The role of reinforcement and contextual cues in the acquisition and expression of motor memories

By

Sarah E. Pekny

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

April 2015

Abstract

Each day we interact with a wide variety objects, from using our computer, to driving a car or preparing a cup of coffee. We can complete the actions necessary to use these objects with little to no error or relearning from the last time we completed the same task. As the phrase ‘just like riding a bicycle’ implies, once learned, certain actions can always be called upon, regardless of how infrequently we execute them.

In the following chapters, we focus on the initial learning of goal-directed motor commands, investigating how people first learn to interact with a novel object and what memory they retain from this interaction. First, we investigated what cues serve as indicators that a particular set of motor commands should be retained for future recall. Our results suggested that decreases in reinforcement signaled participants of a change in the dynamics of the tool, allowing participants to separate and retain multiple motor memories for use of the same tool. From these experiments, we determined that reinforcement of actions also served as a critical cue to recall upon these motor memories. When reinforcement of the current motor commands was withheld, participants switched their actions and recalled previous motor training. We next investigated if the retention of a particular set of motor commands was specific to the tool on which these commands were learned. Here, we found that when participants encountered a new tool similar to the one they had used for training, they relied upon their memory of the trained tool, and generalized some of their previous learning. We then asked if the contents of motor memory were stable or if modifications occurred with continued training. The results of these experiments suggested that with sufficient time away from practice, motor memories become more efficient and we begin to minimize energetic inefficiencies in our movements. Finally, we revisited the idea of reinforcement and action selection. We found that patients diagnosed with Parkinson's Disease (PD), were less sensitive to a lack of reinforcement, which in turn lead to less motor exploration in these patients as compared to healthy controls.

Thesis advisor: Dr. Reza Shadmehr
Thesis committee: Drs Reza Shadmehr, Amy Bastian and Jeffrey Ellenbogen

Acknowledgements

I would like to thank Dr. Reza Shadmehr, the members of our lab and my thesis committee for all of the input and guidance throughout my research. I would also like to thank my family, Mollie and Mike.

Table of Contents

Abstract ii
Acknowledgements iv
List of Tables viii
List of Figures ix
Chapter 1: Introduction 1
  1.1 Assaying the components of motor memory 1
  1.2 Understanding tool-specificity in motor adaptation 2
  1.3 Effects of time on motor memory 4
  1.4 Role of reinforcement in motor exploration 5
Chapter 2: Protection and expression of human motor memories 6
  2.1 Introduction 6
  2.2 Materials and Methods 9
    2.2.1 Experimental Groups 10
    2.2.2 Data analysis 16
    2.2.3 Model 17
  2.3 Results 18
    2.3.1 Experiment 1: memories are protected from unlearning 23
    2.3.2 Experiment 2: protection despite paucity of sudden errors to signal a contextual change 26
    2.3.3 Sudden change in probability of success as a possible contextual cue 29
    2.3.4 Experiment 3: spontaneous recovery of a motor memory following a sudden change in performance errors 31
    2.3.5 Experiment 4: spontaneous recovery of a motor memory following withholding of reward for the competing memory 34
  2.4 Discussion 37
    2.4.1 The multiple components of motor memory 38
    2.4.2 Link to operant conditioning 44
    2.4.3 Limitations 44
Chapter 3: Are motor memories tool-specific? 46
  3.1 Introduction 46
  3.2 Methods 48
    3.2.1 Robotic Tools 48

3.2.2 Behavioral Task 51
3.2.3 Experimental Groups 52
3.2.4 Data Analysis 54
3.3 Results....56
3.3.1 Force field learning transferred across tools....57
3.3.2 Interference was also observed during force field training....62
3.3.3 Tool switching did not reduce interference 66
3.3.4 With repeated switching, participants reduced across-tool transfer 68
3.4 Discussion 70
Chapter 4: Optimizing effort: increased efficiency of motor memory with time away from practice....75
4.1 Introduction....75
4.2 Methods....78
4.2.1 Experimental Groups 81
4.2.2 Data Analysis 83
4.3 Results....84
4.3.1 Learning generalized to the probe target 85
4.3.2 Time away from practice was necessary for the reduction of task-irrelevant forces....87
4.3.3 Sleep was not necessary for reduction of the task irrelevant forces 92
4.3.4 Confounding variables....95
4.3.5 Practice, but not passage of time, improved the task-relevant forces....95
4.3.6 Task irrelevant forces were minimized if they had kinematic consequences....99
4.4 Discussion 100
Chapter 5: Reward dependent modulation of movement variability....107
5.1 Introduction....107
5.2 Methods....109
5.2.1 Behavioral Task 110
5.2.2 Data Analysis 112
5.3 Results....113
5.3.1 Baseline reach variability was comparable between groups....113
5.3.2 Reach variability increased following an unrewarded trial 115
5.3.3 A relationship between probability of reward and reach variability....121
5.3.4 Measuring sensitivity to history of reward 123
5.4 Discussion 127

Chapter 6: Conclusions 133
References..136
Curriculum Vitae 144

List of Tables

Table 1. Clinical characteristics of volunteers with Parkinson’s disease…………..132

List of Figures

Figure 2.1 Experimental design 15
Figure 2.2 Simulation results 22
Figure 2.3 Motor output in the error-clamp block. 25
Figure 2.4 Error magnitudes and reward probabilities 29
Figure 2.5 Spontaneous recovery in Anb..........33
Figure 2.6 Withholding reinforcement produces spontaneous recovery 36
Figure 2.7 A conceptual model of motor memory during BA paradigms. 41

Figure 3.1 Robotic tools and experimental paradigms 50
Figure 3.2 Force field training generalized across robotic tools..........62
Figure 3.3 Field performance improved with repeated switching 65
Figure 3.4 Performance of NoSwitch + Interference group 67
Figure 3.5 Transfer decreased across many sessions. 69

Figure 4.1 Task set-up and experimental groups. 79
Figure 4.2 Generalization to the probe target 86
Figure 4.3 Forces produced to probe target reduced with time, not practice..........91
Figure 4.4 Sleep is not necessary for the reduction of probe forces. 94
Figure 4.5 Changes in speed and accuracy of movements 98
Figure 4.6 Downward segment of reach 100

Figure 5.1 Experimental task and set up..........112
Figure 5.2 Participants respond differentially to positive vs. negative task outcomes..........117
Figure 5.3 PD participants are less sensitive to negative outcomes 120
Figure 5.4 Reward based modulation of variability is impaired in PD 123
Figure 5.5 A state-space model..........126

Chapter 1: Introduction

1.1 Assaying the components of motor memory

One puzzle that has persisted for years in motor control is the question of whether adaptation to a perturbation followed by adaptation to another perturbation leads to catastrophic interference of the initial motor memory. Traditionally, this question has been investigated by studying A-B-A type paradigms. For these protocols, participants first train in one condition *A*, next learn another condition *B*, and then are retested under condition *A*. If performance during retest is similar to naïve, this has been interpreted as the initial memory of *A* being destroyed.

In the experiments described in Chapter 2, we proposed that perhaps the initial memory of *A* is not destroyed, but instead is masked by the competing memory *B*. To study this hypothesis, we trained participants in a reaching task under a series of A-B-A type paradigms, applying opposing velocity-dependent force fields to the participant's hand. However, instead of studying performance during retest on field A, we introduced a new tool to assay the contents of motor memory. After recalling the memory of *A*, we placed the participants in a long block of error-clamp trials, in which the hand was constrained to a straight path to and from the target. We found that the motor output in this long block of error clamp trials was biased towards *A*, suggesting that this memory was protected during subsequent de-adaptation or reverse-adaptation.

Through a series of experiments, we also asked which cues allowed for protection and later expression of these motor memories. In other studies of interference, passage of time and contextual cues have been utilized to help reduce inference between learning opposing conditions. In the experiments of Chapter 2, we did not rely on contextual cues and only used 3 minute set breaks between opposing conditions, but saw evidence of retention of both motor memories. We first discovered that protection from unlearning and expression of memories did not dependent on large performance errors. Participants were able to learn and recall gradually applied perturbations, in which they never experienced large performance errors. Also, errors in the opposite direction of the learned field were enough to recall previous training. Rather, we found that reinforcement appears to be a critical cue that affords protection to motor memories. Additionally, withholding reward encouraged participants to alter their motor commands, resulting in the retrieval of previous field training.

1.2 Understanding tool-specificity in motor adaptation

Through our experiments in Chapter 2, we found that people are able to acquire multiple motor memories. These conclusions were based on the analysis of the motor biases we observed in long blocks of error-clamp trials following adaptation. Importantly, we noticed that the entire history of training was reflected in this motor bias. Though this suggested that all motor memories were retained, we were unable to explicitly recall only one of those memories.

In the work described in Chapter 3, we again trained participants in opposing force field conditions *A* and *B*, but we wanted to find a method by which participants could recall these memories independently of each other. Our idea was that acquisition of a motor memory may be linked to the tool on which training occurred, and by pairing these conditions with separate tools, we could recall a single motor memory. In our lab, we have two separate robotic tools. Each tool is housed in its own room, which differs in location, square footage, set-up, ambient lighting, and extraneous furniture and lab equipment. Additionally, these two robotic tools are not identical in their construction. For example, one robot is a unimanual manipulandum with a soft white grip on the handle and a dark surround blocking the participant from viewing the lab. The other robot is a bimanual manipulandum with an exposed metal grip, and no surround blocking the participant.

Through a series of experiments, we found that motor memories were not linked to the tool on which they were acquired. Instead, participants transferred learning between robots. With repeated switching, this transfer decreased. However, even after 14 sessions, participants were unable to learn the pairing between the robot being used, and the motor memory that would lead to successful countering of the applied conditions.

1.3 Effects of time on motor memory

In Chapters 2 and 3, we investigated the role of various cues and paradigms to achieve protection, recall, and separation of multiple motor memories. In these tasks, we applied force fields to the participant's hand, and in order to achieve success, the participant had to exert forces to counter these perturbations. What we defined as a 'motor memory' was the above-baseline exertion of these learned forces outside of the training environment. In Chapter 4, we wanted to probe the contents of this motor memory, to understand if participants recalled the exact motor commands they had learned during training.

For these experiments, we measured the generalization of the force field training, by measuring the transfer of learning to another target in a different workspace. Though generalization to other movements with similar arm states (position, velocities, etc) is usually considered a hallmark of learning, we designed our task so that this phenomena resulted in an inefficient expenditure of energy. We found that with significant time away from practice (6 or 24 hours but not 3 or 30 minutes) participants decreased the amount of generalization that they expressed overall. This decrease was linked to time away from practice, and not practice itself, as doubling the amount of training without a rest period did not result in a decrease of generalization.

1.4 Role of reinforcement in motor exploration

In Chapter 2, we made the crucial discovery that withholding reinforcement encouraged participants to recall previous training. This relied on the ability of the participant to alter their motor commands once the current motor commands were no longer rewarding. We predicted that if participants are unable to respond to successes and failures, then this process of action selection would be impaired, and reinforcement would no longer play a major role in switching between motor commands.

To study this hypothesis, in Chapter 5 we analyzed the trial-to-trial variability of movements in response to high and low periods of reward feedback. In healthy controls, we found that during periods of low reward probability, reach variability was high. This indicated that following an unsuccessful trial, participants altered their motor commands, presumably in an effort to search for a more rewarding solution. However, during periods of high reward probability, participants maintained their motor commands, and reach variability was low. We then recruited a group of patients diagnosed with Parkinson's Disease (PD) a known example of basal ganglia dysfunction. We found that these participants had levels of reach variability comparable to controls during periods of high reward. However, following unsuccessful trials, the PD patients did not alter their motor commands, indicated an impairment in their ability to modulate motor commands in response to negative feedback.

Chapter 2: Protection and expression of human motor memories[1]

2.1 Introduction

When adaptation to a perturbation is followed by reverse-adaptation, does the brain protect the memory that was acquired during adaptation, or do the errors during reverse-adaptation continue to modify the previously acquired memory? In a typical experiment, a target is presented and the subject produces a movement that is perturbed by amount *A*. With training, the subject learns to respond to that target by producing motor commands that approximately cancel *A*. Let us call the result of this training "the motor memory for *A*". Now suppose that the perturbation is changed to *B*. With training, in response to the same target the brain produces motor commands that cancel *B*. The central question is whether this learning destroys the memory of *A*. Despite a half century of research (Lewis et al., 1952;McGonigle and Flook, 1978;Shadmehr and Brashers-Krug, 1997;Caithness et al., 2004;Krakauer et al., 2005), this question remains unresolved.

For example, Caithness et al. (2004) trained reaching movements in the *AB* paradigm and then retested in *A*. They found that performance was no different than naïve. They wrote: "when people successively encounter opposing transformations (*A* then *B*) of the

---

[1] This work in this chapter has been published in The Journal of Neuroscience: Pekny SE, Criscimagna-Hemminger SE, and Shadmehr R (2011) Protection and expression of human motor memories. J Neurosci 31: 13829-13839

same type (e.g., visuomotor rotations or force fields), memories related to *A* are reactivated and then modified while adapting to *B*." Indeed, there is currently little evidence for protection of *A* in the *AB* paradigm. The same question, however, has produced unequivocal results in other fields of memory research: experiments in classical conditioning (Medina et al., 2001;Stollhoff et al., 2005) suggest that memories that are produced during adaptation are protected during reverse-adaptation (termed extinction). Indeed, extinction training is believed to produce a distinct memory that competes with the original (reviewed in (Bouton, 2002)).

Why have the two fields of research arrived at different conclusions? A major difference is the method that has been used to assay memory. While in the motor learning literature the assay of memory is savings (i.e., faster re-learning), in classical conditioning the assay is *spontaneous recovery*. Spontaneous recovery refers to the observation that extinction training returns performance to baseline, but with passage of time the brain re-expresses the adaptation memory in response to the stimulus. This is taken as evidence that the extinction memory masks expression of the adaptation memory, but passage of time dissolves this mask.

Spontaneous recovery has also been observed in motor learning, but its occurrence has not been viewed as evidence for protection of memories. For example, when a long period of *A* is followed by a brief period of *B*, motor output returns to baseline. However, in the following error-clamp trials in which performance errors are clamped to zero (Scheidt et al., 2000), motor output rises from baseline toward *A* (Smith et al., 2006;Ethier et al., 2008;Criscimagna-Hemminger and Shadmehr, 2008). One view is that this is evidence for multiple timescales of a single context memory (Smith et al., 2006) in which training in *B* can destroy the memory of *A* (Sing and Smith, 2010). Another view is that upon transition to *B*, at least some component of *A* is protected (Lee and Schweighofer, 2009), but then the error-clamp trials make the brain uncertain regarding context, resulting in expression of both *A* and *B*.

Is the memory of *A* protected during learning of *B*? If so, how can we encourage the brain to express it? We show that when the brain is expressing one motor memory, it will switch and express another motor memory merely because of sudden reductions in probability of success. This produces spontaneous expression of a previously acquired memory. In the *AB* paradigm, at least part of the *A* memory is protected and can be retrieved through manipulation of reinforcement.

2.2 Materials and Methods

We recruited 99 neurologically intact, right-hand dominant participants for our study (24.2±4.6 years, mean ± SD, including 45 males and 54 females). All volunteers were naïve to the paradigm and the purpose of the experiment. Each volunteer signed a consent form approved by the Johns Hopkins University School of Medicine Institutional Review Board. Subjects were trained in a force field (Shadmehr and Mussa-Ivaldi, 1994a). They held a two-joint robotic manipulandum while making point-to-point reaching movements with their right hand. The hand was covered by a screen and visual feedback was provided by a small cursor (5 x 5 mm) representing the actual hand position. Participants reached out from a 1 cm^2 starting point to a 1 cm^2 target positioned at a distance of 10 cm directly along the body midline. The trial was considered a success if the movement was completed within 500±50 ms, and the tangential velocity during the reach did not exceed 0.55 m/s or fall below 0.20 m/s. If the trial was successful, the target was animated to represent an explosion. Otherwise, there was no explosion. The explosion was the only 'reinforcement'. Subsequently the robot brought the hand back to the center starting point to indicate start of the next trial. We recorded force at the handle, as well as position and velocity of the hand at a rate of 100Hz. Volunteers were allowed a 2-min break at intervals of 192 trials. A typical experiment lasted ~2hrs.

We placed error-clamp trials (Scheidt et al., 2000) randomly in the baseline and adaptation phases with one-eighth probability (no error-clamp trials were present during brief 20 trial periods, see below). Immediately following training, all groups were given a block of 364 error-clamp trials. During error-clamp trials, the motion of the hand was constrained to a straight line “channel” to the target by a stiff one-dimensional spring (spring coefficient = 2500 N/m; damping coefficient = 25 N.s/m) that counteracted forces perpendicular to the direction of the target. As in all trials, the target was animated (indicating success) if the hand reached the target within 500±50 ms.

2.2.1 Experimental Groups

Our experiments were designed to answer three questions: 1) Does adaptation to a perturbation followed by an equal period of washout or reverse-adaptation result in catastrophic destruction of the motor memory? 2) If motor memories are protected from unlearning, does this protection depend on a contextual cue associated with large errors that arise from a change in the perturbation? And 3) under what conditions do previously acquired motor memories show spontaneous expression?

