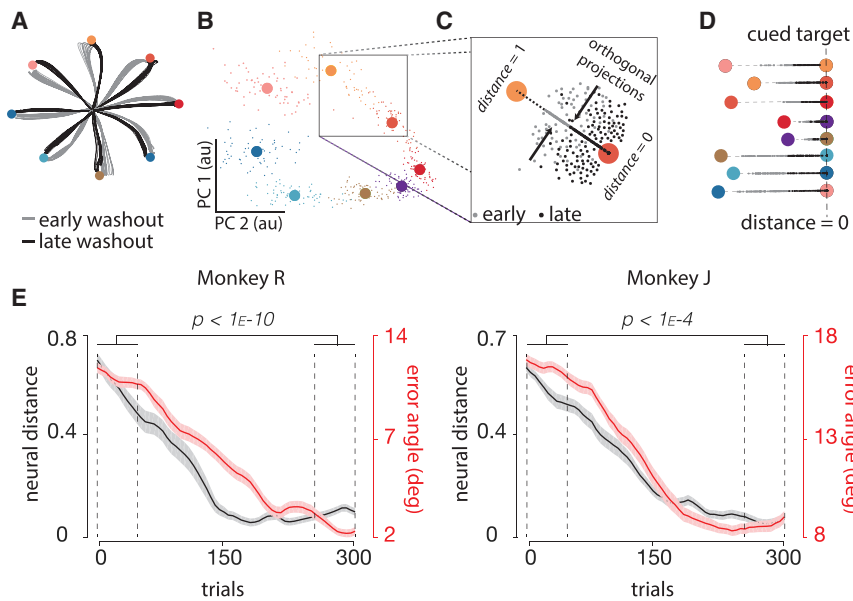
Previous studies have been inconclusive regarding VMR adaptations generalizing across contexts (Krakauer et al., 2000; Taylor et al., 2002; Tanaka et al., 2009). However, we found that monkeys’ overt reaches initially erred in the opposite direction of the preceding covert context VMR (Figure 1C). The corresponding error angles (Figures 1D and 1E) were significantly larger than a control condition where arm reaches followed a no-VMR covert block (Figure 1B, bottom, orange). This indicates that VMR adaptation during the covert context did transfer to the overt context.

While this behavioral effect was robust, its magnitude was smaller than VMR adaptation observed within the overt-to-overt context, both in terms of the error angle, and the rate of washout (Figures S3A and S3B). This suggests that while there was reliable transfer, it was not “complete” transfer. We found no significant relationship between the covert VMR angle and the magnitude of the subsequent transfer to the overt context as measured by initial error angle (Figure S3C). We did find that greater adaptation during the covert context led to longer

washout both for subsequent covert and overt blocks (Figure 1F). Interestingly, this relationship was similar regardless of which context the learning occurred (covert and overt points lie along the same line). This raises the possibility that covert and overt adaptations could engage a similar neural process. Independent from degree of adaptation, we also found that decoders with higher absolute performance led to longer washouts (Figure S3C). This suggests that decoders more similar to the “manual decoder,” i.e., overt reaches, result in greater transfer of learning. Taken together, these findings suggest that covert rehearsal has strong effects on subsequent overt motor behavior (and vice versa, i.e., “overt-to-covert transfer”; Figures S2I–S2L).

### Learning Systematically Changes Motor Preparatory Activity

Having established that transfer occurs, we investigated the neural correlates of this behavior, in hopes of discovering a potential mechanism. We were motivated to examine preparatory



## Figure 2. Motor Adaptation Is Reflected in Population-level Preparatory Activity

(A) Raw arm-controlled cursor trajectories following adaptation to a VMR applied during preceding covert use. Traces are divided into the first 25% (gray) and the last 25% (black) of trials.

(B) 200 ms of instructed-delay period activity from a block of no-VMR baseline overt trials projected into the top two PCs. Point clusters are three-trial-averaged overt trajectories projected into the PC state-space. Larger circles are cluster centroids. Colors denote the eight target conditions and match (A).

(C) Zoom into one condition from (A). Preparatory activity projected into the corresponding region of the PC state-space of (B). Gray/black points correspond to reaches from (A). Neural states are orthogonally projected onto the line connecting the orange and red centroids from (B). Each point is assigned a projection distance, where zero distance denotes being at the cued-target cluster centroid.

(D) Example orthogonal projections for all eight conditions, aligned on the cued-target cluster centroid.

centroid (from B). Early after the switch from VMR covert use, preparatory states were closer to the “anti-VMR” target. (E) Normalized neural distances (combined across reach conditions and sessions) plotted as a function of trials since the switch from covert VMR to overt no-VMR contexts. Red shows the corresponding error angle. Inset compares the first 25% and the last 25% of trials.  $p$  values computed using the Wilcoxon rank-sum test.

neural activity because of a growing body of behavioral (Johnson et al., 2002; Sheahan et al., 2016) and neural (Paz et al., 2003; Stavisky et al., 2017b) evidence showing a link between motor preparation and adaptation. Concretely, we asked whether adaptation transfer from the covert context was reflected in the overt context population-level motor preparatory activity.

This activity is high dimensional, so we visualized it after dimensionality reduction (Cunningham and Yu, 2014). We found a low-dimensional state-space capturing over 80% of the comodulation in the baseline data (Figure 2B). We used this state-space to visualize the delay period activity of overt trials following the switch from a VMR adaptation covert block (Figure 1B cyan shows experimental condition, Figure 2A shows behavior, Figure 2C shows neural projections for one example target). A striking feature of these trials is that early after the switch, preparatory states are shifted toward the neural state corresponding to preparing to move toward the adjacent target (i.e., the “anti-VMR” target). In other words, the monkey’s motor plan is oriented in a direction that opposes the VMR from the previous block. We interpret this shift as evidence of residual adaptation reflected in the preparatory neural state. Quantifying the preparatory states without dimensionality reduction confirmed that immediately following the shift from the covert VMR block (Figure 1, cyan condition), these states were biased in the direction of preparing reaches to the anti-VMR target. Over the course of the washout, preparatory states gradually realigned with the baseline states corresponding to reaches to the cued target (Figures 2D and 2E). Consistent with these population results, we found that single neuron preferred directions (PDs) rotated during learning in the direction corresponding to the VMR. During the washout epochs, PDs reoriented back to baseline (Figure S3D).

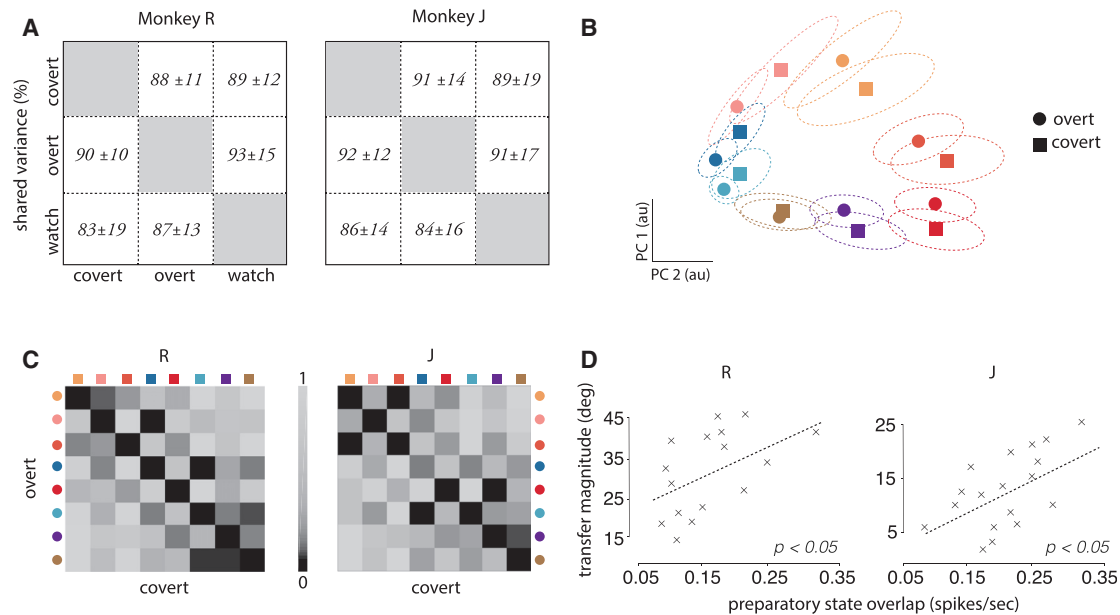
These results demonstrate that: (1) adaptation to the VMR systematically changes motor preparatory activity toward a pattern known to prepare movements angled away from the VMR. These changes enable the subjects to presumably adapt to the VMR, i.e., neural and behavioral changes are strongly correlated on a trial-by-trial basis (Figure 2E). (2) The preparatory changes persist after switching from covert to overt contexts. (3) This neural transfer effect washes out gradually rather than exhibiting a sudden “reset” of the motor circuit, despite visual and proprioceptive feedback (Shadmehr et al., 2010) and the explicit context change of the experimenter removing the monkey’s arm restraint.

## Covert and Overt Movements Share a Common Neural Substrate

Taken together with previous findings (Cisek and Kalaska, 2004; Sobierajewicz et al., 2016), our population analyses suggest that covert rehearsal may involve “practicing” the appropriate motor system preparatory state, whereby behavioral improvements due to learning are accompanied by corresponding changes to the motor plan. This evidence dovetails with recent work arguing that motor cortical preparatory activity functions as advantageous initial conditions for subsequent peri-movement neural dynamics that generate the desired movement (Churchland et al., 2006, 2012; Afshar et al., 2011; Shenoy et al., 2013; Ames et al., 2014). We therefore hypothesized that VMR adaptation transfer was due to covert and overt contexts engaging a similar dynamical system, where in particular both contexts utilized similar initial conditions.

To test this, we first projected baseline overt context preparatory activity into the baseline covert context preparatory activity





### Figure 3. Covert and Overt Movements Share Motor Preparatory States

(A) Percentage of shared variance between the overt context, covert context, and watch (i.e., during decoder training) context neural data. Data from each context are projected into the other contexts before the calculations are done. The subspaces were comprised of PCs (4 for Monkey R, and 6 for Monkey J), which collectively captured over 90% of the neural variance. There is no pairwise statistical significance, as reported by the Wilcoxon rank-sum test.

(B) Example visualization of mean overt (circle) and covert (square) preparatory states projected into a two PC state-space. Colors represent reach conditions. Dotted ellipses represent SEM.

(C) Pairwise statistical population overlap (for the full high-dimensional distributions); zero corresponds to complete statistical overlap. Note that diagonals are close to zero.

(D) The degree of overlap between covert and overt states in the full dimensional space (measured in spikes per second) is correlated with the magnitude of the transfer (error angle measured in degrees for the mean of the first five trials post-switch from covert to overt). Baseline overlap data and subsequent transfer data were collected in the same sessions, with similar parameters (i.e., trial counts, VMR angle, etc.).  $p$  values computed using the F-test, and test the hypothesis of no correlation.

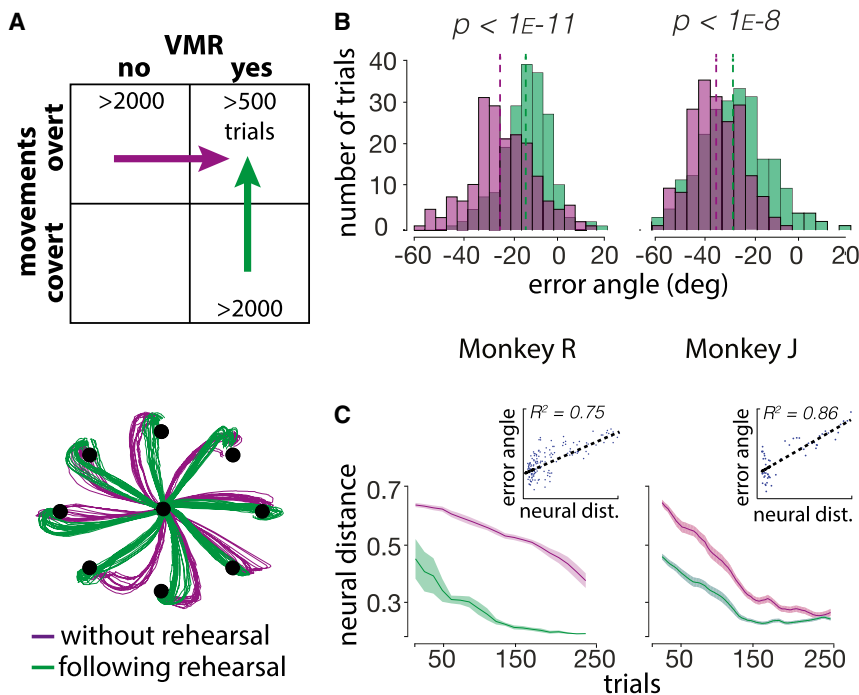
neural state-space. We found that population covariance patterns explaining most of the preparatory activity was shared between the covert and overt contexts (Figure 3A). Furthermore, the neural states corresponding to preparing either covert or overt movements to each target were well aligned (Figure 3B). We quantified this for full-dimensional population activity and found that covert and overt neural states, when preparing movements to a given target, were significantly more similar to each other than to any other target (near-zero diagonal in Figure 3C). Finally, we also found that the degree to which the covert and overt states overlapped (on a session-by-session basis) significantly predicted the magnitude of learning transfer for that session, where greater overlap led to greater transfer (Figure 3D).

Previous findings have demonstrated that monkeys watching cursor movements elicit neural activity consistent with mental rehearsal. Since we trained our decoders using this neural activity, we compared preparatory activity during this cursor-observation period to preparatory activity recoded during covert rehearsal and overt movements. We found that just as covert and overt movements shared preparatory states, observed movements also derive from that common subspace (Figure 3A, “watch” condition). These results suggest that at least at the preparatory level, overt behavior shares neural

operation with both covert rehearsal and putative mental rehearsal.