Experiment 1. We performed this experiment in order to ask whether adaptation to a perturbation followed by an equal period of washout or reverse-adaptation resulted in catastrophic destruction of the motor memory. We trained four groups of subjects in protocols that are illustrated in Fig. 1A. All protocols began with a null field for 192 trials during which no forces were imposed on the hand. Field *A* was clockwise curl in which forces on the hand were dependent on hand velocity $\dot{\mathbf{x}}$ via a viscosity matrix $\begin{bmatrix} 0 & 13; & -13 & 0 \end{bmatrix}$ N.s/m. Field *B* was a counter-clockwise curl. In the BNb group (n=9) we tested whether null training could erase previous training in *B*. Participants were trained in *B* for 384 trials, followed by null for 384 trials, followed by *B* for 20 trials, and then error-clamp for 364 trials. We compared the forces that the BNb group produced in the error-clamp trials with group Nb (n=9). The hypothesis was that if null training produced unlearning of the memory acquired in prior training in *B*, then the forces that subjects produced in error clamp trials following 20 trials in *B* should be identical in the BNb and Nb groups. Alternatively, if the brain protected the memory of *B* during the null training, perhaps it would express this protected memory following brief re-exposure to *B*. Next, we tested this same question in a different protocol. In the BAb group (n=9), we tested whether training in the opposite force field could destroy previous training in *B*. Participants were trained in *B* for 384 trials, followed by training in *A* for 384 trials, followed by *B* for 20 trials, and then error-clamp for 364 trials. We compared the forces that the BAb group produced in the error-clamp trials with group Ab (n=9).

Experiment 2. The results of Exp. 1 suggested protection of memories. A number of previous models have suggested that sudden large errors that occur when the perturbation is changed from one direction to the opposite direction alert the brain that the context has changed, resulting in protection of the currently active memory and spawning of a new memory (Jacobs et al., 1991;Haruno et al., 2001). Here, we performed an experiment in order to ask whether protection of motor memories depended on a contextual cue associated with large errors that arise from a sudden change in the perturbation. We trained four groups of subjects in protocols that are illustrated in Fig. 1B. In the B_gNb group (n=9) we presented *B* gradually over 96 trials, maintained it at full strength for 192 trials, and then gradually returned it to null over 96 trials. The null training continued for another 384 trials, and then *B* was re-introduced at full strength for 20 trials, followed by error-clamp for 364 trials. We compared the forces that the B_gNb group produced in the error-clamp trials with group Nb (n=9). In the B_gAb group (n=9), the gradual presentation of *B* was followed by training in *A* for 384 trials, followed by 20 trials in *B*, and then 364 error-clamp trials. We compared the forces that the B_gAb group produced in the error-clamp trials with group Ab (n=9).

Analysis of the data in the groups that learned *B* gradually demonstrated protection of the *B* memory. To further examine the nature of this protection, we recruited a new group of subjects and introduced field *B* even more gradually. In the B_{gg}Nb group (n=9) we presented *B* gradually over 192 trials, maintained full strength for 192 trials, and then gradually returned to null over 192 trials. The null training continued for another 192 trials, and then *B* was re-introduced at full strength for 20 trials, followed by 364 error-clamp trials.

Experiment 3. In experiments 1 and 2 we observed that 20 *B* trials produced spontaneous recovery of the previously acquired *B* memory, i.e., revisiting *B* produced partial recall of the previously acquired *B* memory, despite the intervening washout and reverse-adaptation. What were the cues that encouraged the brain to express a previously protected motor memory in the error-clamp block? Initially, we imagined that perhaps the critical cue was the fact that the forces in 20 *B* trials were the same as the forces in the initial *B* training. That is, perhaps the forces that subjects experienced in the 20 *B* trials acted as a cue that produced expression of the previously acquired *B* memory. To check for this, we trained subjects in ANb. In the ANb group (n=9), we presented *A* for 384 trials, followed by null for 384 trials, 20 *B* trials, and then 364 error-clamp trials (Fig.

1C). We compared the forces that the ANb group produced in error-clamp trials with the Nb group (n=9). If the critical cue for expression of a memory was similarity between the perturbation forces in the acquisition and re-exposure periods, then subjects should not express *A* in the error-clamp trials after brief exposure to *B*. In fact, we found that the subjects expressed some of the *A* memory in the error-clamp trials.

Experiment 4. One way to account for the results of the above experiments is to imagine that the sudden changes in movement error and probability of success that accompanied the 20 *B* trials made the brain uncertain regarding which motor output was appropriate, *A*, *B*, or null. Perhaps it was this uncertainty that resulted in expression of the previously acquired memory. To test this idea, we trained subjects in *B*, and then after washout, attempted to make our subjects uncertain through artificially manipulating probability of error and/or probability of success. After we manipulated this uncertainty during 20 trials, we then presented them with the usual error-clamp block. In the B_gNR group (n=9) we presented *B* gradually over 96 trials, maintained it at full strength for 192 trials, and then gradually returned it to null over 96 trials (Fig. 1D). Following 96 null trials, a random field was presented for 20 trials. This random field consisting of 7 trials of *A*, 7 trials of *B*, and 6 null trials, randomly interspersed. This was followed by 364 error-clamp trials. We compared the forces that the B_gNR group produced in the error-clamp block with the B_gN group (n=9). The B_gN group did not receive the random field before the error-clamp trials.

To dissociate whether uncertainty arose from sudden changes in the probability of error vs. sudden changes in the probability of success, we considered a final group of subjects. In the B_gNS group (n = 9) we presented *B* gradually over 96 trials, maintained it at full strength for 192 trials, and then gradually returned it to null over 96 trials (Fig. 1D). After an additional 96 null trials, we presented 20 error-clamp trials but withheld reinforcement (target explosions) even if velocity and performance time were within the acceptable limits. This was followed by 364 error-clamp trials with the usual success requirements.

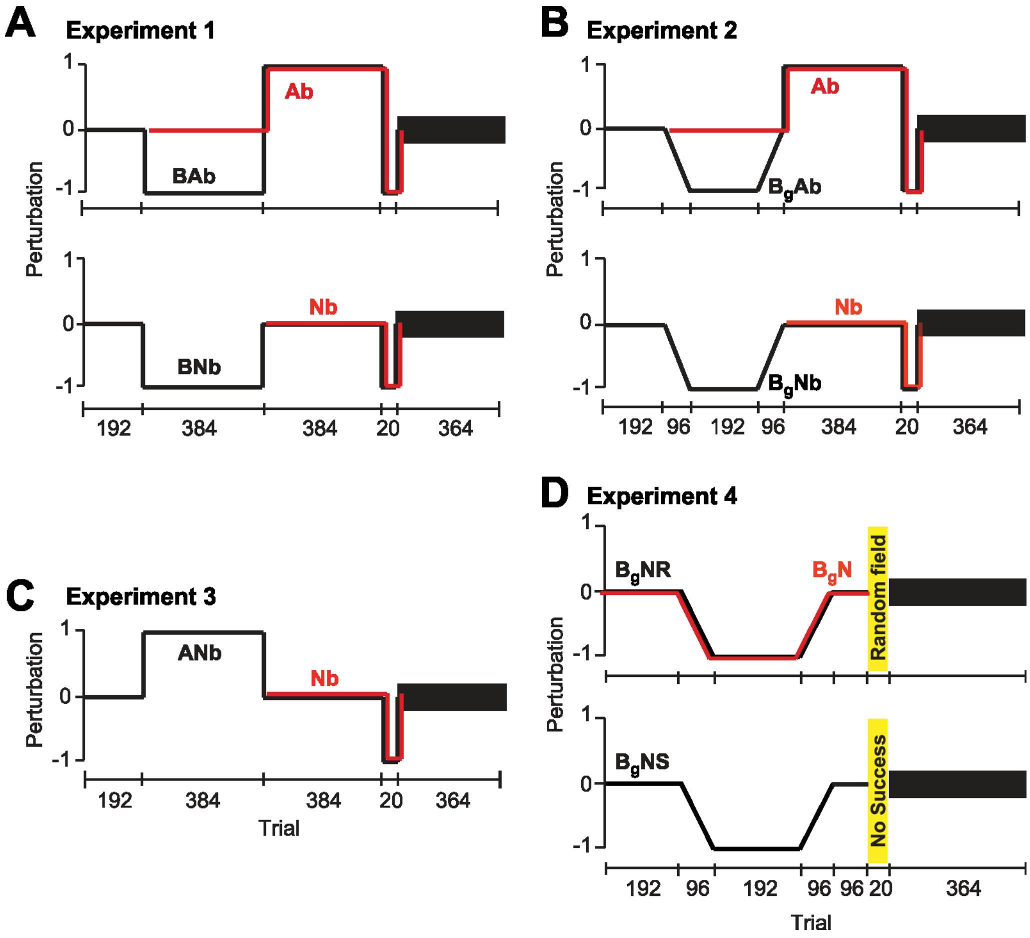


Figure 2.1 Experimental design: subjects held the handle of a light weight robotic arm and reached to a target. For the first 192 trials, the robot produced null (zero) forces. Field *A* is represented by positive values of perturbation, and field *B* by negative values. Abrupt changes in the field are noted by discontinuities in the perturbation line. Gradual changes are noted by negative or positive sloped perturbation lines. The block of error-clamp trials is depicted by the shaded region. In an error-clamp trial, there is no perturbation. Rather, a stiff force channel directs the hand to the target, allowing us to measure forces that subjects produce on that trial, but minimizing learning from movement errors. **A)** Experiment 1: step changes in the perturbations. BNb, Nb, BAb, and Ab groups. **B)** Experiment 2: gradual changes in the perturbations. B_gNb, Nb, B_gAb, and Ab groups. **C)** Experiment 3: ANb and Nb groups. **D)** Experiment 4: manipulation of probability of error and probability of success. In B_gNR group, after washout there are 20 trials in which a random field is present, after which subjects are exposed to an error-clamp block. In B_gNS group, after washout there are 20 error-clamp trials in which there is zero probability of acquiring a reward. After the 20 no-success error-clamp trials, the error-clamp block continues with the usual constraints for reward.

2.2.2 Data analysis

Performance was measured via the force that subjects produced against the channel wall of the error-clamp trials. The force output as a percentage of perturbation was calculated via the ratio of the actual force output as measured at maximum velocity in error-clamp trials, to the ideal force required to cancel the perturbation at that velocity. During the 20 trials of *B*, no error-clamp trials were given. In order to measure performance in these trials, the perpendicular displacement from a straight line to the target at maximum velocity was calculated for each movement. This served as the proxy for movement error in non-error-clamp trials. Repeated-measures ANOVA and post hoc Tukey's test were used to confirm that all groups reached equivalent levels of adaptation during force field trials. Two-tailed t-tests were used to quantify the differences in initial bias of the error-clamp block, and to evaluate the bias in average motor output between groups for the last 100 trials of the error-clamp block. All analyses were done using Matlab 7.0.4 and SPSS.

We performed a boot-strapping procedure to estimate the strength of memories of *A* and *B* that were expressed during error-clamp trials (Criscimagna-Hemminger and Shadmehr, 2008). For example, in order to estimate how null affected the previously acquired memory of *B*, we randomly selected one subject from the BNb group (with replacement) and another from the Nb group (with replacement) and found the difference in force output (percent perturbation) for each trial during the error-clamp block. In other words, the *B* that remained and was expressed after the null field training is $\hat{B} = BNb - Nb$. This is the assay of the memory that was protected from destruction. We repeated this subtraction 100 times, randomly selecting subjects from each group and established a distribution for $\hat{B}$. Similarly, the memory for B_g that remained after subjects experienced a null field was $\hat{B}_g = B_g Nb - Nb$. To account for any differences in the 20 trials of *B* following *A* training, the *Ab* group served as the control to assay the *B* that remained and was expressed after the *A* field training.

2.2.3 Model

We compared the predictions of two previously published models of motor adaptation: a model that allowed erasure of memories (Smith et al., 2006), and a model that used sudden errors to protect memories (Lee and Schweighofer, 2009). Both models are multi-rate, multiple timescale formulations that allow us to compute the expected patterns of spontaneous recovery in error-clamp trials. Both use the following error equation to drive motor adaptation. For force-field trials

$$e^{(n)} = f^{(n)} - x^{(n)} \qquad (1)$$

In the above equation, $e^{(n)}$ is the prediction error on trial *n*, $f^{(n)}$ is the perturbation force, and $x^{(n)}$ is motor output. In error-clamp trials, the perturbation is equal and opposite to the force produced by the subject. Adaptation in the Smith et al. (2006) model is achieved by two internal states, one fast process that adapts quickly but has poor retention, and one slow process that learns slowly but has better retention. The update equations for the net motor output are given by:

$$\begin{aligned} x^{(n)} &= x_f^{(n)} + x_s^{(n)} \\ x_f^{(n+1)} &= a_f x_f^{(n)} + b_f e^{(n)} \\ x_s^{(n+1)} &= a_s x_s^{(n)} + b_s e^{(n)} \end{aligned} \qquad (2)$$

The learning rates for the fast and slow states are $1 > b_f > b_s$ and the forgetting rates for these states are $1 > a_s > a_f$.

In the above model (Eq. 2), both the fast and the slow process are updated by the same prediction error. In this model, memories are not protected. In contrast, in the Lee and Schweighofer (2009) model there is one fast state and many slow states. The slow states are selected based on contextual cues:

$$
\begin{aligned}
x^{(n)} &= x_f^{(n)} + \mathbf{x}_s^{(n)T}\mathbf{c}^{(n)} \\
x_f^{(n+1)} &= a_f x_f^{(n)} + b_f e^{(n)} \\
x_s^{(n+1)} &= a_s x_s^{(n)} + b_s e^{(n)}\mathbf{c}^{(n)}
\end{aligned}
\qquad (3)
$$

The contextual cue is switched based on large errors. This allows for protection of slow memories. In error-clamp trials there is no contextual cue to allow for explicit memory selection, so the value of each element of the context vector is set to $1/m$, where m is the number of contexts. The parameters for both models were set by those given in Joiner and Smith (2008), as a_f=0.85, a_s=0.998, b_f=0.11, and $b_s = 0.021$.

2.3 Results

Our first aim was to determine whether adaptation followed by an equal period of washout or reverse-adaptation resulted in catastrophic destruction of a motor memory. Experiment 1 (Fig. 1A) was designed to answer this question. Let us begin with some simulations to illustrate how the patterns of spontaneous recovery, i.e., motor output in the error-clamp block, should be affected if memories are protected from unlearning.

Consider a training protocol in which a long period of adaptation in *A* (~400 trials) is followed by a brief period (20 trials) of adaptation in *B*, where $B = -A$. This training is then followed by error-clamp trials in which movement errors are clamped to zero (Ab paradigm, left column of Fig. 2A). To simulate learning, we considered two existing models. The first model (Fig. 2B) assumed a single context (Smith et al., 2006) in which errors always produced learning/unlearning. The second model assumed multiple contexts (Lee and Schweighofer, 2009) in which sudden errors produced a contextual change that protected a component of the currently activated memory from unlearning (Fig. 2C). Both models assumed that changes in motor output are due to a fast adaptive process that learned strongly from error but had poor retention, and a slow adaptive process that learned weakly from error but had strong retention. The multiple context model further assumed that the slow component of the memory (but not the fast component) was contextual: a sudden change in error signaled a change in context, resulting in de-activation of the slow trace and instantiation of a new slow trace.

As the left column of Figs. 2B and 2C illustrate, both models predict spontaneous recovery in the Ab paradigm. At the end of Ab training, in the single context model there is a slow memory of *A* and a fast memory of *B*. In the error-clamp block the different rates of decay of the fast and slow processes produce spontaneous recovery of *A*. In contrast, in the multiple context model the sudden errors that occur in the transition from *A* to *B* signal a contextual change. This contextual change de-activates slow *A* (i.e., it no longer contributes to output), and protects it from unlearning, while activating a slow *B*. The multiple context model further assumes that the transition from *B* to error-clamp trials causes re-activation of *A* so that both the slow *B* and slow *A* are present in the error-clamp trials. The important idea is that whereas both models account for the rise of motor output from baseline toward *A* in the error-clamp trials, they do so with very different interpretations: the single context model explains that this rise is due to passive decay of currently activated memories, whereas the multiple context model explains this pattern as a consequence of re-activation of previously inactive and protected memory of *A*.

In previously published work, the Ab paradigm indeed produced rise of motor output from baseline toward *A* followed by a gradual decline (Smith et al., 2006;Ethier et al., 2008;Criscimagna-Hemminger and Shadmehr, 2008). As the above simulations show, both a single context model that allows for erasure, and a multiple context model that protects memories can account for this pattern. However, a simple experiment can dissociate between these two models. Consider a training protocol in which a long period of training in *B* precedes the Ab training (BAb paradigm, right column Fig. 2A). In this scenario, the single context model predicts that because the length of training in *B* is equal to the subsequent training in *A*, the fast and slow memories that are produced by *B* are transformed to fast and slow memories for *A*, i.e., *A* destroys *B* (Zarahn et al., 2008). As a result, in the single context model the motor output in the error-clamp block is identical in the BAb and Ab paradigms (Fig. 2D). In contrast, in the multiple context model the slow *B* is protected from error-dependent learning during training in *A,* but is then expressed in the error-clamp block. Consequently, during the error-clamp block the motor output in BAb is biased toward *B* as compared to Ab (Fig. 2D). In summary, if memories are protected, then we should see that the motor output in the error-clamp block in BAb is biased toward *B* as compared to Ab. We focus on the bias that persists throughout the error-clamp block as evidence of multiple memories, as this assay should indicate retention of the slow memories without contamination from any fast learning or switching that occurs at the transitions.

The organization of the experiments is as follows: In Exp. 1, we will show that in the BAb and similar paradigms, motor output in the error-clamp trials is biased toward *B*, suggesting protection of *B* during adaptation to *A*. In Exp. 2 we will show that protection of *B* is unrelated to sudden errors that might signal a contextual change, raising doubt about models in which contextual change is based on kinematics errors. In Exp. 3 and 4 we will show that spontaneous recovery, i.e., expression of a previously acquired memory in error-clamp trials, is an active process of recall and not passive decay of an already active motor memory. Finally, we will show that this active recall is associated with withholding of reinforcement for a current motor output, resulting in the retrieval of a previously acquired memory.