Together these findings suggest the following potential mechanism for VMR transfer. During covert movements in the presence of a VMR, the sensorimotor system gradually shifts the delay period initial conditions to align with the direction opposing the VMR. We suspect that these changes contribute to the observed VMR adaptation. When the behavioral context switches from covert to overt, two key properties are preserved: (1) the sensorimotor system uses the same neural subspace to prepare cursor movements, and (2) changes in which preparatory state is associated with which cued target persist across the context change. This persistence likely causes subsequent arm movements to be rotated in the direction opposing the VMR, i.e., learning transfer. As monkeys continue to make movements without any VMR, a similar adaptation process reorients the preparatory neural activity back to baseline. The corresponding arm kinematics return to normal once their preparatory initial conditions have reoriented.

This mechanism makes two testable predictions: (1) covert rehearsal should “pre-train” a motor task, and thus require fewer overt trials to learn, and (2) covert rehearsal should achieve this in part by “pre-rotating” the preparatory states to align with the cued target. We tested these predictions by evaluating whether



monkeys adapted to a VMR in the overt context faster if they first covertly rehearsed the VMR (Figure 4A, top). Compared to overtly adapting (i.e., no rehearsal), we found that covert rehearsal resulted in significantly straighter subsequent arm trajectories (Figures 4A and 4B). The neural correlates of this process were also consistent with our prediction: covert rehearsal rotated the preparatory states part way to the fully adapted states, such that overt motor learning required fewer trials to approach the adaptation asymptote (Figure 4C). We observed strong and significant trial-by-trial correlation between the post-rehearsal preparatory state and the behavioral error angle, suggesting that the magnitude of neural changes achieved by rehearsal predicts subsequent improvements in overt motor function (Figure 4C, inset).

The neural mechanism proposed here for the transfer of learning raises an important question: if the preparatory states (i.e., the initial conditions) are similar between overt and covert contexts, shouldn't the underlying dynamical system also be similar? This would suggest a mechanism by which transfer of initial conditions results in movement behavior differences. To test this, we started by repeating the analysis from Figure 3A for the “during movement” epochs and found that population covariance patterns explaining most of the movement period neural activity were shared between the covert and overt contexts (Figure S4A). Next, we explicitly fit a linear dynamical system to *covert* cursor movement data and used the initial condition from the *overt* context to predict the trial-averaged neural trajectory during subsequent reaching. We found the predicted neural trajectories to be highly similar to overt context trajectories (Figure S4B). Finally, we found that the brief but strong oscillatory component in the neural population responses observed during overt reaches (Churchland et al., 2012)

by a common neural substrate, but also potentially engage a similar dynamical machine.

## DISCUSSION

In this study, we investigated whether learning in a covert paradigm, where no movements are made, can transfer to overt performance. Concomitantly, we studied the neural mechanism that mediates such transfer. Our experiments revealed that learning in a BMI-guided covert context does indeed transfer to overt performance, both in terms of post-adaptation after-effects, as well as “pre-training” a motor skill in order to accelerate overt improvements. Our analyses also revealed a dynamical systems mechanism for motor learning transfer. We note that our data show correlation, not causation, in support of this mechanism. However, we believe the significant correlation that we observed on a trial-by-trial basis between the degree of preparatory state rotation and the kinematic error angle (Figures 2E and 4C), the overlap between covert, overt, and observation-only preparatory states (Figures 3A–3C), as well as the significant correlation between the preparatory state overlap and the subsequent magnitude of transfer (Figure 3D), point to the preparatory states likely being necessary for motor learning transfer. This causality could be evaluated in future studies by inactivating preparatory activity prior to movement onset and measuring its effect on learning transfer.

Our preparatory activity findings also touch upon a more fundamental role of motor preparation (Allen et al., 2017; Chen et al., 2017; Makino et al., 2017). The current dynamical systems view of motor cortex ascribes motor preparation to setting the initial state from which neural activity naturally evolves (i.e., guided by lawful rules), presumably to cause movement. This

## Figure 4. Covert Rehearsal Can Enhance Overt Motor Learning

(A) Top: experimental flow. Bottom: purple traces show overt arm trajectories with a  $-45^\circ$  VMR applied following a no-VMR overt block. Green traces show overt  $-45^\circ$  VMR arm trajectories following a covert  $-45^\circ$  VMR block. The rehearsed green trajectories are more direct.

(B) Histograms of error angles for the two conditions. Colors matched to (A). p values were obtained from two-tailed Student's t tests.

(C) Comparison of rehearsed and non-rehearsed normalized preparatory neural state-space distance (as described in Figures 2B and 2C) as a function of trial number. Preparatory states start more adapted after covert rehearsal. Data averaged across sessions and reach conditions. Insets show strong correlation between individual trials' normalized preparatory distance and subsequent error angle ( $p < 0.05$ ).

were also present in the covert data, albeit predominantly the low-frequency components. Together, these findings lend support for an interpretation that overt and covert movements not only are generated

interpretation argues different initial states as being used to produce different movements (Shenoy et al., 2013). Moreover, previous results found that the natural variability in the setting of the preparatory state correlates with reaction time (Afshar et al., 2011). That is, a state “closer” to the desired movement results in faster movements. While the causality of these initial conditions has not yet been established, recent results show that incorrect motor preparation needs to be “corrected” before executing desired movements, albeit not by complete re-planning (Ames et al., 2014). Taken together with our present findings, perhaps a major part of learning can be viewed as the process by which the motor system finds the optimal set of initial conditions that produces the best possible subsequent movement. The fact that we observe motor learning transfer suggests that motor cortex has the ability to perform this manipulation even in the absence of muscle activity. This view is consistent with studies arguing that motor cortex performs important work in “muscle-null” neural dimensions (Kaufman et al., 2014; Stavisky et al., 2017a). Furthermore, recent human results demonstrate that motor preparation, and not execution, helps separate interfering motor memories (Sheahan et al., 2016). The neural correlates of this process could be consistent with the mechanism described in the present study, especially given recent findings that the dynamical systems features of motor cortex are conserved between humans and non-human primates (Pandarinath et al., 2015).

Our study also presents evidence that a similar dynamical machine is in operation for both overt and covert movements (Figure S4). While on one hand this could explain why preparatory activity plays a central role for learning transfer, on the other hand it raises questions about the precise role of M1 (Miri et al., 2017). Why should a dynamical system for a context where no overt movements are made be largely similar to one that generates strong activity for overt movements? One explanation is that these brain regions are primarily concerned with high-level movement intentions, which are shared between the tasks (i.e., directing the cursor to the target). Certainly, the prosthetics community has benefitted from using the strong velocity-related signals present in M1 (Kao et al., 2014). Future studies recording from other and/or deeper areas could reveal more pronounced differences between overt and covert movements and help contextualize the present results. Another possibility is that our undoubtedly simplified model of motor cortex as a low-dimensional linear dynamical system does not capture differences between the contexts that explain a relatively low fraction of the variance but have important effects upon the true, nonlinear dynamical system.

Another contribution of this study is the use of BMIs to describe a “covert rehearsal” paradigm. While we cannot directly assert that covert rehearsal is the same as mental rehearsal, or any other covert process for that matter, we believe that this paradigm is a reasonable strategy for studying motor-related internal processes. First, the fact that covert rehearsal enhances overt performance suggests that, at the very least, it engages some common subset of motor skills and their associated neural machinery, without the need for overt practice.

Second, we found that the performance of the BMI decoder strongly correlated with degree of learning transfer (Figure S3C).

This resonates with reports of mental rehearsal being more effective when performed with more vivid imagery (Ryan and Simons, 1982), and more realistic rehearsal results in better transfer (Hwang et al., 2013). In particular, we found that even in rare cases of poor decoding performance, monkeys were still able to learn the rotation, albeit with a small degree of learning transfer. Our results suggest that the reduction in the magnitude of transfer is likely a result of the poor performing decoder causing the monkeys to practice a noisier version of the true neural pattern, while still allowing adaptation; increasing the trial counts to overcome this noise could potentially boost the magnitude of transfer. This is consistent with the view that the decoder was still built in a “biomimetic” fashion (Shenoy and Carmena, 2014). A non-biomimetic decoder, on the other hand, would likely result in the monkeys cognitively learning the rotation, but the rehearsed pattern would fundamentally differ between contexts, thus resulting in a small degree of transfer regardless of trial count. Future studies could readily investigate this prediction. This would also reconcile the difference between our study and those that find that learning does not generalize across certain contexts, e.g., reach direction (Krakauer et al., 2000). We speculate that in those studies, the learning focused on rehearsing patterns of activity that were not common to the new context. If so, even with a predominantly biomimetic decoder, no transfer should be expected. In contrast, in our study, the same task was presented and performed in both contexts, and our analyses revealed that generalization likely occurred due to a similar dynamical machine being engaged during both contexts.

Third, we found that not only do covert and overt movements share a large degree of preparatory activity variance, but the same holds for neural activity recorded during observation of cursor movements. Previous findings defined such observation as mental rehearsal (Cisek and Kalaska, 2004). Thus, at least at the preparatory level, covert movements, overt movements, and mental rehearsal derive from a common substrate. Finally, we constructed our decoders using neural activity from the observation epochs. The fact that monkeys could use decoders to make successful cursor movements implies that during covert rehearsal they engage neural activity similar to that during decoder training (i.e., putative mental rehearsal), albeit with visual feedback. Taken together, this evidence suggests that mental rehearsal could engage a similar dynamical systems mechanism to the one described here, even if only at the motor preparation level, for facilitating motor learning transfer. Of course, learning could just be a process of refining motor preparation (via above-mentioned argument). From this view, the covert rehearsal paradigm provides a much-needed new avenue for studying covert processes. We note however that in contrast to classical mental rehearsal, the covert rehearsal paradigm is closed-loop. Thus, subjects use visual feedback to help guide producing neural activity patterns associated with overt movements. Future studies will need to evaluate the role of sensory feedback (Liu and Scheidt, 2008; Shabbott and Sainburg, 2010; Suminski et al., 2010), both in terms of its relation to mental rehearsal, and in terms of how best to take advantage of covert rehearsal to guide overt improvement.



One important caveat in this study is that we used VMR adaptation as a specific instance of motor learning because it is amenable for reductionist experiments and has a rich prior literature. It is possible that VMR adaptation is learned and potentially transferred by a different mechanism than complex skilled movements, and thus the generalization of our results to other forms of motor learning remains an open question. However, previous studies have implicated motor preparatory activity as initializing peri-movement neural dynamics in more complex arm reaching behavior (e.g., Churchland et al., 2012). Thus, we speculate that complex motor skill learning, including learning from covert rehearsal, also involves changes in preparatory activity, as seen in VMR adaptation. Future studies could compare whether motor preparatory state is the substrate for learning transfer for more complex skilled movements.

BMI tasks designed to be similar to desired overt motor skills can also be valuable tools in and of themselves with a variety of translational applications. For example, in a rehabilitation application, a patient could be instructed to imagine or attempt to make movements. The patient would receive sensory feedback about how accurately they are modulating their neural activity, perhaps by their arm being moved by external means (Ajiboye et al., 2017), and could iteratively refine subsequent attempts. A second application could use covert rehearsal as a tool to accelerate motor skill learning (for example, a surgical or flight simulator) by allowing safe, targeted, and frequent practice, augmenting existing simulators that provide feedback only on the movement output, rather than the neural output associated with the motor skill. In summary, we used BMIs as a window into previously inaccessible covert mental processes to discover a common neural substrate between covert and overt movements that facilitates motor learning transfer.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENTS AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Task design
  - BMI decoder fitting
  - Residual movement tracking
  - Preparatory neural state analysis
  - Statistics

## SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and can be found with this article online at <https://doi.org/10.1016/j.neuron.2018.01.040>.

## ACKNOWLEDGMENTS

We thank Mackenzie Risch, Michelle Wechsler, Liana Yates, Aliyyah Noel, and Rosie Steinbach for expert surgical assistance and veterinary care. We thank Beverly Davis for administrative support. We thank W.L. Gore Inc. for donating Preclude artificial dura used as part of the chronic electrode array implantation

procedure. We thank members of the Shenoy group and Prof. Surya Ganguli, Prof. Chethan Pandarinath, and Prof. Scott Delp for useful discussions and feedback on this manuscript. We thank Shantini Vyas for numerous useful discussions and detailed comments on this manuscript. S.V. was supported by an NIH F31 Ruth L. Kirschstein National Research Service Award 1F31NS103409-01, an NSF Graduate Research Fellowship, and a Ric Weiland Stanford Graduate Fellowship; N.E.-C. was supported by a Bio-X Fellowship; S.D.S. was supported by an ALS Association Milton Safenowitz Postdoctoral Fellowship, grant number 17-PDF-364; K.V.S. was supported by the following awards: National Institutes of Health (NIH) National Institute of Neurological Disorders and Stroke (NINDS) Transformative Research Award R01NS076460, NIH National Institute of Mental Health Grant (NIMH) Transformative Research Award R01MH09964703, NIH Director's Pioneer Award 8DP1HD075623, Defense Advanced Research Projects Agency (DARPA) Biological Technology Office (BTO) "REPAIR" award N66001-10-C-2010, DARPA BTO "NeuroFAST" award W911NF-14-2-0013, the Simons Foundation Collaboration on the Global Brain awards 325380 and 543045, and the Howard Hughes Medical Institute.

## AUTHOR CONTRIBUTIONS

S.V., N.E.-C., and S.D.S. designed the study. S.V. performed the experiments, analyses, and wrote the manuscript with input and editing by all authors. S.I.R. performed the surgical implantations. P.N. provided significant guidance on data presentation. K.V.S. supervised all aspects of this work.