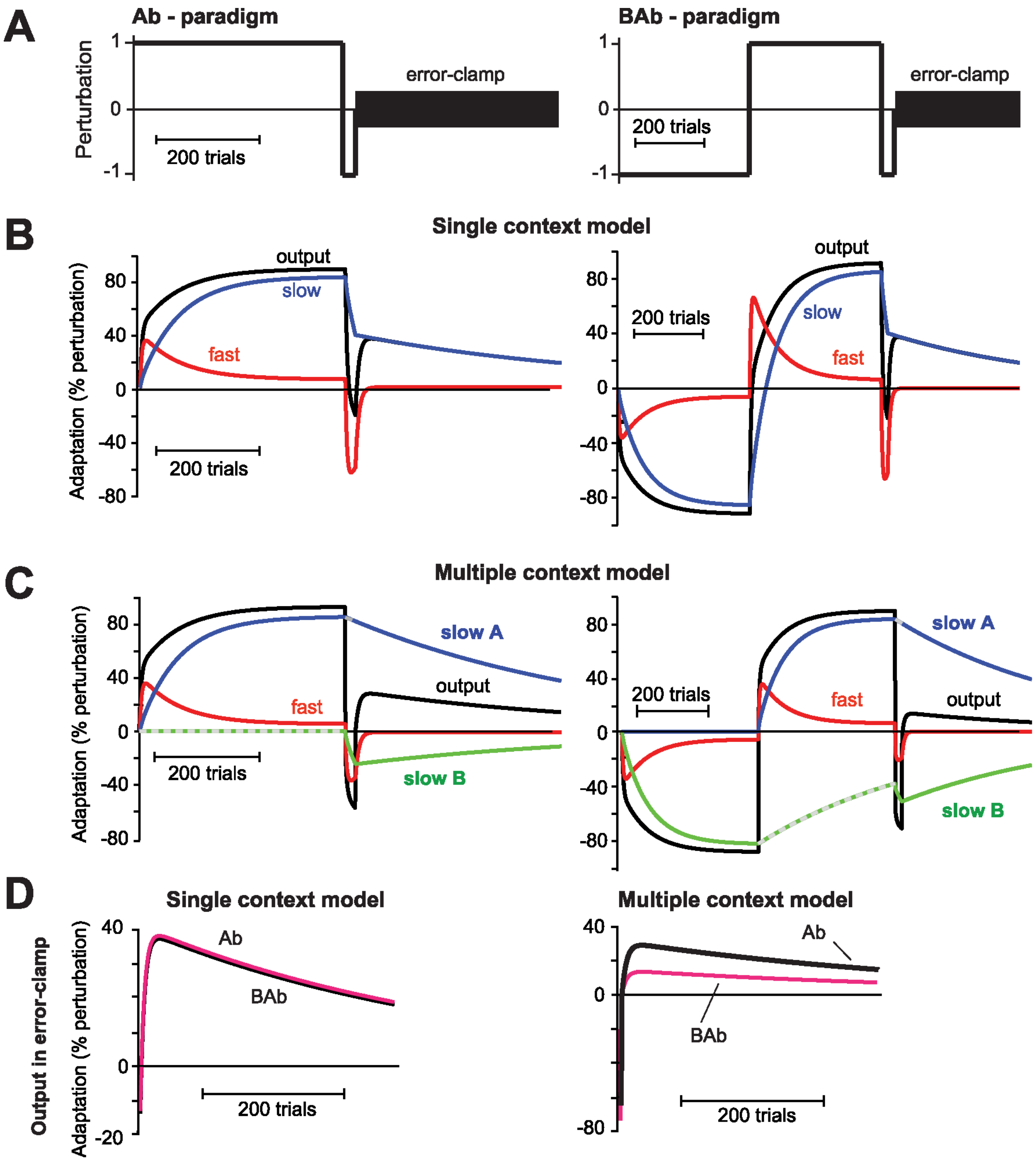


Figure 2.2 Simulation results. **A)** Perturbation schedule for Ab and BAb. **B)** Simulation results for single context model. In this model, there is one slow and one fast state. **C)** Simulation results for the multiple context model. In this model, there is one fast state and two slow states, one for each possible 'context'. **D)** Force output in the error-clamp block for the single context and multiple context models. In the single context model, force output is identical in the BAb and Ab paradigms. This is not the case for the multiple context model. In the Ab paradigm the multiple context model (like the single context model) exhibits spontaneous recovery of *A*. However, in the BAb paradigm, the multiple context model (unlike the single context model) exhibits a motor output that is biased toward *B* as compared to Ab.

2.3.1 Experiment 1: memories are protected from unlearning

The design of this experiment is shown in Fig. 1A. To determine whether the memory of *B* is protected during subsequent training in *A*, we compared the forces that subjects produced in the error-clamp block in the BAb and Ab groups. We noted that by end of training in *A*, performance of the Ab and BAb groups were indistinguishable (comparison of last five error-clamp trials in *A*, $F(1,16) = 0.041$, $p > 0.5$). However, in the error-clamp block the motor output in the BAb group was biased toward *B* as compared to Ab (Fig. 3A). For example, the forces on the first error-clamp trial were significantly more biased toward *B* in BAb than Ab (t-test, $p = 0.005$). Furthermore, the forces reached a much lower plateau in the BAb group as compared to the Ab group (average of last 100 trials, t-test, $p < 0.001$). In comparing the BAb and Ab groups, the motor output in the error-clamp block is similar to the predictions of the multiple context model (right column, Fig. 2C), suggesting that memory of *B* was protected during training in *A*, and then re-expressed in the error-clamp block.

To determine whether the memory of *B* could be destroyed by subsequent training in null, we compared performance in the error-clamp block in the BNb group vs. the Nb group. We noted that by the end of the null trials performance of the Nb and BNb groups were indistinguishable (comparison of the last five error-clamp trials in null, $F(1,16) = 0.121$, $p > 0.5$). However, in the error-clamp block the motor output in the BNb group was biased toward *B* as compared to Nb (Fig. 3B). For example, the forces on the first error-clamp trial were significantly more negative in BNb than Nb (t-test, $p = 0.005$).

This bias, however, vanished beyond the 50th trial of the error-clamp block. Therefore, training in null appeared to have a greater effect on memory of *B* than training in *A*. [It is interesting that in the Nb group, 20 trials of *B* are sufficient to produce a memory that does not decay to zero even after 300 error-clamp trials. This is a consistent finding that we have found regardless of whether the 20 trials are in a clockwise or a counterclockwise field (Keisler and Shadmehr, 2010).] Taken together, the patterns of motor output during error-clamp block of Exp. 1 suggested that during adaptation to *A* or washout in null, the previously acquired memory of *B* was at least partially protected. Brief re-exposure to *B* produced re-expression of the previously acquired *B* memory.

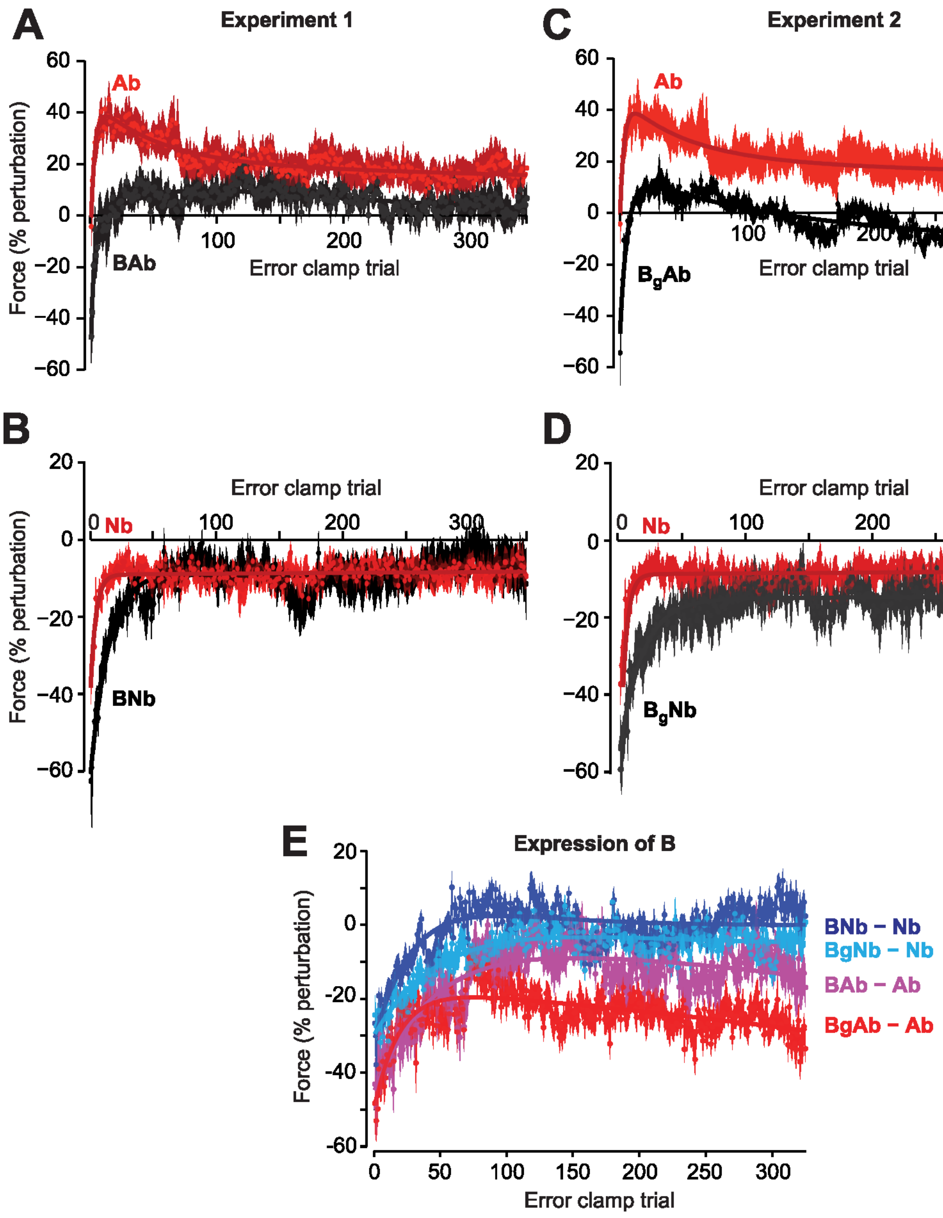


Figure 2.3 Motor output in the error-clamp block. Dots are across subject mean of individual trials. Shading is SEM. **A)** Ab and BAb comparison. Prior training in B biases the motor output toward B. Compare this plot with the predictions of the multiple-context model in Fig. 2C. **B)** Nb and BNb comparison. **C)** Ab and BgAb comparison. **D)** Nb and BgNb comparison. **E)** Expression of B in the error-clamp block. Each data set represents estimate of B as computed through a between group subtraction. For example, the blue data set represents the difference in the motor output between the BNb and Nb groups in the error-clamp block: $\hat{B} = BNb - Nb$.

2.3.2 Experiment 2: protection despite paucity of sudden errors to signal a contextual change

One possibility is that in Exp. 1, the large movement errors that accompanied introduction of *B*, or the large errors that accompanied transition to null, acted as context cues that facilitated protection of *B*. Indeed, such kinematic errors are the basis for contextual change in theoretical models (Jacobs et al., 1991;Haruno et al., 2001;Doya et al., 2002). Without the large errors to mark a change in context, the theoretical models predict that memories will be erased. Furthermore, expression of the *B* memory in the error-clamp trials may be due to the fact that the errors induced by the 20 *B* trials were similar to the initial errors experienced during adaptation to *B*. Thus, we performed Experiment 2 (Fig. 1B) in order to ask two questions: whether protection of the *B* memory required sudden change in errors to signal a change in context, and whether recall of this memory relied on cues that were error dependent.

In the B_gAb group, field *B* was introduced gradually, and then after a period of constant perturbation, was gradually returned to null, following which *A* was introduced abruptly (Fig. 1B). This was followed by 20 *B* trials and then a long sequence of error-clamp trials. We imagined that if formation of the *B* memory required a sudden perturbation to ‘label’ it, or if its recall during re-exposure (20 *B* trials) required a similarity between errors during initial learning and re-exposure, then the forces in the error-clamp block would be similar in B_gAb and Ab. Instead, we found a strong bias toward *B* in the B_gAb group as compared to Ab. For example, the forces on the first error-clamp trial were more negative in B_gAb vs. Ab (t-test, $p = 0.004$). Furthermore, the forces reached a much lower plateau in the B_gAb group as compared to the Ab group (average of last 100 trials, t-test, $p < 0.001$). These results suggested two ideas: 1) that a sudden change in error was not required for establishing a motor memory that could be protected, and 2) that recall of a motor memory did not require errors during re-exposure that were similar in magnitude to errors experienced during acquisition. This last point is crucial as it suggests that expression of *B* in the error-clamp block is not based on a comparison between errors acquired during acquisition and re-testing. Finally, because the forces had a lower plateau in the B_gAb group vs. the BAb group, it would appear that a memory that is acquired without sudden errors (gradual *B*) is more resistant to destruction than a memory that is acquired with sudden errors (abrupt *B*) (Huang and Shadmehr, 2009).

The transition from *B* to *A* in the above experiment was abrupt, inducing sudden errors. Is this abrupt transition crucial for protection of the *B* memory? To check for this crucial assumption of multiple context models, we considered performance of the B_gNb group in comparison to Nb group (Fig. 1B). In the B_gNb group, the initial *B* memory was acquired without sudden changes in error and its transition to null was also without sudden changes. After the brief re-exposure to *B*, we again found a strong bias toward *B* in the B_gNb group vs. the Nb group. For example, the forces on the first error-clamp trial were more negative in B_gNb vs. Nb (t-test, $p = 0.030$).

In summary, results of Exp. 2 suggested three ideas. First, the protection of the *B* memory did not rely on a sudden change in errors that may have signaled a change in context. Second, the recall of the *B* memory did not rely on cues such as error size that might be shared in initial exposure and re-exposure. This implies that sudden movement errors were not necessary to contextually label a memory so that it could be protected or later recalled. Finally, gradual adaptation to *B* produced a motor memory that was more resistant to subsequent training in *A* or null (as compared to abrupt adaptation to *B*), as evidenced by a stronger bias toward *B* in Exp. 2 vs. Exp. 1.

Expression of memories in error-clamp trials

While the results from Experiments 1 and 2 suggest that the memory of *B* was not destroyed by subsequent training in null or *A*, it is useful to quantify how much of this memory was expressed in the error-clamp trials. We estimated expression of the *B* memory in the error-clamp block using a bootstrapping method. For example, to quantify *B* that was expressed in the BNb group, we subtracted the forces produced by the Nb group from the BNb group, i.e., $\hat{B} = BNb - Nb$ (Fig. 3E). This analysis was critical in order to determine the contributions of the slow memory of B retained and re-expressed, over the bias induced by the 20 trials of B exhibited in the Nb condition. The results suggested that while in all experiments a significant amount of *B* memory was expressed in the error-clamp trials, there was a trend towards stronger expression in the BAb and B_gAb groups (calculated by $BAb - Ab$) than BNb and B_gNb groups. That is, somewhat surprisingly the expression of the *B* memory was more affected by the null washout trials than by adaptation to the opposite perturbation.

2.3.3 Sudden change in probability of success as a possible contextual cue

Fig. 4A plots the movement errors in the groups that were abruptly introduced to *B* and then transitioned abruptly to null. The largest trial-to-trial change in error occurred when *B* was introduced (~22mm), and when null was re-introduced (~25mm). In comparison, gradual introduction of *B* and gradual re-introduction of null produced trial-to-trial changes that were no larger than 3mm and 6mm, respectively (these errors occurred following set breaks). If a sudden change in movement error signals a contextual change, in the gradual condition these cues were less available. Yet, protection and expression of the *B* memory was more robust in the gradual condition than the abrupt condition (BgNb vs. BNb, Fig. 3). Therefore, it seems unlikely that sudden changes in movement errors act as contextual cues.

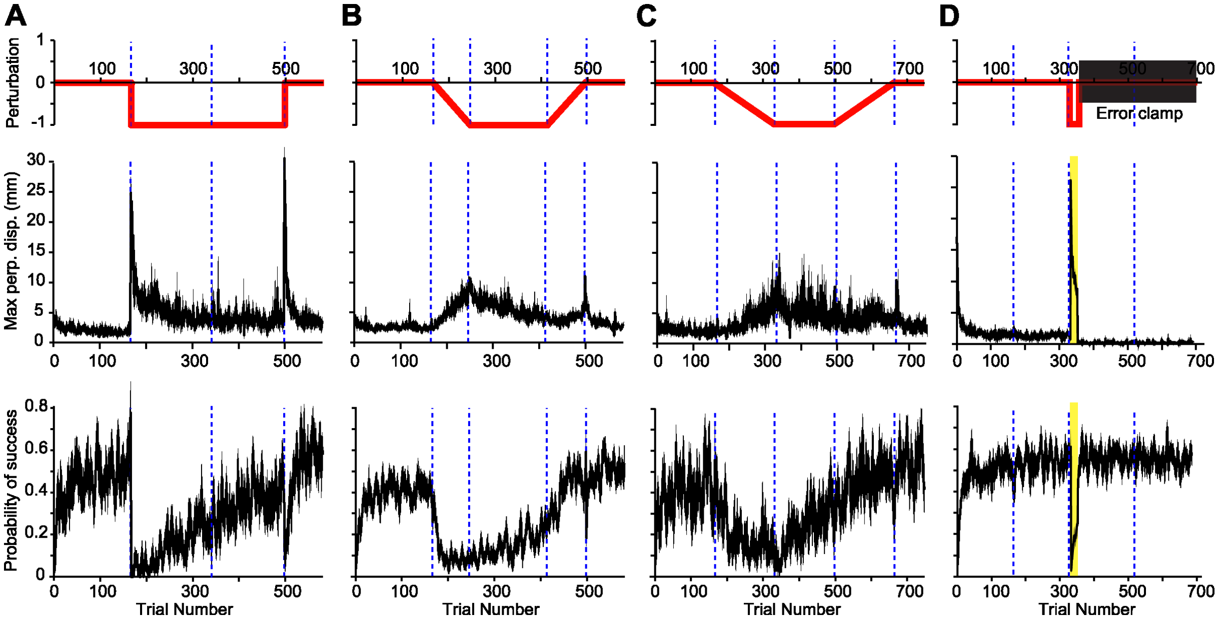


Figure 2.4 Error magnitudes and reward probabilities during adaptation to gradual and abrupt perturbations. The graphs represent across subject mean and SEM. Dashed lines indicate set breaks. **A)** Abrupt perturbation. Data is from the BNb group. **B)** Gradual perturbation. Data is from $\mathrm{B_gNb}$. **C)** Very gradual perturbation. Data are from the $\mathrm{B_{gg}Nb}$ group. **D)** The effect of sudden change in perturbation immediately before the error-clamp block. Data are from the Nb, BNb, and ANb groups.