## DECLARATION OF INTERESTS

K.V.S. is a consultant to Neuralink Corp. and on the Scientific Advisory Boards of CTRL-Labs Inc. and Heal Inc. These entities did not support this work.

Received: August 16, 2017  
 Revised: November 21, 2017  
 Accepted: January 20, 2018  
 Published: February 15, 2018

## REFERENCES

- Afshar, A., Santhanam, G., Yu, B.M., Ryu, S.I., Sahani, M., and Shenoy, K.V. (2011). Single-trial neural correlates of arm movement preparation. *Neuron* 71, 555–564.
- Ajiboye, A.B., Willett, F.R., Young, D.R., Memberg, W.D., Murphy, B.A., Miller, J.P., Walter, B.L., Sweet, J.A., Hoyer, H.A., Keith, M.W., et al. (2017). Restoration of reaching and grasping movements through brain-controlled muscle stimulation in a person with tetraplegia: a proof-of-concept demonstration. *Lancet* 389, 1821–1830.
- Allen, W.E., Kauvar, I.V., Chen, M.Z., Richman, E.B., Yang, S.J., Chan, K., Gradinaru, V., Deverman, B.E., Luo, L., and Deisseroth, K. (2017). Global representations of goal-directed behavior in distinct cell types of mouse neocortex. *Neuron* 94, 891–907.e6.
- Ames, K.C., Ryu, S.I., and Shenoy, K.V. (2014). Neural dynamics of reaching following incorrect or absent motor preparation. *Neuron* 81, 438–451.
- Athalye, V.R., Ganguly, K., Costa, R.M., and Carmena, J.M. (2017). Emergence of coordinated neural dynamics underlies neuroprosthetic learning and skillful control. *Neuron* 93, 955–970.e5.
- Buch, E., Weber, C., Cohen, L.G., Braun, C., Dimyan, M.A., Ard, T., Mellinger, J., Caria, A., Soekadar, S., Fourkas, A., and Birbaumer, N. (2008). Think to move: a neuromagnetic brain-computer interface (BCI) system for chronic stroke. *Stroke* 39, 910–917.
- Chase, S.M., Kass, R.E., and Schwartz, A.B. (2012). Behavioral and neural correlates of visuomotor adaptation observed through a brain-computer interface in primary motor cortex. *J. Neurophysiol.* 108, 624–644.
- Chen, T.W., Li, N., Daie, K., and Svoboda, K. (2017). A map of anticipatory activity in mouse motor cortex. *Neuron* 94, 866–879.e4.
- Churchland, M.M., Afshar, A., and Shenoy, K.V. (2006). A central source of movement variability. *Neuron* 52, 1085–1096.

- Churchland, M.M., Cunningham, J.P., Kaufman, M.T., Foster, J.D., Nuyujukian, P., Ryu, S.I., and Shenoy, K.V. (2012). Neural population dynamics during reaching. *Nature* 487, 51–56.
- Cisek, P., and Kalaska, J.F. (2004). Neural correlates of mental rehearsal in dorsal premotor cortex. *Nature* 431, 993–996.
- Cunningham, J.P., and Yu, B.M. (2014). Dimensionality reduction for large-scale neural recordings. *Nat. Neurosci.* 17, 1500–1509.
- Denis, M. (1985). Visual imagery and the use of mental practice in the development of motor skills. *Can. J. Appl. Sport Sci.* 10, 4S–16S.
- Feltz, D.L., and Landers, D.M. (1983). The effects of mental practice on motor skill learning and performance: A meta-analysis. *J. Sports Psychol.* 5, 25–57.
- Gallego, J.A., Perich, M.G., Miller, L.E., and Solla, S.A. (2017). Neural manifolds for the control of movement. *Neuron* 94, 978–984.
- Ganguly, K., and Carmena, J.M. (2009). Emergence of a stable cortical map for neuroprosthetic control. *PLoS Biol.* 7, e1000153.
- Ganguly, K., Dimitrov, D.F., Wallis, J.D., and Carmena, J.M. (2011). Reversible large-scale modification of cortical networks during neuroprosthetic control. *Nat. Neurosci.* 14, 662–667.
- Gilja, V., Nuyujukian, P., Chestek, C.A., Cunningham, J.P., Yu, B.M., Fan, J.M., Churchland, M.M., Kaufman, M.T., Kao, J.C., Ryu, S.I., and Shenoy, K.V. (2012). A high-performance neural prosthesis enabled by control algorithm design. *Nat. Neurosci.* 15, 1752–1757.
- Golub, M.D., Yu, B.M., and Chase, S.M. (2015). Internal models for interpreting neural population activity during sensorimotor control. *eLife* 4, e10015.
- Golub, M.D., Sadtler, P.T., Oby, E.R., Quick, K.M., Ryu, S.I., Tyler-Kabara, E.C., Batista, A.P., Chase, S.M., and Yu, B.M. (2018). Learning by neural reassociation. *Nat. Neurosci.* 21, <https://doi.org/10.1038/s41593-018-0095-3>.
- Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci.* 21, 1229–1243.
- Huang, V.S., Haith, A., Mazzone, P., and Krakauer, J.W. (2011). Rethinking motor learning and savings in adaptation paradigms: model-free memory for successful actions combines with internal models. *Neuron* 70, 787–801.
- Hwang, E.J., Bailey, P.M., and Andersen, R.A. (2013). Volitional control of neural activity relies on the natural motor repertoire. *Curr. Biol.* 23, 353–361.
- Jarosiewicz, B., Chase, S.M., Fraser, G.W., Velliste, M., Kass, R.E., and Schwartz, A.B. (2008). Functional network reorganization during learning in a brain-computer interface paradigm. *Proc. Natl. Acad. Sci. USA* 105, 19486–19491.
- Johnson, S.H., Rotte, M., Grafton, S.T., Hinrichs, H., Gazzaniga, M.S., and Heinze, H.J. (2002). Selective activation of a parietofrontal circuit during implicitly imagined prehension. *Neuroimage* 17, 1693–1704.
- Kao, J.C., Stavisky, S.D., and Sussillo, D. (2014). Information systems opportunities in brain-machine interface decoders. *Proceedings of the IEEE* 102, 666–682.
- Kaufman, M.T., Churchland, M.M., Ryu, S.I., and Shenoy, K.V. (2014). Cortical activity in the null space: permitting preparation without movement. *Nat. Neurosci.* 17, 440–448.
- Krakauer, J.W., and Mazzone, P. (2011). Human sensorimotor learning: adaptation, skill, and beyond. *Curr. Opin. Neurobiol.* 21, 636–644.
- Krakauer, J.W., Pine, Z.M., Ghilardi, M.F., and Ghez, C. (2000). Learning of visuomotor transformations for vectorial planning of reaching trajectories. *J. Neurosci.* 20, 8916–8924.
- Liu, X., and Scheidt, R.A. (2008). Contributions of online visual feedback to the learning and generalization of novel finger coordination patterns. *J. Neurophysiol.* 99, 2546–2557.
- Liu, C., Adviser-Freeman, W.T., and Adviser-Adelson, E.H. (2009). Beyond pixels: exploring new representations and applications for motion analysis. In *Proceedings of the 10th European Conference on Computer Vision: Part III*, pp. 28–42.
- Makino, H., Ren, C., Liu, H., Kim, A.N., Kondapaneni, N., Liu, X., Kuzum, D., and Komiyama, T. (2017). Transformation of Cortex-wide Emergent Properties during Motor Learning. *Neuron* 94, 880–890.e8.
- Mathis, M.W., Mathis, A., and Uchida, N. (2017). Somatosensory cortex plays an essential role in forelimb motor adaptation in mice. *Neuron* 93, 1493–1503.e6.
- Miri, A., Warriner, C.L., Seely, J.S., Elsayed, G.F., Cunningham, J.P., Churchland, M.M., and Jessell, T.M. (2017). Behaviorally selective engagement of short-latency effector pathways by motor cortex. *Neuron* 95, 683–696.e11.
- Oby, E.R., Perel, S., Sadtler, P.T., Ruff, D.A., Mischel, J.L., Montez, D.F., Cohen, M.R., Batista, A.P., and Chase, S.M. (2016). Extracellular voltage threshold settings can be tuned for optimal encoding of movement and stimulus parameters. *J. Neural Eng.* 13, 036009.
- Orsborn, A.L., Moorman, H.G., Overduin, S.A., Shaneechi, M.M., Dimitrov, D.F., and Carmena, J.M. (2014). Closed-loop decoder adaptation shapes neural plasticity for skillful neuroprosthetic control. *Neuron* 82, 1380–1393.
- Pandarath, C., Gilja, V., Blabe, C.H., Nuyujukian, P., Sarma, A.A., Sorice, B.L., Eskandar, E.N., Hochberg, L.R., Henderson, J.M., and Shenoy, K.V. (2015). Neural population dynamics in human motor cortex during movements in people with ALS. *eLife* 4, e07436.
- Pandarath, C., Nuyujukian, P., Blabe, C.H., Sorice, B.L., Saab, J., Willett, F.R., Hochberg, L.R., Shenoy, K.V., and Henderson, J.M. (2017). High performance communication by people with paralysis using an intracortical brain-computer interface. *eLife* 6, e18554.
- Papaxanthis, C., Schieppati, M., Gentili, R., and Pozzo, T. (2002). Imagined and actual arm movements have similar durations when performed under different conditions of direction and mass. *Exp. Brain Res.* 143, 447–452.
- Paz, R., Boraud, T., Natan, C., Bergman, H., and Vaadia, E. (2003). Preparatory activity in motor cortex reflects learning of local visuomotor skills. *Nat. Neurosci.* 6, 882–890.
- Prsa, M., Galiñanes, G.L., and Huber, D. (2017). Rapid Integration of Artificial Sensory Feedback during Operant Conditioning of Motor Cortex Neurons. *Neuron* 93, 929–939.e6.
- Ranganathan, R., Wieser, J., Mosier, K.M., Mussa-Ivaldi, F.A., and Scheidt, R.A. (2014). Learning redundant motor tasks with and without overlapping dimensions: facilitation and interference effects. *J. Neurosci.* 34, 8289–8299.
- Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* 2, 661–670.
- Ryan, E., and Simons, J. (1982). Efficacy of mental imagery in enhancing mental rehearsal of motor skills. *J. Sport Psychol.* 4, 41–51.
- Sadtler, P.T., Quick, K.M., Golub, M.D., Chase, S.M., Ryu, S.I., Tyler-Kabara, E.C., Yu, B.M., and Batista, A.P. (2014). Neural constraints on learning. *Nature* 512, 423–426.
- Saposnik, G., Teasell, R., Mamdani, M., Hall, J., McIlroy, W., Cheung, D., Thorpe, K.E., Cohen, L.G., and Bayley, M.; Stroke Outcome Research Canada (SORCan) Working Group (2010). Effectiveness of virtual reality using Wii gaming technology in stroke rehabilitation: a pilot randomized clinical trial and proof of principle. *Stroke* 41, 1477–1484.
- Shabbott, B.A., and Sainburg, R.L. (2010). Learning a visuomotor rotation: simultaneous visual and proprioceptive information is crucial for visuomotor remapping. *Exp. Brain Res.* 203, 75–87.
- Shadmehr, R., Smith, M.A., and Krakauer, J.W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annu. Rev. Neurosci.* 33, 89–108.
- Sheahan, H.R., Franklin, D.W., and Wolpert, D.M. (2016). Motor planning, not execution, separates motor memories. *Neuron* 92, 773–779.
- Shenoy, K.V., and Carmena, J.M. (2014). Combining decoder design and neural adaptation in brain-machine interfaces. *Neuron* 84, 665–680.
- Shenoy, K.V., Sahani, M., and Churchland, M.M. (2013). Cortical control of arm movements: a dynamical systems perspective. *Annu. Rev. Neurosci.* 36, 337–359.
- Silvoni, S., Ramos-Murguialday, A., Cavinato, M., Volpato, C., Cisetto, G., Turolla, A., Piccione, F., and Birbaumer, N. (2011). Brain-computer interface in stroke: a review of progress. *Clin. EEG Neurosci.* 42, 245–252.

- Sobierajewicz, J., Szarkiewicz, S., Przekoracka-Krawczyk, A., Jaśkowski, W., and van der Lubbe, R. (2016). To what extent can motor imagery replace motor execution while learning a fine motor skill? *Adv. Cogn. Psychol.* *12*, 179–192.
- Stavisky, S.D., Kao, J.C., Ryu, S.I., and Shenoy, K.V. (2017a). Motor cortical visuomotor feedback activity is initially isolated from downstream targets in output-null neural state space dimensions. *Neuron* *95*, 195–208.e9.
- Stavisky, S.D., Kao, J.C., Ryu, S.I., and Shenoy, K.V. (2017b). Trial-by-trial motor cortical correlates of a rapidly adapting visuomotor internal model. *J. Neurosci.* *37*, 1721–1732.
- Suminski, A.J., Tkach, D.C., Fagg, A.H., and Hatsopoulos, N.G. (2010). Incorporating feedback from multiple sensory modalities enhances brain-machine interface control. *J. Neurosci.* *30*, 16777–16787.
- Tanaka, H., Sejnowski, T.J., and Krakauer, J.W. (2009). Adaptation to visuomotor rotation through interaction between posterior parietal and motor cortical areas. *J. Neurophysiol.* *102*, 2921–2932.
- Taylor, D.M., Tillery, S.I.H., and Schwartz, A.B. (2002). Direct cortical control of 3D neuroprosthetic devices. *Science* *296*, 1829–1832.
- Warner, L., and McNeill, M.E. (1988). Mental imagery and its potential for physical therapy. *Phys. Ther.* *68*, 516–521.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Organisms/Strains		
Rhesus macaque ( <i>Mucacca mulatta</i> )	Wisconsin and Yerkes Primate Centers	N/A
Software and Algorithms		
MATLAB	Mathworks	<a href="https://www.mathworks.com/products/matlab.html">https://www.mathworks.com/products/matlab.html</a>
Simulink RealTime	Mathworks	<a href="https://www.mathworks.com/products/simulink-real-time.html">https://www.mathworks.com/products/simulink-real-time.html</a>
Other		
Cerebus system	Blackrock Microsystems	<a href="http://blackrockmicro.com/neuroscience-research-products/neural-data-acquisition-systems/cerebus-daq-system/">http://blackrockmicro.com/neuroscience-research-products/neural-data-acquisition-systems/cerebus-daq-system/</a>
Utah microelectrode arrays	Blackrock Microsystems	<a href="http://blackrockmicro.com/electrode-types/utah-array/">http://blackrockmicro.com/electrode-types/utah-array/</a>
Polaris optical tracking system	Northern Digital	<a href="https://www.ndigital.com/medical/products/polaris-family/">https://www.ndigital.com/medical/products/polaris-family/</a>