Another source of information that can signal a contextual change is probability of success. During initial null field training, all abrupt and gradual BNb conditions displayed comparable levels of success rate ($F_{(2,42)} = 1.546$, $p = 0.225$). In the abrupt condition, the probability of a successful trial dropped sharply at the onset of *B* (Fig. 4A). Interestingly, the probability of success also dropped significantly in the gradual condition (Fig. 4B). For example, when field *B* was at 25% of full strength the probability of success had dropped by more than 80%. Therefore, while in the gradual condition there were no sudden changes in performance errors, the gradual accumulation of the small errors and the non-stationary nature of the environment led to substantial reductions in reinforcement.

Perhaps this decrease in probability of success in the gradual condition was because we increased the strength of the field too quickly (over ~100 trials), leading to larger trial-to-trial variance and displacement error. That is, perhaps the gradual *B* in Exp. 2 was not gradual enough. To check for this, we recruited a new group of subjects for a paradigm in which the perturbation was introduced very gradually, over twice as many trials as before (the $B_{gg}Nb$ group, as illustrated in Fig. 4C, top row, in which the field reached full strength after 192 trials). Once again we observed a large drop in the probability of success, despite the fact that the perturbation only produced a minimal increase in movement errors. For example, when the perturbation had reached 25% full strength, movement errors had increased by about 3mm from baseline, but probability of success had dropped by 75%. After washout in null, this new ultra-gradual group also exhibited strong expression of *B* in the error-clamp block: forces were strongly biased toward *B* in the B_{gg}Nb group as compared to Nb (t-test, $p = 0.020$). In summary, in the gradual condition we observed small incremental increases in movement errors, but much sharper declines in probability of success. It is possible that a large change in probability of success acted as a cue that signaled a contextual change, initiating a search for better motor commands (Izawa and Shadmehr, 2011a).

Let us now consider the events that took place immediately before the error-clamp block. Fig. 4D displays the movement errors and probability of success in the various groups that experienced null, then brief exposure to *B*, and then the error-clamp block. In the 20 *B* trials movement errors suddenly increased and were then eliminated by the transition to the error-clamp block. Similarly, probability of success suddenly decreased, and then recovered. Therefore, one of the critical events that took place in the 20 *B* trials was that previously reinforced motor commands (in null or *A*) were no longer reinforced. It is possible that this withholding of reinforcement for a current motor memory resulted in the expression of the competing motor memory. We will test this idea directly in Experiments 3 and 4.

2.3.4 Experiment 3: spontaneous recovery of a motor memory following a sudden change in performance errors

Why is it that the *B* memory is being expressed in the error clamp trials? Is it because the 20 *B* trials are in the same field as the *B* that was experienced before? Or is it that the sudden introduction of movement errors and change in probability of success that takes place in the 20 *B* trials encourage a switching from expression of one memory to another? To decide between these possibilities, we consider a scenario in which the errors that came before the error-clamp block were unrelated to the errors that were experienced during acquisition of the memory. In the ANb group (Fig. 1C), training in *A* was followed by a long period of training in null, and then 20 trials in *B*. During these 20 trials the brain will experience a sudden decline in performance. That is, previously reinforced motor commands (appropriate for null) will no longer be reinforced. Will the 20 trials in *B* produce spontaneous recovery of *A*?

By the end of the null trials, performances in the ANb and Nb groups were indistinguishable (average of last five error-clamp trials, $F(1,16) = 0.108$, $p > 0.5$). Furthermore, as Fig. 5A illustrates, performance of these two groups were indistinguishable during the 20 *B* trials (perpendicular displacement, $F(1,16) = 0.80$, $p > 0.5$). Therefore, during the training in *B* there was no evidence of prior training in *A* in the ANb group. Finally, the forces in the first trial of the error-clamp block were indistinguishable between ANb and Nb (t-test, $p = 0.398$). Remarkably, as the trials in the error-clamp block continued, the ANb group produced forces that became biased toward *A* (average of last 100 trials, t-test, $p<0.001$). We performed a bootstrap analysis to quantify expression of *A* in the error clamp trials: $\hat{A} = ANb - Nb$. The results (Fig. 5C) demonstrated that the *A* memory exhibited spontaneous recovery during the error-clamp block, despite the fact that this block occurred hundreds of trials after acquisition of *A*, and was preceded with training in *B*. Along with Experiment 2, the data in Fig. 5C suggest that the mere presence of sudden errors and/or sudden changes in probability of success produce spontaneous expression of a previously acquired and presumably de-activated motor memory.

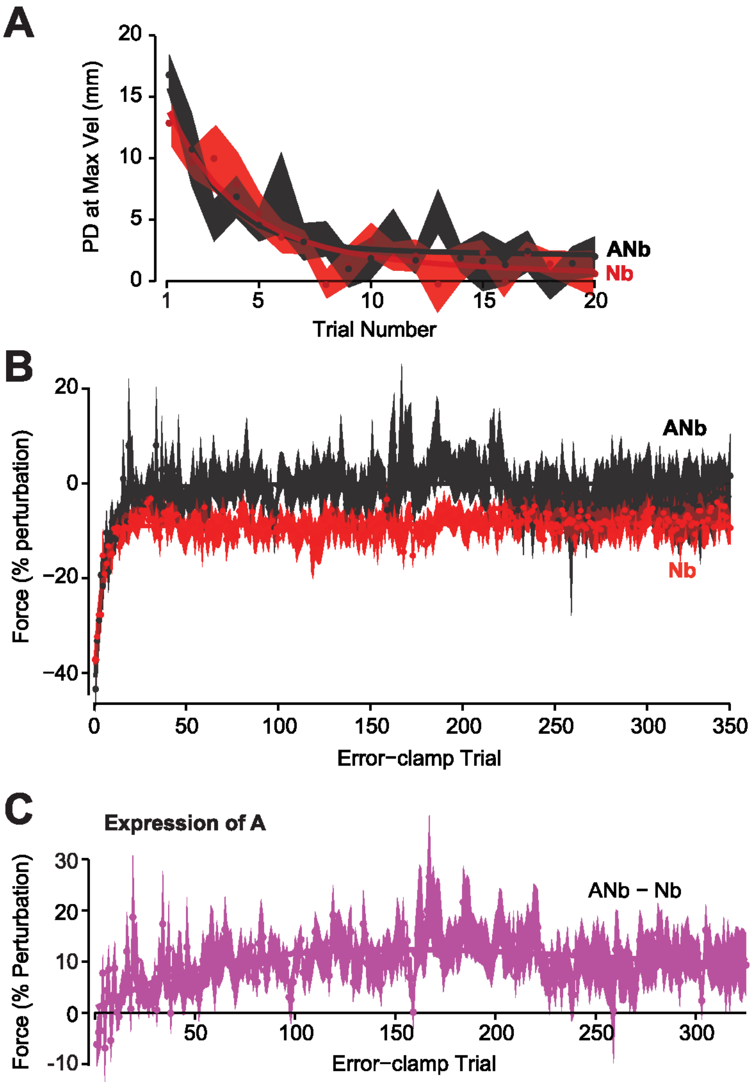


Figure 2.5 Spontaneous recovery in Anb. The data in this figure depict the results of Exp. 3 in which we compared the ANb and Nb groups. **A)** Movement errors (perpendicular displacement at maximum velocity) during the 20 *B* trials that precede the error-clamp block. The two groups appear indistinguishable. Error bars are SEM. **B)** Force in the error-clamp block. Error bars are SEM. **C)** Expression of A in the error-clamp block, as estimated through bootstrap analysis $\hat{A} = ANb - Nb$. Error bars are 95% confidence intervals.

2.3.5 Experiment 4: spontaneous recovery of a motor memory following withholding of reward for the competing memory

In the B_gNR group (Fig. 1D, top plot), the null training was followed by 20 trials in which the field was random on any given trial (*A*, *B*, or null). As a control, we considered the B_gN group in which *B* was introduced gradually and removed gradually, followed by washout in null and then error-clamp trials. Note that in the B_gN group the null training directly leads to the error-clamp trials. Therefore, the B_gN group experiences no sudden change in error and/or probability of success before the error-clamp block. In contrast, in the BgNR group 20 high error and low success trials immediately precede the error-clamp block. Indeed, in the B_gNR group we observed robust expression of the *B* memory in the error-clamp block (Fig. 6A) (average of last 100 trials, t-test, $p < 0.001$). Note that *B* was learned gradually and without large errors, yet it was re-expressed after washout when the subjects encountered a sequence of random large errors. In comparison, the forces produced by the B_gN group in error-clamp trials were indistinguishable from zero (first 250 trials, $F(1,16) = 0.617$, $p > 0.4$). Therefore, the error-clamp block by itself was not sufficient to produce expression of a previously acquired memory. Rather, a small number of trials in which there were large errors and low probability of success produced a condition in which a previously acquired memory showed spontaneous recovery.

In BgNR group, the random condition consisted of a number of trials in which field *B* was present. It is possible that expression of *B* was due to occasional presence of this perturbation immediately before the error-clamp trials. The alternate hypothesis is that the brain expressed *B* because the motor memory for null was no longer producing a rewarding outcome in the random field. Our final experiment was designed to test the idea that the brain switched between motor memories merely because of sudden changes in probability of success.

In the B_gNS group the null trials were followed by 20 no-success error-clamp trials in which regardless of the movement, target explosions were withheld (schematic in Fig. 1D, success probabilities in Fig. 6B). These no-success error-clamp trials were followed by the usual error-clamp block. Before the no-success trials motor output was comparable between the B_gNS and B_gN groups (Fig. 6A, open circles). However, as reinforcement was withheld, subjects in the B_gNS group began producing forces appropriate for *B*. Indeed, trial after trial the withholding of reinforcement encouraged greater expression of *B*. By the 23rd trial, expression of *B* was similar in the B_gNS and B_gNR groups. As the error-clamp trials continued, the motor output in the B_gNS group continued to be biased toward *B* as compared to the BgN group (all 364 error-clamp trials, t-test, $p < 0.001$). Therefore, withholding of reinforcement during expression of the null field memory produced spontaneous recovery of the memory for *B*.

In summary, Experiments 3 and 4 demonstrated that sudden removal of reinforcement encouraged expression of a previously acquired motor memory. This suggests that in the

BAb, BNb, and similar experiments in which memory of *B* was spontaneously expressed in error-clamp trials, a critical factor was the fact that the current motor commands (*A* or null) suddenly became unsuccessful in acquiring reinforcement. This sudden change in probability of success encouraged expression of a previous acquired memory, i.e., the memory of *B*.

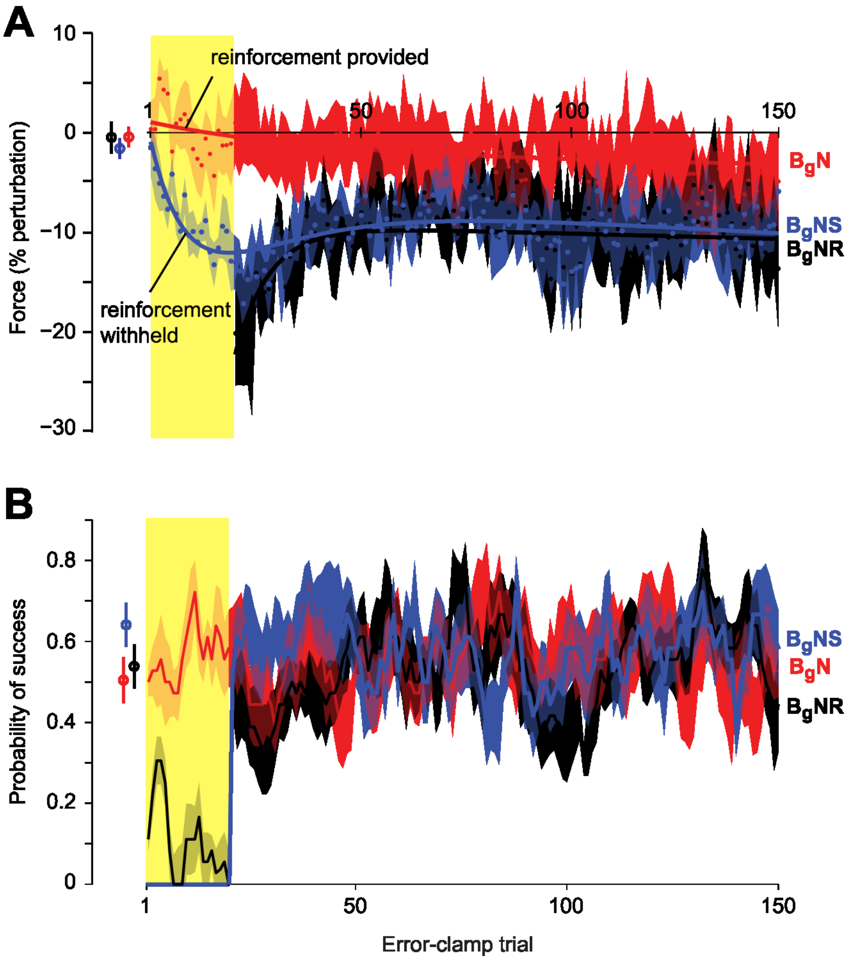


Figure 2.6 Withholding reinforcement produces spontaneous recovery of *B* in the error-clamp block. The data depicts the results of Exp. 4 in which we compared the performance of the BgN, BgNR, and BgNS groups. **A)** Open circles indicate the forces that subjects produced in the null field (representing the last 3 error-clamp trials in the null field), i.e., after completion of training in *B* and washout. These forces are similar between groups. Upon entering the error-clamp block, reinforcement is withheld for the first 20 trials for the B_gNS group (yellow region), but is present for the B_gN group. For the B_gR group, the first 20 trials are in a random field (yellow region, in which there were no error-clamp trials). Beyond the first 20 trials, all groups are reaching in an error-clamp. **B)** Probability of success for each group. Error bars are SEM.

2.4 Discussion

In numerous experiments people have adapted their movements to perturbation *B*, and then adapted to the opposite perturbation *A*. To determine whether adaptation to *A* destroyed the memory of *B*, they were re-tested in *B*. When the temporal distance between the training episodes was zero, as in experiments here, performance in retest was usually no different than naïve (Lewis et al., 1952;Flook and McGonigle, 1977;Shadmehr and Brashers-Krug, 1997;Bock et al., 2001;Caithness et al., 2004;Krakauer et al., 2005;Overduin et al., 2006), suggesting catastrophic interference. Here, we found that *BA* training produced two competing memories. When motor output associated with one memory was denied reinforcement, the competing memory was retrieved.

In Exp. 1 subjects trained in *B* and then in null or *A*, followed by brief exposure to *B*. We found that in BAb and BNb, motor output during the error-clamp block was biased toward *B* as compared to Ab and Nb, respectively. Therefore, adaptation followed by de-adaptation did not result in catastrophic destruction. In Exp. 2 we asked whether protection of a memory required sudden errors to mark a context-change. We found that despite gradual presentation of the perturbation, forces in the error-clamp block were biased toward *B*. Therefore, protection was not based on contextual change signaled by large kinematic errors. In Exp. 3 we trained subjects in *A* and then after washout, presented 20 *B* trials. The sudden errors in *B* produced spontaneous recovery of *A*. Therefore, spontaneous recovery was an active process of retrieval and not passive decay of an already active memory. Finally, in Exp. 4 we found that random errors could produce spontaneous recovery. Most interestingly, we found that when we trained in *B* and then denied reinforcement following washout, the brain retrieved the memory of *B*. Therefore, when current motor commands produced the expected kinematic outcome but were un-rewarded, the brain expressed another set of motor commands that were previously rewarded.

2.4.1 The multiple components of motor memory

When one learns to produce motor commands that compensate for a perturbation, and then that perturbation is removed, what prevents erasure of the motor memory? There are a number of computational models of learning in which the system is composed of multiple modules, each an expert with a forward model that predicts behavior in a particular context of the environment, paired with an inverse model or controller that produces motor commands (Wolpert and Kawato, 1998;Haruno et al., 2001;Doya et al., 2002). The forward and inverse models are tightly coupled during acquisition and use. Importantly, switching between modules takes place due to a responsibility selector that assigns credit to each module based on the accuracy of predictions made by its forward model. Such models produce protection of acquired memories when there are sudden large errors in behavior. Our results appear inconsistent with these models: first, we found that motor memories were protected even when perturbations were introduced gradually, preventing large errors. Second, we found that the brain switched from expressing one memory to another merely because current motor commands were not acquiring reward, despite paucity of kinematic errors.

A possible approach is to change the Lee and Schweighofer (2009) model so the switch between slow states is based on probability of success, rather than large performance error. In the resulting model, slow states learn from performance errors, but contextually switch based on probability of success. However, as this learning depends entirely on performance errors, we cannot account for the fact that during adaptation, repeated reinforcement of a movement produces a memory independent of the memory produced via error-dependent learning (Orban de Xivry et al., 2011;Huang et al., 2011a).