### CONTACT FOR REAGENTS AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Saurabh Vyas ([smvyas@stanford.edu](mailto:smvyas@stanford.edu)).

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

Recordings were made from motor cortical areas of two male adult monkeys (*Macaca mulatta*), R (15 kg, 12 years old) and J (16 kg, 15 years old), while they performed an instructed delay cursor task in one of two contexts (Figure 1A). Use of two animals is standard practice in the field. Each monkey had two chronic 96-electrode arrays (1 mm electrodes, spaced 400  $\mu$ m apart, Blackrock Microsystems), one implanted in the dorsal aspect of the premotor cortex (PMd) and one implanted in the primary motor cortex (M1). The arrays were implanted 5 years and 7 years prior to these experiments for monkey R and J respectively. Voltage signals were band-pass filtered from each electrode (250 Hz – 7.5 KHz). These signals were processed to detect multi-unit “threshold crossing” spikes. We detected spikes whenever the voltage crossed below a threshold of  $-4.5$  times the root-mean-square voltage. In this study, we do not spike sort, or assign spikes to individual neurons. In our view, this is justified due to three reasons: first, multi-unit spike trains are the standard signal used in the BMI literature; second, our scientific claims rely on the motor-cortical population activity, which includes both single- and multi-unit activity; finally, decoding both multi-unit spikes as well as well-isolated single units can potentially increase the amount of information available on chronically implanted electrode arrays (Pandarinath et al., 2015, 2017; Oby et al., 2016; Stavisky et al., 2017a). Animal protocols were approved by the Stanford University Institutional Animal Care and Use Committee.

### METHOD DETAILS

#### Task design

Monkeys performed instructed-delay movements in one of two contexts as described in Figure 1A. Our standard methods have been previously described (Gilja et al., 2012; Shenoy et al., 2013; Ames et al., 2014). In the ‘overt’ context, both monkeys performed a Radial 8 Task, where they reached using their contralateral-to-arrays arm in order to move a computer cursor in virtual reality (latency of  $7 \pm 4$  ms). Eight targets were arranged radially in a 2D circle, along with an additional target at the center of the circle. Monkeys started by holding the cursor on the central target continuously for 500 ms. After a variable instructed delay period (sampled uniformly from 400 – 800 ms), monkeys moved the cursor within a  $4 \times 4$  cm acceptance window of the cued target. This target also had to be held continuously for 500 ms. The target changed color to signify the hold period. If the cursor left the acceptance window, the timer was reset, but the trial was not immediately failed. Monkeys had 2 s to acquire the target. Success was accompanied with a liquid reward, along with a success tone. Failure resulted in no reward, and a failure tone. Regardless, the center target was then presented,

which the monkeys had to acquire in a similar fashion as above. The period starting from the instructed delay and ending with the target acquisition (or failure) constituted a trial.

In order to train monkeys to perform the instructed-delay task, we used the following protocol. First, we jittered the cued target during the entire instructed delay period as a means to indicate the beginning and end of the period. Second, monkeys were required to limit cursor velocity to a maximum of 1 mm/s during the delay period. Violating this condition automatically resulted in a failure. During training, we increased the time between a failed trial and the beginning of the next trial to 5 s to discourage failures. By comparison, the time between a successful trial and the subsequent trial was 20 ms. In the ‘covert’ context, the same task design was used, with the exception that monkeys’ arms were comfortably restrained and they had to control the cursor velocity using a BMI by modulating their neural activity.

The key manipulation on the Radial 8 Task was to introduce a visuomotor rotation (VMR), parametrized by angle  $\theta$ . This consisted of rotating the cursor position  $\theta$  degrees counter-clockwise around the workspace origin. In order to acquire targets in these cases, monkeys had to move their arm (overt context) or modulate neural activity (covert context) in a fashion that would ordinarily move the cursor at an angle  $-\theta$  relative to the cued target. Thus, monkeys had to apply e.g., a clockwise rotation in order to counter the effect of the counter-clockwise VMR. Monkeys received constant visual feedback, so they could correct their cursor trajectories during learning. We used four VMR angles in this study:  $-45^\circ$ ,  $45^\circ$ ,  $60^\circ$ , and  $90^\circ$ .

In order to encourage the monkeys to “try hard” to adapt to the VMR (rather than accepting making highly curved and inefficient movements to targets), we employed the following strategy. First, we decreased the maximum reach time every 500 trials (on average) from 10 s initially to 1.5 s in the last 500 trial block. In later sessions, we started from 3 s. Second, we introduced a path efficiency check, where we automatically failed a trial if the maximum orthogonal deviation (relative to the straight-line distance to the cued target) exceeded a bound. The experimenter manipulated this bound, making it more challenging, as a function of time. This factor in particular played a critical role in eliciting precise behavior which led to the effects described in [Figure 1](#).

In order to perform the experiments described in [Figure 1](#) (experimental flow in [Figure 1B](#)), we used the following protocol. For the control conditions, monkeys were arm-restrained and used a BMI (under no VMR) for  $2000 \pm 1000$  trials of the Radial 8 task. This typically took 90 min. After this adaptation phase, the experimenter went into the monkey’s room and removed the restraint of the arm contralateral to the arrays. This readied the overt context (still with no VMR), in which the monkey performed the Radial 8 task for 500 trials. For the experimental conditions the same procedure as the control experiment was followed, with the one change that during BMI use, a VMR was introduced. Both control and the experimental conditions were performed within the same experiment session. We analyzed a minimum of four sessions for each analysis, with alternating order of control and experimental blocks. The experiments described in [Figure 4](#) followed a similar flow, with two changes. First, the control comprised of only overt arm reaches in the presence of a VMR. Second, the experiment comprised of first using the BMI (in the presence of the same VMR as the control) for  $2000 \pm 1000$  trials of the Radial 8 task, followed by a block of overt reaches in the presence of the same VMR. The precise number of trials for all of these experiments was based in part on the monkey’s degree of VMR adaptation as qualitatively assessed by the experimenter at each session. To minimize day-to-day or residual adaptation effects, monkeys started and ended each session with overt arm reaches without VMR. Sessions were excluded from analysis if effects from savings or interference were observed during the first 500 trial baseline block.