We approach our problem by considering a different model of motor memory. Suppose the process of generating a movement involves two computations: one that transforms a target state x_t into motor commands u, i.e., a control policy, possibly in the motor cortex, and one that transforms motor commands into predicted sensory consequences $\hat{x}$, i.e., a forward model, possibly in the cerebellum (Shadmehr and Krakauer, 2008). Upon exposure to an abrupt perturbation *B*, sensory prediction errors produce adaptation of the forward model, a process that depends on the cerebellum (Synofzik et al., 2008). At this early stage of learning, motor commands improve not because the controller has changed, but because commands are corrected via internal feedback through the forward model (Chen-Harris et al., 2008). This accounts for the fact that early in training, despite large improvements in performance there is little or no change in the motor cortex (Paz et al., 2003) and disruption of the motor cortex does not affect the initial rapid phase of adaptation (Orban de Xivry et al., 2011). As training continues, certain motor commands repeat and are reinforced by success. This reinforced repetition produces a distinct motor memory (Diedrichsen et al., 2010;Huang et al., 2011a;Verstynen and Sabes, 2011) that depends on the motor cortex (Orban de Xivry et al., 2011), producing plasticity in the controller so that motor outputs are associated with reward (Fig. 7B). The motor command in any given trial is the one that is most likely to be rewarded (the mode of this distribution). At the end of *B* training, we have acquired a new forward model and controller (Figs. 7B&C).

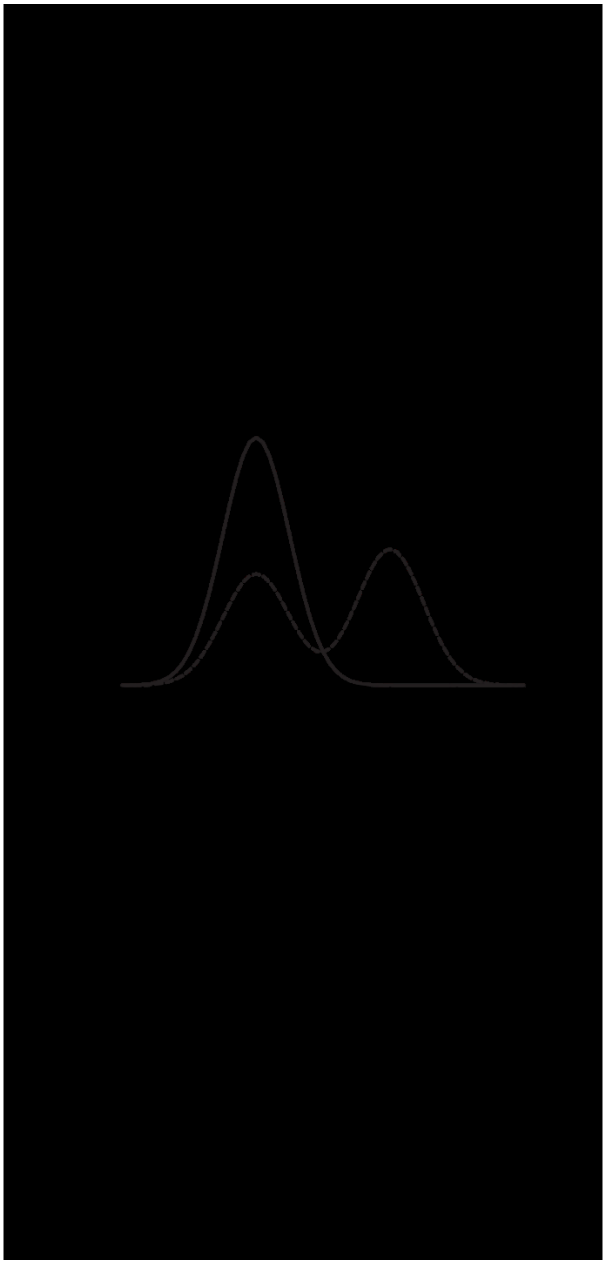


Figure 2.7 A conceptual model of motor memory during BA paradigms. A) Given a target state x_t and estimate of current state $\hat{x}$, the control policy produces motor command u. Selection of control policy depends on history of reward, which is learned by the policy selector, possibly in the basal ganglia. Motor commands produce predictions of sensory consequences, a function of the forward model. Sensory prediction errors produce adaptation of the forward model, whereas repetition and reward produce a change in the controller. B) The controller. At the end of training in perturbation B, the motor commands are associated with probability of reward as shown by the solid line. When the perturbation switches to A, previously rewarded motor commands are no longer rewarded, resulting in a search for new motor commands that can produce reward. At the end of BA training the probability of reward is bimodal, reflecting a peak for each learned field, as shown by the dashed line. C) The forward model. In the null field (thin line), the motor command $u = 0$ produce straight ahead movements (represented by $\hat{x} = 0$). By end of training in B, the forward model has learned that $u \approx -1$ produces straight ahead movements. Following training in A the forward model predicts that $u \approx +1$ produces straight ahead movements. Therefore, after AB training, the prediction errors produce catastrophic interference in the forward model.

When the perturbation changes to *A*, previously successful motor commands are no longer successful. This encourages a search for new motor commands (Izawa and Shadmehr, 2011a) and possible disengagement of the controller representing *B*. The critical hypothesis is at the end of *BA* training, we have multiple motor commands associated with success (Fig. 7B), with the mode reflecting motor commands that were successful in *A*. Therefore, reversal of the perturbation is unlikely to produce erasure in the controller because reward is simply shifted to new motor commands, creating a bimodal distribution reflecting the history of all learning. However, it is possible that by the end of *BA* the forward model acquired in *B* is catastrophically affected by training in *A* (Fig. 7C). The novel prediction is that when a motor memory shows spontaneous recovery, it reflects the output of the controller (which learns from reinforcement) and may not be accompanied with the appropriate forward model. This also leads to the prediction that spontaneous recovery will be absent in basal ganglia patients, but present in cerebellar patients.

This model can account for a number of observations in our data. We observed that memories were protected in the gradual condition. This protection may not rely on kinematic errors, but changes in success probability, which were substantial even in the gradual condition. Selection of the controller based on reward would also account for switching that we observed when reinforcement was withheld. Finally, we observed that BN training produced much less recovery of B than BA training (Fig. 3E). This is accounted for as the peak of the two resulting associations between motor commands and reward in the controller are closer in BN than in BA training, resulting in greater interference.

This model may relate to state-space models, in that the motor cortex may serve as the site for ‘slow states’ of learning, with ‘fast’ learning taking place in the cerebellum. Evidence for this case has been observed through double dissociation by applying tDCS over M1, which increases retention of learned motor memories with no effects on adaptation, with tDCS over the cerebellum increasing the rate of adaptation, without affecting retention (Galea et al., 2010).

2.4.2 Link to operant conditioning

In BAb and similar experiments protection of *B* was not due to a contextual cue from sudden errors, and spontaneous recovery of *B* was not due to a similarity between errors during testing and initial adaptation. Rather, the brain expressed *B* because the current motor commands (in null or *A*) were suddenly unsuccessful. This parallels observations in operant conditioning. For example, Mazur (1995) investigated the role of reward in pigeons that were trained to peck at two different keys. Each key delivered reward at a constant probability for a set number of trials, but then changed to a new reinforcement schedule. The pigeons were able to adjust to the new schedule, but at the start of a new block they reverted back to the previous schedule, indicating spontaneous recovery of prior training. That is, previously rewarded behaviors are not erased when new behaviors are rewarded, consistent with the view that in the AB paradigm, reinforcement of *B* motor commands does not erase the association of *A* commands with reward.

2.4.3 Limitations

Traditional methods of assaying motor memory have relied on measures of savings, i.e., training then re-testing on the same task. The implicit assumption has been that if the memory is present, then there are contextual cues in re-testing that should be sufficient to express it. This approach has produced the conclusion that motor memories are erased because there is no evidence of savings. Our work here shows that this conclusion is false, but does not explain why previous methods of assaying motor memory failed to observe protection. Though our model predicts that spontaneous recovery reflects the output of the controller, our current design cannot determine the actual component of motor learning present in error-clamp trials. Furthermore, our work does not address two important issues in motor learning: passage of time appears to strengthen motor memories (Brashers-Krug et al., 1996), and repetition appears to increase resistance to interference (Krakauer et al., 2005). It is unclear whether denial of reinforcement would produce spontaneous recovery that increases with passage of time, and with increased repetition. Finally, though we linked changes in success rates to spontaneous recovery of motor memory, these changes also altered levels of cognitive attention.

Chapter 3: Are motor memories tool-specific?

3.1 Introduction

When faced with a novel tool, people can learn the dynamics of the instrument and motor patterns necessary to interact with the tool and achieve a goal. However, studies of motor adaptation have found that people are unable to retain and flexibly switch between multiple motor patterns using the same tool, especially when no contextual cues are given to signal changes in tool dynamics (Brashers-Krug et al., 1996;Krakauer et al., 1999). Coupling distinct contextual cues with each perturbation can minimize interference during the acquisition and recall of multiple motor memories, however results vary on the efficacy of these cues.

The most effective contextual cues often rely heavily on proprioceptive and visual feedback of the movement. For example, in reaching tasks, using different workspace locations can allow for participants to learn opposing dynamical force fields (Hwang et al., 2006;Howard et al., 2013). Pairing opposing force fields with different arm postures also facilitates learning of the fields (Gandolfo et al., 1996). Alternately, when using the same hand space and limb posture, distinct visual locations can serve as effective cues for learning and switching between multiple motor memories (Hirashima and Nozaki, 2012;Howard et al., 2013).

Though some visual and proprioceptive cues aid learning of motor behaviors, other simple cues inexplicably are not effective in motor adaptation. In everyday life, we can learn to associate different colors with different behaviors, as in stopping at a red traffic signal, but accelerating when the signal again turns green. However, utilizing color as a context cue has yielded mixed results in motor adaptation, proving effective in only a few cases (Gandolfo et al., 1996;Krouchev and Kalaska, 2003;Wada et al., 2003;Cothros et al., 2009;Addou et al., 2011;Howard et al., 2013). Participants are also unable to learn from simple temporal structures, as when opposing fields are alternated after each trial (Karniel and Mussa-Ivaldi, 2002). Distinct haptic feedback has also proven to be ineffective. Neither the use of handles with different shapes nor adjusting thumb posture between applied force fields work to reduce interference in reaching tasks (Gandolfo et al., 1996;Cothros et al., 2008).

All of these studies have focused on acquiring and switching between multiple motor memories using the same tool. For these tasks, contextual cues are utilized to either signal a switch in the dynamics of the tool, or to create the illusion of switching between tools. In the present study, we investigated the training tool itself as a contextual cue. We utilized two different robotic tools, each with its own defining set of characteristics and paired each with an experimentally imposed dynamical environment. Participants alternated between the two robots across four days of training, completing the same center-out reaching task on each tool. Though the task was identical across the tools, we applied a counter-clockwise force field to the arm on one robot, and a clockwise force field to the arm on the other robot. We hypothesized that participants would learn to treat each tool as a separate entity, and that learning would not generalize across the robots. We also expected that participants who trained in additional experimental sessions beyond four days would learn to associate each tool with the given perturbation and would be able to immediately recall the appropriate, tool-specific behavior necessary to achieve success.

3.2 Methods

We recruited 42 healthy, right-hand dominant participants for our study (22.6 ± 3.4 years, mean ± SD, 27 males and 15 females). All participants were naïve to the purpose of the experiment, and had never previously participated in a reaching experiment in the lab. All volunteers provided consent by signing a form approved by the Johns Hopkins University School of Medicine Institutional Review Board.

3.2.1 Robotic Tools

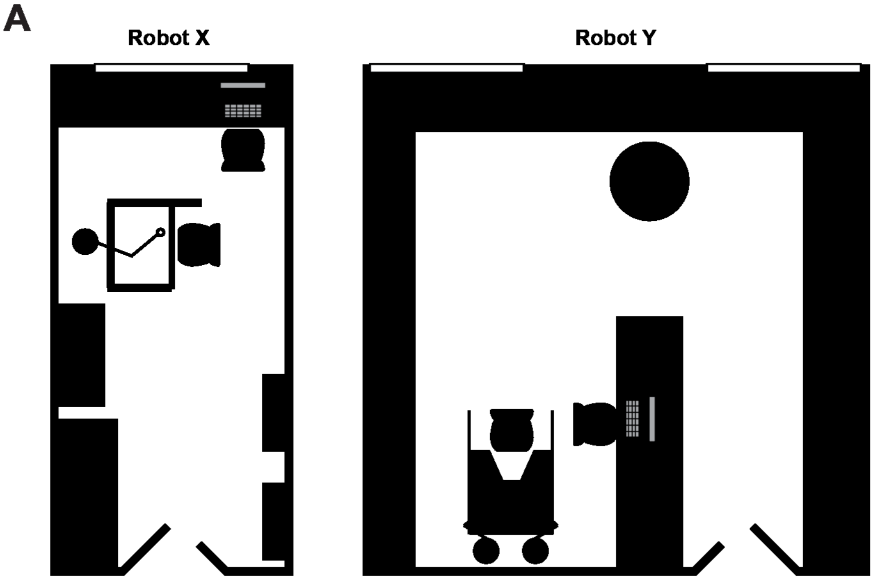
When participants arrived on the initial day of testing they were taken to one of two rooms, each housing a distinct robot, as shown in Figure 1A.

> **Robot X** is our unimanual robot. The participant interacted with this robot by holding onto a metal handle wrapped with a white, soft tennis grip, located at the end of the manipulandum. The outer frame of Robot X was enclosed by a dark surround which

prevented the participant from viewing the experimenter, who was seated to the right of the participant.

Robot Y is our bimanual robot. For this experiment, the participant only held onto the handle of the robot's right manipulandum. This arm has the same configuration as Robot X, but with an uncovered metal handle located at the end of the manipulandum. Here the outer frame of the robot was not enclosed, and the experimenter sat to the left of the participant.

The rooms housing Robot X and Robot Y differed in their physical location in the building, each having a separate entrance. These rooms also varied in square footage, ambient lighting, and the presence of additional furniture and lab equipment. Additionally, the preparatory activities prior to the onset of the experiment differed between the two robots. Robot Y requires a calibration routine using both of the robotic arms, which the participants witnessed. Participants also watched as the experimenter moved Robot Y's left manipulandum out of reach, beyond the edge of the experiment workspace, leaving only the right manipulandum for use. Finally, the structure separating the participant from the experimental workspace was different between the two robots. Robot X is open between the participant and the floor, whereas Robot Y has a screen below the participant's arm. To account for this and maintain arm posture between robots, a temporary metal bar was placed across both robots, to provide a lower and upper boundary through which the participant could reach. Additionally, the distance between the participant's shoulder and the floor and the distance between the participant's shoulder and the experimental set up was controlled each day, in order to maintain similar postures between robots and across sessions.



B

		Day 1	Day 2	Day 3	Day 4
NoSwitch + NoInterference	A_1A_1 B_1B_1				
Switch + Interference	$A_1B_2A_1B_2$ $B_1A_2B_1A_2$		Field B, Robot Y Field B, Robot X Field A, Robot Y Field A, Robot X		Field B, Robot Y Field B, Robot X Field A, Robot Y Field A, Robot X
NoSwitch + Interference	$B_1A_1B_1A_1$		Field A, Robot X Field A, Robot Y		Field A, Robot X Field A, Robot Y

Figure 3.1 Robotic tools and experimental paradigms (A) Participants held onto a robotic manipulandum to make 10 cm center-out reaching movements to one of four targets. Participants performed this task on one of our two robotic tools: our unimanual Robot X or our bimanual Robot Y. Each robot was housed in a separate room, each with a unique set of defining characteristics. **(B)** Across groups, participants were counterbalanced for field direction (clockwise or counterclockwise) and also for tool used. Participants in the NoSwitch+NoInterference group only learned one given field, and trained in this field using the same robot in two sessions separated by 48 hours. Participants in the Switch + Interference group trained on both fields and both robots, alternating field and robot each day for four consecutive days. In the NoSwitch + Interference group, participants learned both fields, but trained on the same robot across four days of training

3.2.2 Behavioral Task

Though we utilized different robots for this experiment, the behavioral task that the participant completed was identical across the two tools and across all days of participation. For both robots, a screen was placed above the participant's arm, blocking the view of the participant's hand and the robotic manipulandum. Visual feedback was provided at all times via a cursor (radius, 2.5mm) tracking the hand movement, which was projected onto the screen. Participants made reaching movements to guide this cursor from a circular start position (radius, 0.5 cm) to a circular target (radius, 0.5 cm) located 10 cm away. Only one target was present for any trial, appearing pseudo-randomly at 30°, 70°, 110°, or 150° from horizontal. If the duration of the participants reach fell between 500 ± 50 ms, the participants were informed of their success through an animated target 'explosion'. Unsuccessful trials were indicated through color feedback when participants were too fast or too slow. Following each trial, the robot guided the participant's hand back to the start location to initiate the next trial.

At the start of each day of training, participants first reached through an error-clamp (spring coefficient = 2500 N/m; damping coefficient = 45 N·s/m) to the target located at 110°, for 15 consecutive trials. Participants were informed that they would be reaching to four different targets over the course of the experiment, but that these initial 15 trials were repeated towards the same target in order to ensure that the participant learned the correct speed necessary to achieve target explosions. We confirmed that all participants were able to achieve at least one target explosion during this initial period.

Following these initial trials, all four targets were randomly presented for 168 trials. During these movements, a velocity-dependent curl field was applied to the participant's arm. These forces were dependent on hand velocity $\dot{x}$ via a viscosity matrix which could either be clockwise curl (Field A, [0 13; -13 0] N·s/m) or counterclockwise curl (Field B, [0 -13; 13 0] N·s/m). After completing this block, participants sat at the robot for an imposed 3 minute set break. Participants then reached through to all four targets again with the same curl field condition for an additional block of 192 trials. These two blocks of trials were repeated across each day of training for the participant's given experimental protocol. Error-clamp trials were randomly interspersed at a frequency of about 1 in 7 trials throughout these two blocks in order to probe force field compensation.