### BMI decoder fitting

In order to train the BMI decoder at each session, we adapted the recalibrated feedback-intention trained Kalman filter (ReFIT) procedure ([Gilja et al., 2012](#)). At the start of each session, monkeys observed 200 trials of Radial 8 automated cursor movements from the center of the workspace to one of 8 radially arranged targets at a distance of 12 cm. We performed three such blocks of 200 trials, each block with cursor velocity of 8, 10, and 12 cm/s. We used the neural and kinematic data from these blocks to estimate a position and velocity Kalman filter. Since BMI performance is sensitive to velocity gain, we manipulated this gain (i.e., scaling from neural activity to cursor velocity) on a session-by-session basis so as to help the monkey balance three factors: being able to successfully hold the center during the delay period (which benefits from low gain), moving in as straight a line as possible during the movement period (which benefits from low gain), and reaching the target as quickly as possible (which benefits from high gain). This procedure was followed for all Monkey J sessions.

Monkey R has poorer signal quality, and thus on roughly 50% of the sessions, the initial decoder (estimated the same as was done for Monkey J) was used by monkey R in closed-loop to perform another 200 Radial 8 task trials. These data were used by the final Kalman Filter decoder, estimated using the ReFIT algorithm. We again manipulated the gain of the decoder to strike a balance between performance and ease of accomplishing the instructed delay period (i.e., staying below the 1 mm/s requirement during the delay period). For the other 50% of sessions for Monkey R, we built a decoder in a similar fashion as was done for Monkey J. We found no difference in the scientific findings by using the ReFIT decoder (versus the standard Kalman filter) for Monkey R. In our decoders, no distinction was made between PMd and M1; all neural data were used jointly.

### Residual movement tracking

While monkeys performed covert movements (via the BMI), we tracked the position of various body parts using infrared video cameras at a rate of 24 frames per second. Three cameras were positioned such that the dorsal forearm, the hand, the fingers, the rhomboids and deltoids regions, the pectorales and deltoid regions, and the shoulder, biceps, and triceps regions were visible,



and took up most of the field-of-view. We used an open source implementation of the Lucas-Kanade optical flow algorithm (Liu et al., 2009) in order to estimate a velocity for each pixel from frame-to-frame. In order to ascribe a single velocity for each frame, we averaged the horizontal and vertical components (individually) of the velocity for all pixels in the field-of-view. We used intentionally placed physical landmarks visible in the videos to convert pixel coordinates into real-world coordinates. Finally, we used the millisecond-level timestamp from each frame to associate it with each individual trial. This analysis is presented in Figure S1.

An alternative way to perform these control experiments could be to use an EMG preparation as done in many of our prior studies, e.g., (Kaufman et al., 2014). After consideration, we believe that a computer vision approach is more appropriate in this study primarily because we want to *simultaneously* monitor multiple muscles. While it is possible that the animals' muscles sometimes undergo isometric co-contractions, which would result in minimal visible movement while still producing muscle activity, this is quite unlikely for three primary reasons: (1) given the number of hours of BMI experiments the animals must do, it is quite unlikely that they sustain co-contractions for that long as it is energetically expensive (i.e., very tiring), (2) even if there are isometric co-contractions, the shape of the muscles changes slightly, which can be measured at our camera working distance, which has hundreds of pixels for each anatomy of interest, and finally (3) co-contracting would not help the animals anyway because in order to affect the cursor movement, the muscle activity would need to systematically vary as a function of target direction; such distinct patterns of co-contractions would be picked up by our cameras. Thus, we believe that our approach well assesses the degree to which correlated movements are made during the covert context, and has the key advantage over EMG of being less likely to "miss a muscle" when measuring.

### Preparatory neural state analysis

All of the analyses in Figure 2 and Figure 3 examine motor cortical preparatory neural states using a standard application of Principal Components Analysis (PCA). In brief, neural data were arranged into a data matrix comprising of neural firing rates of every neuron for every condition for every time point. We only looked at 200 ms worth of time points at the end of the instructed delay period. We then applied PCA on this data matrix. This results in a low-dimensional representation capturing the naturally occurring co-modulation in the data. Figure 2B is a visualization produced by applying this procedure on a block of no-VMR overt trials. For each condition, trials (averaged in a bin of size three) were projected onto the top two PCs. These are 500 baseline trials that were collected at the beginning of the experimental session.

In the behavioral data (Figure 2A) we observed that overt no-VMR trials following a block of covert trials with a VMR initially erred in the direction corresponding to the adjacent target in the opposite direction of the VMR. In order to do a similar analysis on the neural data, we fit a line between the centroids of neural data (recorded while the animal prepared reaches) to pairs of targets from Figure 2B. We then projected trial-averaged (5 trials) firing rates corresponding to overt reaches (after adapting to a VMR under the covert context), directly onto that line. These projections are shown in Figure 2C. We established a normalized distance metric so that trials can be combined across reach conditions and sessions. Concretely, trials that were projected directly onto the centroid corresponding to the cued target are assigned a distance of 0, whereas trials projected directly onto the centroid of the target in the opposite direction of the cued target are assigned a distance of 1. Figure 2D shows this visually for all eight conditions, and Figure 2E shows the statistics across all reach conditions and sessions.

In Figure 3A we use PCA in the same fashion to compute the cumulative variance captured by each individual PC. We find that 4 PCs (in Monkey R) and 6 PCs (in Monkey J) capture over 90% of the variance in the data. In order to find the amount of shared neural variance between overt and covert behaviors, we project the 4 PCs (6 in Monkey J) from the overt context onto the 4 PCs (6 in Monkey J) of the covert context, and compute the inner product, normalized by the amount of total variance captured by each PC. Vice-versa gives the shared variance between covert and overt. A similar procedure was performed (not reported here) where both sets of data were combined and a joint subspace was found. We found no significant difference between the joint analysis and the analysis presented in Figure 3. The same procedure was followed for the pairwise comparisons for the 'watch' condition. The same procedure was followed for Figure S4A.

### Statistics

For all histograms (Figure 1, Figure 4, and Figure S2), the significances of the differences in the distributions were determined with two-tailed Student's *t* tests, assuming non-equal variances of the two samples. We confirmed that each histogram followed a normal distribution using the Kolmogorov-Smirnov test. For data that did not follow a normal distribution, we used the Wilcoxon rank-sum test (Figure 2, Figure 3, and Figure S3), using the paired (i.e., signed) test where appropriate. For all linear regressions (Figure 1, Figure 3, Figure 4, Figure S1, and Figure S3) we used the *F*-statistic to assess the significance level of the slopes being different from zero. Partial correlations were used to rule out influence from other experimental parameters. In Figure S3, we compared the slopes using a two-tailed Student's *t* test. In Figure 3 we measured the statistical overlap between the populations using the Bhattacharya coefficient, normalized such that 1 indicates no statistical overlap. For all tests, we used  $p = 0.05$  as the significance threshold.