3.2.3 Experimental Groups

Experiment 1 A subset of participants (n=16) in Experiment 1 experienced our NoSwitch + NoInterference protocol, which is displayed in Figure 1B. This group served as a control in which we measured the retention of force field learning on the same robotic tool after passage of time. These participants came in on Day 1 and trained on one of two robots (X or Y) and adapted to one of two force fields (A or B). These subjects then returned after 48 hours (actual time 48.0 ± 0.40 hours, mean ± SD), and trained in the same field on the same robot that they had previously trained on. These participants were split evenly across four conditions to counterbalance the trained field and robot used. For example, the A_1A_1 group indicates that these participants all learned the clockwise Field A first, on 'Robot 1', and then retrained on Field A again on Robot 1.

Robot 1 was Robot X for four of these participants, and Robot Y for the remaining four participants.

Another subset of participants (n=18) trained in the Switch+Interference protocol. After training in field A or B on Day 1, on Robot X or Y, these participants returned after 24 hours (24.06 ± 0.38 hours, mean ± SD). For the session on Day 2, the participants trained on the opposite field and opposite robot they had experienced on Day 1. Following another 24 hour break (23.98 ± 0.23 hours, mean ± SD) they again switched both field and robot. Therefore, training on Day 3 was utilized the same robot and field combination the participant had experienced on Day 1. Following a final 24 hour break (23.95 ± 0.62 hours, mean ± SD), participants switched again, training on the same field and robot they experienced on Day 2. Again participants were split across four conditions to counterbalance the order of both robot and field. For example, in the $A_1B_2A_1B_2$ group, participants came in on Day 1 and learned Field A on an initial, 'Robot 1'. For four of these subjects this was Robot X and for five of these subjects, this was robot Y. The next day, participants now learned Field B on the second tool, 'Robot 2'. If the participants had previously trained on Robot X, they now performed Field B on Robot Y. Similarly, the other participants in the $A_1B_2A_1B_2$ group who trained on Robot Y on Day1 now trained on Robot X on Day 2. These participants switched two more times, so that Days 3 and 4 of training were identical to Days 1 and 2, respectively, with regards to the applied field and robot used. Two of these subgroups consist of n=5 participants, as two participants were added to this group later to serve as participants for our Experiment 3.

Experiment 2: With our next group of participants, we wanted to determine if switching between the two robots offered any benefit for training in the two possible curl fields. In the NoSwitch + Interference group (n=8), participants learned the counter-clockwise Field B on Day 1. Four of these participants started training on Robot X, with the other four participants training on Robot Y. On Day 2 of training, the participants now experienced Field A, but on the same robot they had trained on for Day 1. Using the same robot again, they then trained in Field B on Day 3 and Field A on Day 4. Therefore, these subjects experienced both force fields, but only trained on a single robot across the four days of experimentation.

Experiment 3: We observed that transfer across the robots decreased with an increasing number of training sessions. To investigate this further, we asked four subjects in the Switch+Interference group to continue their training through additional sessions. Two of these participants (23 year old female, 28 year old male) came in for 12 total sessions, and two of these participants (24 year old female, 26 year old male) came in for 14 sessions. Two of these participants experienced the $A_1B_2A_1B_2$ paradigm, one starting on Robot X and one on Robot Y. The other two participants experienced the $B_1A_2B_1A_2$ paradigm, with one starting on Robot X and the other on Robot Y.

3.2.4 Data Analysis

We recorded position, velocity, and force data at the robotic handle at a rate of 100 Hz and analyzed this data offline using Matlab R2009b and IBM SPSS Version 22. To measure compensation for the applied force field, we calculated an adaptation index (AI) for each error-clamp trial of training. The AI is equal to the ratio between the actual force produced by the participant and the ideal force for a given trial. Ideal force is measured by the applied curl field multiplied by the hand velocity. Using the average of the last 5 AI per day and the average of the first 15 AI per day (first 15 consecutive trials to the target at 110°), we also calculated a transfer measure across each day of participation. Here the transfer value was found by:

$$Transfer\ Value = 1 - \frac{AI_{End\ of\ Day\ n} - AI_{Start\ of\ Day\ n+1}}{AI_{End\ of\ Day\ n}}$$

For this equation, a transfer value of +1 indicates that almost all field learning was retained and expressed during the first 15 trials of the new session, as $AI_{End\ of\ Day\ n} \approx AI_{Start\ of\ Day\ n+1}$. A value of 0 indicates that very little learning transferred to these initial trials, as performance towards the end of every session was near full compensation of the field, meaning $AI_{End\ of\ Day\ n} \approx 1$. For the participants in the NoSwitch + NoInterference group, transfer values could only be measured between Days 1 and 3. For the remaining participants, transfer was calculated between all days of training.

We also measured performance during the fielded trials via the maximum perpendicular displacement of the hand from a straight line path connecting the start position to the target. In order to account for the different target locations, we combined these measures of displacement into bins of four movements across each block.

3.3 Results

When participants arrived to the lab on the initial day of testing, they were seated at one of our two robotic tools, Robot X or Robot Y, as shown in Figure 1A. Once seated, the participants first reached through an error-clamp to a target located at 110° from horizontal for 15 consecutive trials. As our two robotic tools are not identical, we first needed to verify that there were no differences between the robots which may have influenced the initial behavior of the participants. Using a t-test analysis and focusing on the reaches that all participants (n=42) made when they first sat down on Day 1, we observed that there were no statistical differences between the robots for mean peak velocity ($t_{(40)}$=-0.283, p = 0.779) or the mean maximum displacement of the first 15 trials ($t_{(40)}$=-1.470, p = 0.149).

As these first 15 reaches were in error-clamp, we also analyzed the average adaptation index in these trials and saw no influence of robotic tool on the magnitude of these forces ($t_{(40)}$=-0.325, p = 0.747).

Following these initial trials, participants then immediately reached through either Field A (a clockwise curl field) or Field B (a counterclockwise curl field) for the remainder of the experimental session. We also wanted to verify that performance was similar in each field across the two robots, and to also investigate if there were statistically significant behavioral differences due to field direction or an interaction between robot and applied field. Using a two-way ANOVA, we found that there was no effect of robot ($F_{(1,41)}$ = 0.808, p = 0.374), of field ($F_{(1,41)}$ = 0.056, p = 0.814) or an interaction between robot and field ($F_{(1,41)}$=0.684, p = 0.413) on the initial hand displacement during fielded trials on

Day 1 (maximum displacement in first bin of trials). Analyzing the end of training performance after participants adapted to the field on Day 1, a two-way ANOVA indicated no effect of robot ($F_{(1,41)}$=0.811, p = 0.373), of field ($F_{(1,41)}$=3.144, p = 0.084) or of field*robot interaction ($F_{(1,41)}$=0.335, p = 0.566) on the last 5 AI of training.

Finally, we wanted to confirm that between day transfer of learning was unaffected by the robotic tool used and field applied to the hand. In order to analyze transfer values, we looked at two-way ANOVAs within the NoSwitch + NoInterference group (n=16) separately from the Switch + Interference Groups (n=18). For the NoSwitch + NoInterference group, we found no effect of robot ($F_{(1,15)}$ = 3.051, p = 0.106), field ($F_{(1,15)}$ = 0.895, p = 0.363) or robot*field interaction ($F_{(1,15)}$ = 0.144, p = 0.711) on the transfer values measured between Day 1 and Day 3. For the Switch + Interference Groups we again saw no effect of robot ($F_{(1,17)}$ = 3.394, p = 0.087), field ($F_{(1,17)}$ = 0.126, p = 0.728) or field*robot interaction ($F_{(1,17)}$ = 0.028, p = 0.869) for transfer between Days 1 and 2.

3.3.1 Force field learning transferred across tools

As no differences were found in the reach kinematics or learning behavior due to applied field or robot used, we combined the experimental paradigms shown in Figure 1B to yield a NoSwitch + NoInterference group and a Switch + Interference group. With the Switch + Interference paradigm, we asked whether force field training would transfer across robots, leading to interference when adapting to opposing fields. Our NoSwitch + NoInterference group served as a control to quantify force field retention and faster relearning on the same robot and same field combination.

For both groups, participants adapted to an applied force field over the course of two blocks on Day 1. As shown in Figure 2A, by the end of the first day of training participants in both groups adapted to the applied field, achieving an average adaptation index of 0.86± 0.07 and 0.88 ± 0.10 (average of last 5 AI, mean ± SD) in the NoSwitch + NoInterference and Switch + Interference groups, respectively.

Participants in the NoSwitch + NoInterference group came back for another training session after a 48 hour break and were seated at the same robot that they had used for training on Day 1. At the start of this Day 3 training, these participants again reached to the target at 110° for 15 consecutive trials in the error-clamp condition. Shown in Figure 2A, these participants now had an initial bias in their forces that was larger in magnitude than on Day 1 ($p < 0.001$, paired t-test). Importantly, the forces exerted against the channel wall on Day 3 were in the same direction necessary to counter the forces experienced on Day 1, indicating that participants had retained a portion of their force field training and exhibited this retention even before revisiting the applied perturbation. Using the AIs calculated from these first 15 trials and the AIs from the end of performance on Day 1, we calculated a transfer index to find the percentage of learning that each individual participant retained from the previous training session. As shown in Figure 2B, these participants exerted forces equivalent to 37.20±0.46% of the forces they produced at the end of learning on Day 1, an amount that was significantly greater than 0 ($p<0.001$, z-test).

The participants in the Switch + Interference group were also retested on Day 3 using the same robot and reaching through the same field they had experienced on Day 1. These participants however, experienced an interfering task in between these two test sessions. These participants were tested 24 hours after Day 1, but during this session participants were seated in front of the robot they had not trained with on Day 1. We first asked if the participants would generalize their learning from Day 1 to this new robotic tool they had never previously seen or interacted with.

During the first 15 trials of Day 2, the participants in the Switch + Interference group exerted forces against the error-clamp that were greater in magnitude than their initial bias on Day 1 ($p<0.001$, post-hoc comparison of Day 1 vs Day 2 following a significant repeated measures ANOVA on average AI in first 15 trials of each day, $F_{(3,51)} = 12.144$, $p<0.001$). Crucially, as with the NoSwitch + NoInterference group, these participants exerted forces in the same direction necessary to counter the field learned on Day 1. This behavior again indicated that participants retained learning from the previous day's session, and generalized a portion of the learning, even before experiencing any perturbations using the new tool. What differed in this group was the fact that participants generalized learning to a new tool that they had never seen or experienced before. We again calculated a transfer value to determine what percentage of learning was retained and transferred, as displayed in Figure 2B. We found that these participants exerted forces equivalent to 61.15±32.07% of the forces at the end of Day 1 training. Again this average transfer value was significantly different than 0 (z-test, $p<0.001$).

After the initial first 15 error-clamp trials, participants in the NoSwitch + NoInterference group now reached through the curl field that was in the opposite direction of the field they had experienced on Day 1. Participants were able to adapt to this opposing field, achieving an adaptation index of -0.83±0.09 (average of last 5 AI, mean±SD) by the end of training.

When participants came back for an experimental session on Day 3, they again switched robots, returning to the robot they had experienced on Day 1. We expected one of two options. First, we imagined that participant's would continue to transfer learning from the previous day's session, as we had observed between Days 1 and 2. For this scenario, we would again expect a large positive transfer value, indicating that 24 hours after practice, participants expressed forces in the same direction as the previous days training. Alternatively, participants could have learned to associate a particular field with a given robot. If this were the case, we expected a negative transfer value at the beginning of Day 3. This was due to the fact that participants trained on opposing fields for Day 1 and Day 2. A negative transfer value would indicate that the participants had bias in the direction opposite of the Day 2 training, but in the same direction as on Day 1.

What we observed however, was neither of these two options. When participants came in on Day 3, they had an average transfer value of 0.07±0.27, values which were not significantly different from 0 transfer (z-test, $p = 0.13$). This suggested that perhaps participants were expressing an average of all training sessions, rather than just the previous day's learning or the learning they had associated with that particular robot.

Following the initial 15 trials, participants adapted to the same field they had experienced on Day 1, reaching an average AI of 0.89±0.13 (mean±SD, last 5 AI) by the end of training. This end of day force compensation was not significantly different from the end of learning on Day 1 (no effect of session, repeated-measures ANOVA for average of last 5 AI across Days 1-4, $F_{(3,51)} = 1.588$, $p = 0.204$).

These participants then came back for a final testing session on Day 4. Again they switched robots, using the opposite robot they had trained on during Day 3, but the same robot they had used on Day 2. At the beginning of Day 4, these participants exhibited transfer values of 0.38±0.27 (mean±SD), an amount significantly greater than 0 transfer (z-test, $p<0.001$). As these values were positive, this indicated that participants were producing forces similar to the end of the previous day's training, not in the same direction of the field they had previously trained on with this robot. These values however, were statistically smaller than the transfer values observed after Day 1 training, (post-hoc comparison, $p<0.001$, following a significant effect of session, repeated-measures ANOVA $F_{(2,34)}=14.428$, $p<0.001$).

Overall, participants retained a portion of their learning from force field adaptation, and expressed this retention by exerting these forces as soon as they sat down at the robot. We measured this via 15 consecutive trials to a single target in error-clamp, prior to re-introducing the field each day. Participants exhibited these forces even when using a tool they had previously not encountered. With continued task and robot switching, participants expressed a running average of the fields they had learned. However, it appeared that participants began to transfer less learning the more they switched robots and fields.

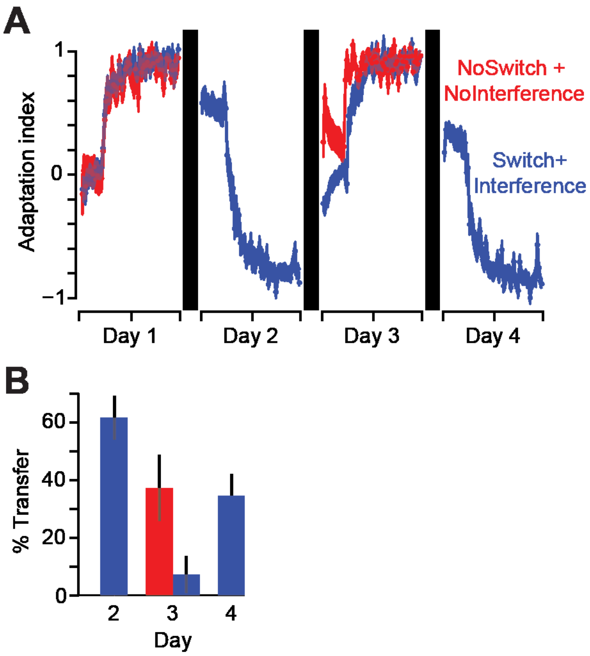


Figure 3.2 Force field training generalized across robotic tools (A) Participants in the NoSwitch + NoInterference group retained training from Day 1, expressing a force bias at the start of Day 3. Participants in the Switch + Interference group also exhibited retention of Day 1 training, expressing a force bias on Day 2. This generalization was across our robotic tools, as participants were now training on the opposite robot they had used on Day 1. When participants in the Switch + Interference group returned to the same robot on Day 3, they did not express a force bias similar to the NoSwitch + NoInterference group on Day 3. This suggested that participants were expressing a bias reflecting the total history of training **(B)** Transfer values were calculated to find the amount of between day transfer, normalized by each participants end of session performance. Notably, participants in the Switch + Interference group transferred less between Day 3-4 as they did between Day 1-2. Data are all plotted as mean ± SE.

3.3.2 Interference was also observed during force field training

We also measured hand displacement due to the applied perturbation on each day's experimental session in order to see if each participant's behavior during fielded trials reflected the same patterns observed when analyzing the initial force biases per day. Plotted in Figure 3A, we observed that on Day 1, participants were initially displaced by the applied field, but gradually reduced their errors and compensated for the field. To quantify this behavior, we analyzed two measures. First, we analyzed the initial displacement, measured as the first bin of movements and error reduction. Similar to other studies of savings (Krakauer et al., 2005), we saw a quick reduction of error over the first few bins of training. To quantify this learning, we measured error-reduction as the average maximum displacement across bins 2-11.

Comparing the NoSwitch + NoInterference group with the Switch+Interference group, we saw no differences between the groups with regards to their initial error or error reduction (t-test, $p = 0.461$ and $p = 0.944$) on Day 1. This was expected as it was the first session for all participants. In the NoSwitch + NoInterference group, participants returned after 48 hours to train on the same field. Upon reintroducing the field, the participants were initially displaced less (paired t-test on bin 1, $p<0.001$) and reduced their error faster (paired t-test on average displacement bins 2-11, $p<0.001$) than they had on Day 1, as shown in Figure 3B. This was in agreement with the observations made from the force bias on Day 3, suggesting that participants had retained a portion of their training from Day 1 and now exhibited savings, or faster relearning.

We also observed significant differences between Day 1 and 3 for both initial displacement (post-hoc comparison, $p = 0.004$, following significant effect of session, repeated measures ANOVA $F_{(3,51)}=27.463$, $p<0.001$) and error reduction (post-hoc comparison, $p = 0.041$, following significant effect of session, Repeated measures ANOVA, $F_{(3,51)}=24.968$, $p < 0.001$) for participants in the Switch + Interference group.

However, as displayed in Figure 2A, this difference was in the opposite direction of what was observed in the NoSwitch + NoInterference group. On Day 3, the subjects in the Switch + Interference actually performed worse in the field when they were relearning it, exhibiting a larger initial displacement and slower relearning than on Day 1. Again this was in agreement with our previous observation that participants retained what they had learned on the previous day of training and generalized this learning to the new tool. However in this case, participants had learned the opposing field on Day 2 and retention of this training now resulted in interference.

However, savings were observed between Days 2 and 4 for the participants in the NoSwitch + NoInterference Group. We saw statistically significant differences in the initial displacement (post-hoc comparison, $p<0.001$, following significant effect of session, Repeated measures ANOVA, $F_{(3,51)}=24.968$, $p < 0.001$) and for the error reduction (post-hoc comparison, $p< 0.001$ following significant effect of session on repeated measures ANOVA $F_{(3,51)}=27.463$, $p<0.001$). This further supported the observation that with increased number of task and robot switching, participants began to transfer less.

In our task, participants were never exposed to a null condition, in order to allow for washout of the previous days training. As participants started off with a bias in their behavior upon entering the field each day, we could not measure true savings, in the sense that participants did not start with the same amount of displacement between sessions. However, if participants had learned to couple the robot with its given field, we could have expected that after initial displacement, participants would recall the necessary motor commands to counter the field, and rapidly reduce their error. For the NoSwitch+NoInterference group, this did seem to be the case. After initial exposure to the field, participants quickly reduced their errors on Day 3. However, participants in the Switch+Interference group showed significant relearning on Days 3 and 4, despite having encountered these fields before.

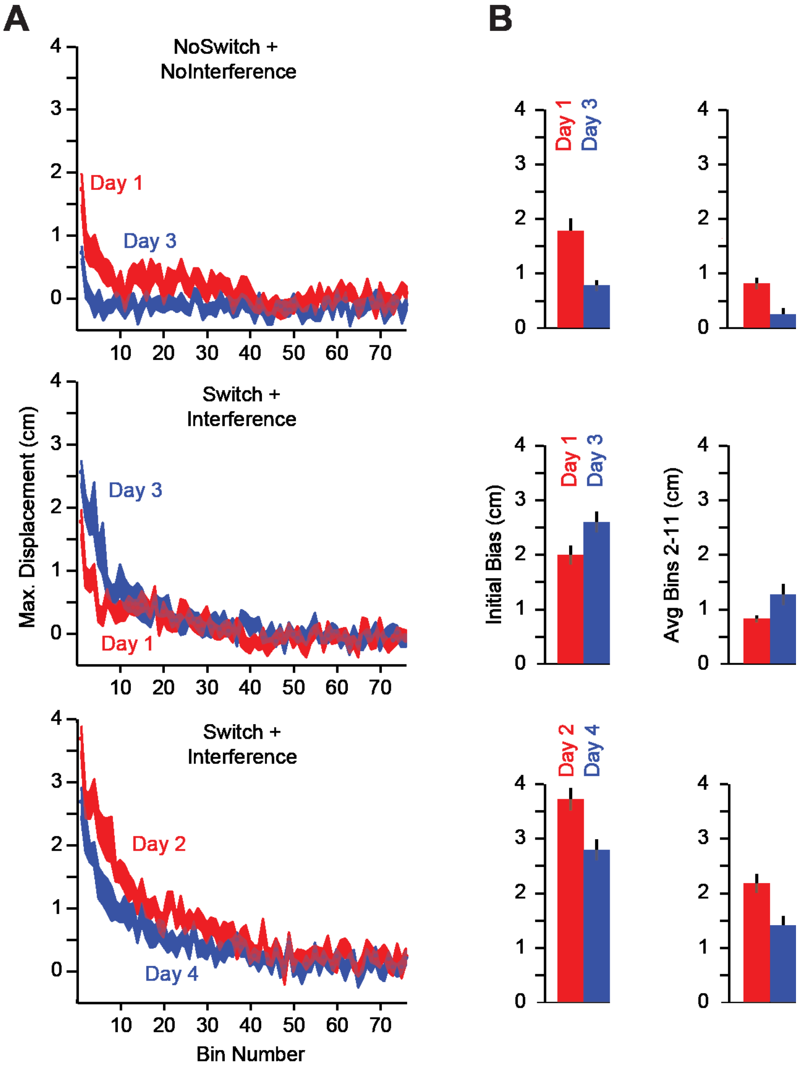


Figure 3.3 Field performance improved with repeated switching (A). Participants in the NoSwitch + NoInterference group displayed faster relearning on Day 3. In the Switch + Interference group, participants performed worse on Day 3 than they had on Day 1, during initial training. However, participant's performance on Day 4 was improved as compared to Day 2, indicating a benefit of continued switching. **(B)** Mean maximum displacement in field during the first bin of four trials **(C)** Average error maximum displacement between blocks 2-11. Data are all plotted as mean ± SE.

3.3.3 Tool switching did not reduce interference

We next wanted to understand if pairing force fields with different robotic tools offered any benefit in reducing interference between learning the opposing fields. To answer this question, we recruited a group of participants for the NoSwitch + Interference group (n=8). For this paradigm, participants alternated fields across days of training, but they did not switch robots. Instead these participants trained on the same robot across all four days of experimentation.

The performance of these participants is plotted in Figure 4A. In the NoSwitch + Interference group, participants only trained in the B-A-B-A field order, so the performance of participants who trained in the $B_1A_2B_1A_2$ paradigm of the Switch + Interference group is also plotted in Figure 4 for comparison. We first verified that the ability to compensate for both force fields was equivalent across the two groups. Analyzing the end of day force compensation, we found no statistically significant effect of session ($p = 0.273$, repeated-measures ANOVA on average last 5 AI, $F_{(3,45)}=1.340$), or group*session interaction ($p = 0.131$, $F_{(3,45)}=1.978$), indicating that all participants performed equally well across the different fields, with or without switching tools.

Using a repeated-measures ANOVA, we next analyzed the average AI of the first 15 trials per day to compare the initial force bias per session. We found that the initial performance across groups was not statistically different. While we found a significant effect of training day ($p = 0.000$, $F_{(3,45)} = 62.410$, $p = 0.000$), there was no significant group*training day interaction (p=0.948, $F_{(3,45)}$=0.119). Finally, we also analyzed transfer values and found a significant effect of session (p = 0.028, repeated-measures ANOVA $F_{(2,30)}$=4.023) but no session*group interaction (p = 0.417, repeated-measures ANOVA $F_{(2,30)}$ = 0.901).

Overall, we found no evidence that participants in the Switch + Interference group benefited from switching robots every time the field was switched. Instead, the level of transfer was equivalent to participants in the NoSwitch + Interference group, who did not switch robots and instead learned both fields on the same robotic manipulandum.

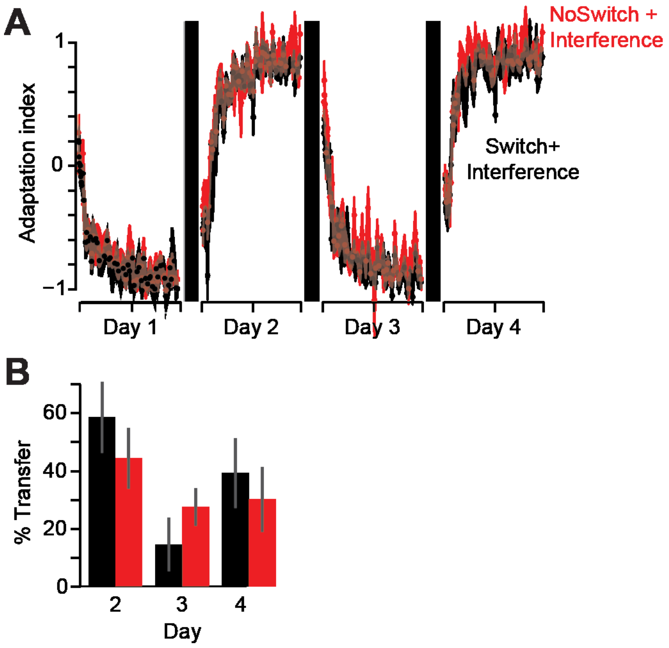


Figure 3.4 Performance of NoSwitch + Interference group (A) Participants in the NoSwitch + Interference group trained on both force fields, alternating fields across days, using the same robot for all four days of training. Performance was statistically indistinguishable from participants in the Switch + Interference group, who paired each field with a different robot. **(B)** Transfer values for each day of training. Data are plotted as mean SE.

3.3.4 With repeated switching, participants reduced across-tool transfer

We also wanted to further investigate our observation that in the Switch + Interference group, there was noticeably less transfer of learning between Days 3 and 4 as compared to Days 1 and 2. Participant's appeared to benefit from having previously experienced the switch between fields and robots. To go after this question, we requested that 4 participants from our Switch + Interference group continue their training across more sessions. Two of these participants came in for a total of 12 sessions and the other two participants came in for 14 sessions.

The results of this extended training group is plotted in Figure 5. Across training sessions, the between day change in initial bias gradually decreased, with participants settling on a near constant initial force bias. For 3 of the 4 participants, transfer was almost completely eliminated, and these participants reduced their initial force bias to near baselines levels. The remaining participant also eventually settled on a constant initial bias, however this bias was significantly in the direct of compensation for the counter-clockwise field.

Despite reducing transfer, we did not find that participants were able to immediately recall the motor commands necessary for a given robot, upon sitting down at the tool. Participant's also never achieved explicit knowledge of the differences between the tools. Upon completion of the task, when questioned, participants were unable to recall any differences between their experiences on the two robots, and were surprised to learn about the alternating fields.

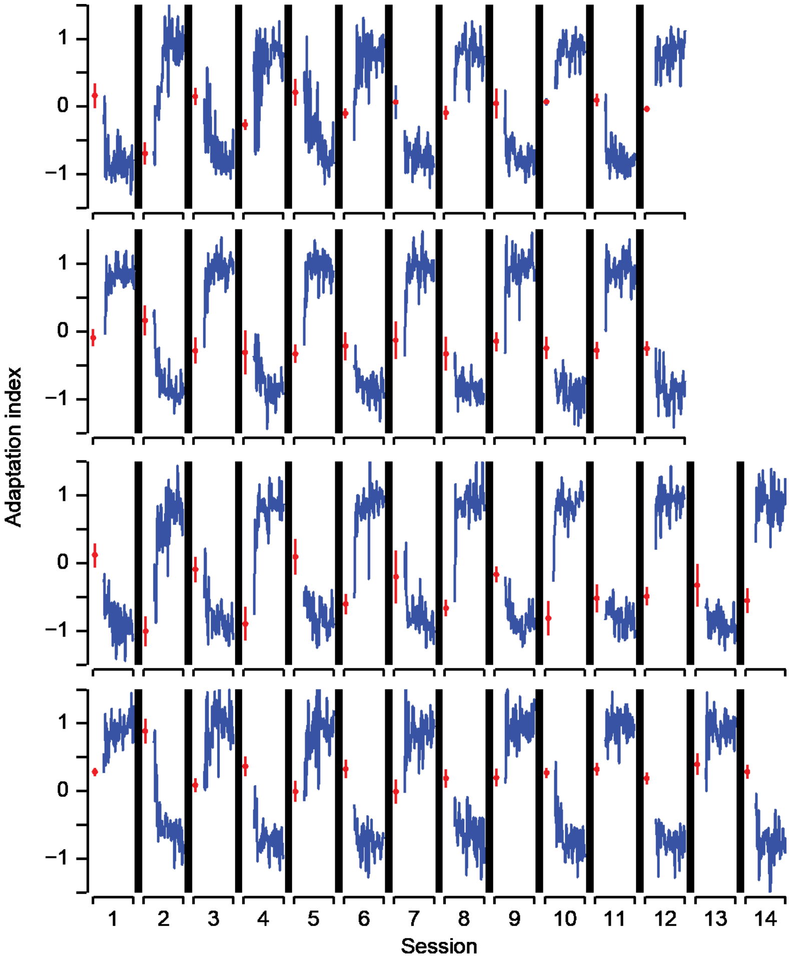


Figure 3.5 Transfer decreased across many sessions. Four participants in the Switch + Interference group continued training for 12-14 days. Across these sessions, participants decreased their between day transfer of learning.

3.4 Discussion

For our study, we asked whether force field adaptation in a reaching task is specific to the tool on which training occurred. To investigate this question, we trained participants to counter two opposing velocity-dependent force fields, alternating fields across days of training. For a subset of participants, the switch in force fields was coupled with a switch in the robotic tool. We found however that the contextual cues distinguishing the two robotic tool as separate entities were not effective, as participants in this group transferred just as much learning across days as participants who never switched tools.

This result was somewhat surprising, given that the way we interact with and manipulate objects is strongly influenced by visual cues. For example, visual input about the shape and orientation of an object is enough to allow participants to anticipate and adapt necessary grip forces (Jenmalm and Johansson, 1997). We stretch and shape our fingers while reaching towards an object in view, and begin to close our fingers before making contact (Jeannerod, 1984). When we compare objects we expect that weight scales with size, so that we predict a smaller object will weigh less than an object with greater volume. This prediction, known as the size-weight illusion, also influences lift forces. When picking up a small, heavy object, participants initially underestimate the object's weight. With repeatedly lifting, participants learn to adjust the necessary load forces, however the illusion itself persists beyond this learning (Flanagan et al., 2008). Our robots had several distinguishing visual features, including the frame surrounding the manipulandum, the number of arms on each robot, the robotic handle itself, the location of the experimenter relative to the participant, and the entire environment of the robot room. Despite the existent of many cues available to differentiate the tools, it appeared that participants still treated the robots as identical tools.

Though significant transfer was observed between robots across the initial four days of testing, we did observe that transfer began to attenuate by the fourth day of testing. Building on this observation, we extended training for four participants in our study. However, even after 14 days of testing, we found that participants were unable to rapidly switch between robots. In other studies of task switching, long-term training has proven to be effective. After many days of training, colors cues can serve as an effective contextual cue for humans and monkeys to anticipate a switch in experimentally imposed environments (Krouchev and Kalaska, 2003;Wada et al., 2003;Addou et al., 2011). In a study of prism adaptation, participants learned to flexibly switch between prism-on and prism-off conditions over the course of six weeks (Martin et al., 1996). Early in training, participants gradually adapted to the prisms, learning to accurately throw a ball with distorted visual feedback of the target location. Once removing the prisms, they exhibited significant after effects. However, after the long amount of training, participants were able to put the prism goggles on and accurately throw on the first trial. When they removed the goggles, again they were immediately accurate. Notably, the authors did find that the goggles themselves probably served as the contextual cue for switching. When introduced to a novel prism within the same goggle frame, participants incorrectly recalled the motor commands necessary to adjust to the well-trained prism. It is possible that our training session was not long enough, and with additional sessions, participants may have eventually learned the correct association between robotic tool and applied force field.

Other studies of motor adaptation have also looked at tool-specificity by measuring the transfer of learning from robots and treadmills to natural movements, such as reaches in free space and over-ground walking. Adaptation on a split-belt treadmill results in limited transfer to natural movements made in over-ground walking (Torres-Oviedo and Bastian, 2010). The authors also found that greater transfer occurred when vision was removed from the participant. This suggested that when participants were denied the visual cues indicating they were walking on treadmill, this led to a memory that was not tool specific. Following force field adaptation in reaching tasks, after-effects have been observed when participants let go of the robot to make reaches in free space, and when participants made reaches holding onto the robotic handle after it was detached from the robotic manipulandum (Cothros et al., 2006;Kluzik et al., 2008). However, these aftereffects are smaller in magnitude than reaches made in a washout or null condition, with the handle attached to the robotic arm. Additionally, these reaches in free space, with or without the detached handle, did not cause unlearning of the force field training. This suggests that force field adaptation is linked to the context of the experimental setup, with the robotic handle being attached to the robotic manipulandum. Given the specificity observed in these studies, it is again surprising that participants were unable to dissociate the two robotic tools used in our study.

Understanding which contextual clues link training to tools and developing strategies to encourage participants to generalize learning outside of the training context would be of great importance for rehabilitation. The use robots and treadmills as a means for rehabilitation after stroke, has resulted in some evidence of adaptation as an effective therapy in reducing motor impairments (Patton et al., 2006;Reisman et al., 2009). These results have only been observed on a short timescale however. For example, for at least 2 hours after training, prism adaptation can result in after effects that help restore function in patients with right hemisphere stroke suffering from hemispatial neglect (Rossetti et al., 1998).

Identifying which contextual cues are tied to behaviors and memories is also crucial for reconsolidation strategies. By introducing the correct trigger, we can cause an individual to recall a memory. Once recalled, this stored information becomes labile again (Nader et al., 2000). During the labile period, new information can be introduced to allow for the extinction of the previously consolidated memory. This intervention strategy has proven effective erasing in fear memories in humans, which could extend to future therapies for anxiety disorders such as post-traumatic stress disorder (Schiller et al., 2010). Retrieval-extinction has also proven effective in reducing cravings in heroin addicts (Xue et al., 2012). From these results, it is reasonable to conclude that a robotic manipulandum itself is the strongest contextual cue, and can recall previous training on that tool, regardless of the surrounding environment or extraneous tool features.

In conclusion, we found that force field adaptation transferred across two similar, but distinctly separate robotic tools. With repeated switching, participants were able to reduce this transfer. However, participants were unable to learn the correct association between robot and the experimentally imposed environment, even after 14 days of experimentation.

Chapter 4: Optimizing effort: increased efficiency of motor memory with time away from practice[2]

4.1 Introduction

In making a movement, there are numerous muscle activation patterns that can produce success. For example, to hit a tennis ball into the service court, some players twist and turn their bodies as they reach to hit the ball, while others reach more gracefully. Theoretical approaches suggest that learning of an action should favor production of the less effortful movement, i.e., effort should carry a cost (Salimpour and Shadmehr, 2014). Indeed, there is some evidence for this idea. With practice, there is often a reduction in muscle co-contraction (Thoroughman and Shadmehr, 1999;Franklin et al., 2003;Darainy and Ostry, 2008) and metabolic expenditures (Huang et al., 2012). Furthermore, given a choice between reaching movements that require various levels of effort, people (Cos et al., 2011;Wang and Dounskaia, 2012) and monkeys (Pasquereau and Turner, 2013) have a preference toward the less effortful reach.

However, there are also examples in which despite practice, more effort is expended during a movement than is necessary. For example, as people reach, muscles about the

---

[2] The work in this chapter was published in The Journal of Neurophysiology: Pekny SE and Shadmehr R (2015) J Neurophysiol 113:445-454

wrist joint are activated in order to counter the torques resulting from rotations of the elbow and shoulder. However, when the wrist is mechanically immobilized to eliminate the effects of these ‘interaction’ torques, the wrist activations persist (Koshland et al., 2000). When people train to reach in a field in which a straight point to point trajectory requires more force than a very curved trajectory, they choose the straight trajectory, even after hours of practice (Kistemaker et al., 2010) (but see (Izawa et al., 2008) and (Uno et al., 1989)). Finally, following force field training, experimenters can replace the field with an error-clamp in which the hand is constrained to a straight path to the target and the production of the field-compensating motor commands is no longer necessary. In these error-clamp trials, participants continue to produce field-specific forces (Scheidt et al., 2000), even after hundreds of trials (Pekny et al., 2011;Vaswani and Shadmehr, 2013). These examples highlight instances in which there are less effortful motor commands that can produce success, but the brain appears unable or unwilling to select them.

One possibility is that straight reaching movements may be the habitual response in certain conditions, more resistant to change, and therefore a less than ideal paradigm for short-term studies of the processes that may be involved in optimization of effort. Here, we thought to approach the question of effort optimization from a different perspective. We designed a task that involved learning a new motor behavior in which there was a natural tendency to produce inefficient motor commands. We then quantified the conditions that were required for reducing these inefficiencies.

We designed a task in which learning a movement resulted in robust inefficiencies, i.e., production of forces that were unnecessary for success. We then altered the conditions in which the learning took place in order to uncover the elements that were important for reduction of inefficiencies.

For our study, we considered a standard force field paradigm in which the participants reached to a target and a field was applied to their hand (Shadmehr and Mussa-Ivaldi, 1994b). In order to successfully complete the task and hit the target, the participants learned to produce forces that countered the field. However, this learning led to generalization, i.e., extrapolation of training to novel regions of the task space (Hwang and Shadmehr, 2005). We measured generalization by asking the subjects to reach to a probe target and measured the forces that they produced during that reach. Importantly, we designed the task in such a way as to make production of these forces unnecessary for success.

To achieve this inefficient but natural behavior, the probe target was always presented in an error-clamp. In these error-clamp probe trials there were no kinematic consequences of the effortful motor commands, as the reach was clamped to a straight line trajectory to and from the target. Therefore, the production of forces against the clamp had no bearing on success. Without performance errors or explicit knowledge of the task design, the participants were left with an implicit measure of effort as the only cue to indicate the inefficient nature of the forces that they produced. Could these task-irrelevant motor commands be reduced?

Surprisingly, we found that extended practice did not result in a decrease of the task-irrelevant force production. Rather, a critical factor was time away from practice; following a break of 6 or 24 hours in duration, but not of 3 or 30 minutes, the task-irrelevant force production spontaneously decreased. Therefore, while increasing the number of practice trials alone did not result in optimization of effort, it was time away from practice that afforded the opportunity to increase the efficiency of behavior.

4.2 Methods

Human volunteers (n= 41, 24.84 ± 5.44 years old, mean±1SD, including 18 men and 23 women) learned to make out-and-back reaching movements. All volunteers were naïve to the purpose of the experiment, were right hand dominant, and reported no known neurological disorders at the time of testing. Each participant provided written consent. Our procedures were approved by the Johns Hopkins University School of Medicine Institutional Review Board.

Participants held the handle of a two joint robotic manipulandum and made out-and-back reaching movements along the body midline from a 1 cm^2 start target to a 1 cm^2 goal target located 10 cm away. The goal target appeared at one of two locations: 90 □ (train target) or 270 □ (probe target), a reach to the goal target and then back to the start position in a continuous motion without stopping. Visual feedback was provided at all times via a 5 mm^2 cursor representing real-time hand position projected onto a screen covering the participant's hand. Reaches were considered successful when the total movement time was within 830-970ms, peak tangential velocity was within 0.20-0.60m/s, and participants had their turn-around point within a 6 mm diameter of the center of the goal target. Information regarding success or failure was provided once the hand returned to the start position. The goal target was animated to resemble an explosion in the case of a successful trial.

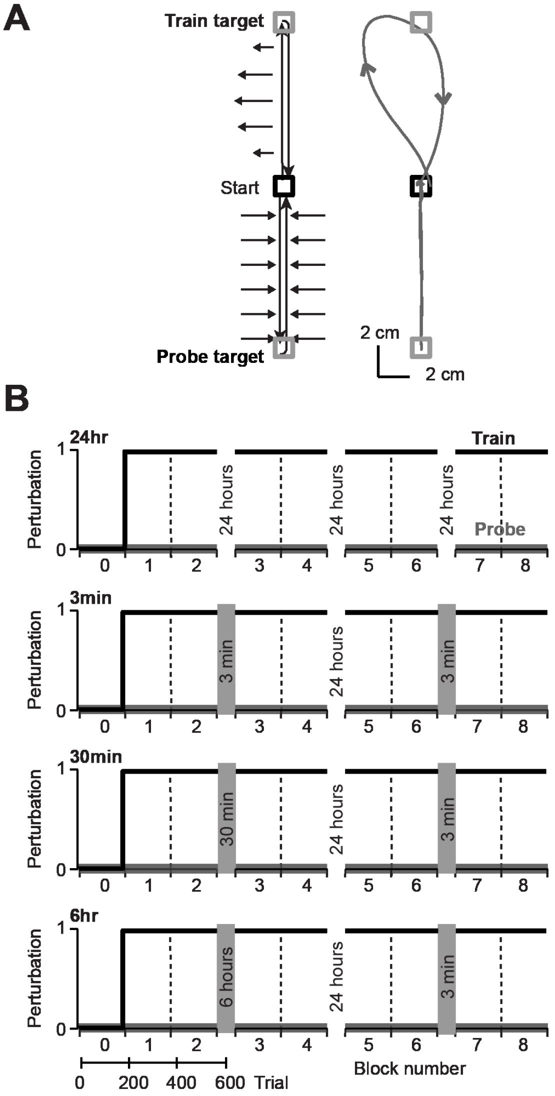


Figure 4.1 Task set-up and experimental groups. A) Participants held the handle of a robotic manipulandum and made 10 cm out-and-back reaching movements to two different targets. For the train target, a velocity dependent curl force field was applied on the outward reach. The field was turned off once the turn-around point of the reach back to start was detected. For the probe target the reaches were always in error-clamp. **B)** Perturbation protocols. Participants were provided a short set break (1 min) between blocks (indicated by dashed lines). Each block consisted of 192 trials, with 169 train targets and 24 probe targets, randomly interspersed.

Targets appeared in one of two locations, as shown in Fig. 1A, and were labeled as train target and probe target. For the train target, a counter-clockwise velocity-dependent curl force field with a gain of 13 N.s/m (Brashers-Krug et al., 1996) was imposed on the hand during the outward phase of the reach. Once the hand reached the train target and a turnaround point was detected, forces were shut off (Figure 1A). That is, for the train target the field was present only in the outward reach and not in the return back to start. Furthermore, on randomly selected reaches to the train target (1 out of 8) we imposed an error-clamp. In these trials, the hand was constrained to a straight path both out to the train target and during the return to start. Error-clamp was a channel centered between the start and goal positions, and implemented via a stiff one-dimensional spring (spring coefficient = 2500 N/m; damping coefficient = 25 N•s/m).

The probe target was presented in 1 out of 8 randomly selected trials (otherwise the train target was presented). Crucially, for the probe target the reaches were always in error-clamp. Therefore, any forces that subjects produced in the probe trials against channel walls were unnecessary as there was never a force field for the probe target. However, we expected that learning of the train target would be generalized to the probe target. The question was whether the inefficiencies inherent in this generalization could be reduced. Each block was 192 trials. Subjects were given a set break of ~1min between each block.

4.2.1 Experimental Groups

24Hr group: The protocol is summarized in Figure 1B (24hr group). Participants (n=10, 27.00 ± 7.05 years old, mean±1SD, including 5 men and 5 women) completed a baseline block in the null field (192 trials, no perturbation), and were then presented with two blocks of practice (192 trials each) in which they reached in a force field to the train target, and reached in an error-clamp to the probe target. They then left the lab and returned at approximately 24 hours (24.24 ± 1.86 hours, mean ± 1SD). On Day 2, subjects trained in two more blocks of trials (192 trials each). Importantly, the first two trials on Day 2 were in error-clamp. At the start of block 3, immediately following the 24 hour break, participants first reached to an error-clamp in the train direction, followed by an error-clamp to the probe direction. This was consistent across all participants, and all groups, and allowed us to measure the retention of force production following training but prior to the experience of any further performance errors. This pattern was repeated on Days 3 and 4. The sessions were performed at approximately 48 and 72 hours after the start of initial training (47.96 ± 1.36 hours, 71.96 ± 1.60 hours, mean ± SD), for a combined total of four consecutive days.

3min group: The results of the 24hr group suggested that passage of time may provide the conditions necessary to reduce generalization, as evidenced by overnight reductions in the forces that subjects produced to the probe target. To dissociate between effects of practice vs. effects of time/overnight sleep, we recruited a new group of subjects, the 3 minute group (3min, Figure 1B). Participants (n=10, 21.20 ± 3.36 years old, mean±1SD, including 5 men and 5 women) began their training in one block of null and two blocks of field trials, exactly as in the 24hr group. However, they continued their training with an additional two blocks of field trials after a 3 minute break. This paradigm effectively doubled the amount of practice on the first day as compared to the 24hr group. This group then returned after 24 hours (23.15 ± 2.12 hours, mean ± 1SD) and completed another four blocks of trials. Again, all participants reached through an error-clamp to the train target, followed by an error-clamp to the probe target at the start of blocks 3, 5 and 7.

30min and 6hr groups: The results of the 3min group suggested that increasing the number of trials on Day 1 was not sufficient to allow for the reduction of generalization. To dissociate between effects of time vs. leaving the experimental set up and time vs. overnight sleep, we recruited two new groups of subjects, a 30 minute group (n=11, 24.45 ± 3.56 years old, mean±1SD, including 4 men and 7 women) and a 6 hour group (n=10, 26.80 ± 4.05 years old, mean±1SD, including 4 men and 6 women). Following the initial block of null, and then two blocks of field training (exactly as in the 24hr and 3min groups), subjects in the 30min group left the experimental setup and sat quietly in the lab for 30 minutes before returning to the robotic arm and completing another two blocks (30min group, Figure 1B). In contrast, subjects in the 6hr group left the lab for at least six hours (actual wait time was 6 hours, 15.0±11.5 minutes, mean ± SD) and then returned to complete an additional two blocks of trials (6hr group, Figure 1B). Participants were allowed to perform their normal activities during this six hour period but were instructed not to sleep, play video games, or exercise. Both groups of subjects returned 24 hours later (23.95 ± 2.46 hours for 30 min group, 26.12 ± 2.35 hours for 6 hour group, mean ± 1SD) and performed another four blocks of trials. Consistent with the other participants, at the start of blocks 3, 5 and 7 all participants in these groups reached through an error-clamp to the train target, followed by an error-clamp to the probe target.

4.2.2 Data Analysis

To quantify the motor commands that subjects learned to generate, a force index was calculated from the forces $f(t)$ that they produced against channel walls during an error-clamp trial. This index, labeled with variable a, reflected a measure of compensation as a function of ideal forces:

$$f(t) = a\left(13\dot{x}(t)\right) \tag{1}$$

In Eq. (1), $\dot{x}(t)$ is hand velocity parallel to the direction of target, and 13 reflects the size of the velocity-dependent, curl force field that we applied to the hand. We found a least-squares estimate of the variable a for each error-clamp trial by using the measured forces $f(t)$ and velocities $\dot{x}(t)$. To do this, the reach trajectory for a given trial was divided in two parts at the point of maximum extent (reach out and reach back). For the train target, we computed the force index a_t for the reach component from start point to the target. For the probe target, we computed the force index a_p for the reach component from target back to the start point. A probe to train ratio a_p/a_t was calculated for each participant by dividing the force index of the probe target by the force index to the train target. This was calculated as a proxy for the percentage of learning that was generalized, as complete compensation for the field is rarely achieved (i.e., $a_t < 1$).

To measure an individual's reach consistency, pairs of reach trajectories were compared using the time series of velocity vectors, resulting in a correlation coefficient (Shadmehr and Mussa-Ivaldi, 1994b). To measure changes in correlations with training, all field trials across all sessions were separated into mini-sets of seven consecutive trial bins. The cross correlation of each pair within this 7 trial bin was computed, resulting in 21 correlations. The average of the correlations was then reported as a measure of consistency for that bin of trials for that subject. Reach reaction time (RT) was measured by finding the time point at which the hand's tangential velocity exceeded 0.02 m/s.

All data analysis was completed in MATLAB 7.0.4 and all statistical analysis was performed with SPSS Statistics 22. Individual reach trajectories were excluded from analysis if the participant did not reach the target (within 8mm of target center). Additionally, reach durations of 2.4 seconds or more, and reaches with force production greater than 150% of necessary force compensation were removed from analysis. Two participants were removed from the data pool and are not included in the overall subject number reported (n=41) as they had an average a_p / a_t that was outside the 3 standard deviation window of the population median.

4.3 Results

Volunteers (n=41) made out-and-back reaching movements to a train target and a probe target (Figure 1A). After baseline training in a null field, all participants completed two blocks of trials in a force field (192 trials per block, Fig. 1B).

4.3.1 Learning generalized to the probe target

For the train target, the field produced forces that were perpendicular to reach direction, pushing the hand away from the goal on the outward reach, but not on the inward reach. In response, subjects learned to produce forces perpendicular to the direction of reach on the upward segment of their motion to the train target, as illustrated in Fig. 2A. For the probe target, there were never any perturbations; rather, reaches were always in an error-clamp. Nevertheless, as subjects learned to produce forces for the train target, they also produced forces for the probe target. In particular, subjects produced significant forces during the upward segment of their reach in the probe trials (the upward segment is the positive velocity component of the probe target, Fig. 2A).

We used a force index (Eq. 1) to quantify these forces. For the train target, the index a_t was computed as the hand reached from start to the turn-around point (an index of 1 implies exact compensation for the velocity dependent field). This was our proxy for the task-relevant forces that subjects produced. For the probe target, the index a_p was computed as the hand reached from the turn-around point back to start (for the probe target, an index of 1 implies that the subjects produced forces that were as large as that for the train target). This was our proxy for the task-irrelevant forces that subject produced. We found that a_t increased with practice, reaching a value of 0.89 ± 0.07 by the end of block 2 (mean ± 1SD across last 5 error-clamp trials, top plot, Fig. 2B). Similarly, a_p increased with practice, reaching an average of 0.33 ± 0.16 (mean ± 1SD across last 5 error-clamp trials). To compare the rates of increase in the forces for the train and probe targets, we computed a_p/a_t , as shown in the bottom sub-plot of Fig. 2B. This is a measure of percentage of learning that was generalized to the probe target. We found that a_p/a_t rapidly declined. Of the total change from the beginning to the end of Day 1 training, 96.29% of the total drop in a_p/a_t occurred in the first 50 trials, remaining relatively constant and decreasing only an additional 3.71% for the subsequent 330 trials. Therefore, participants generalized their experience from the train target to the probe target.

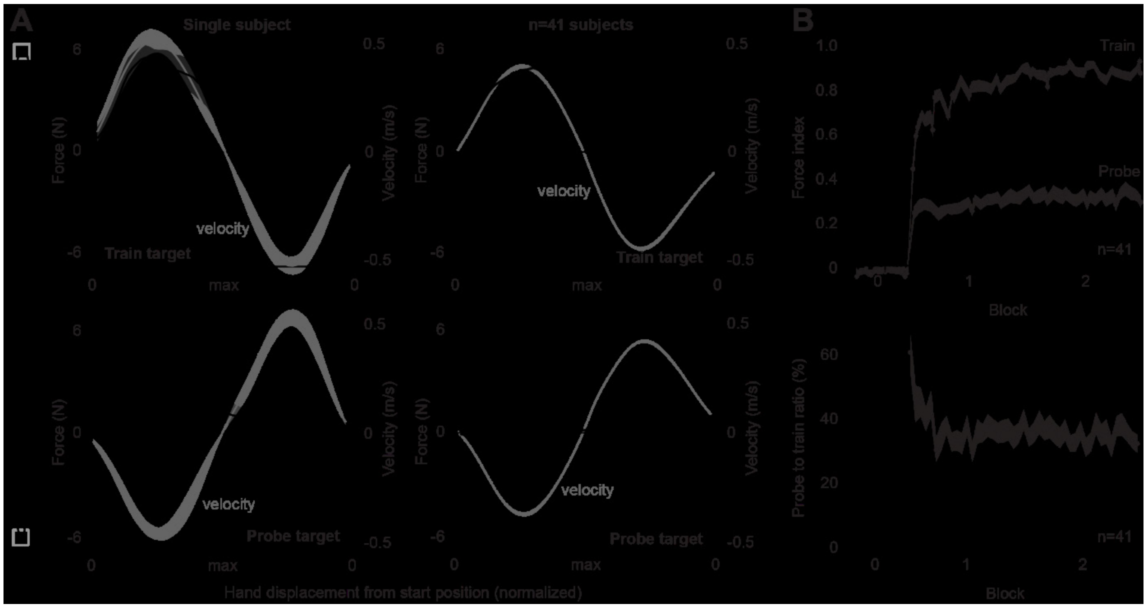


Figure 4.2 Generalization to the probe target. **A)** The traces show the forces and velocities recorded from the last 20 error-clamp trials in block 2 on Day 1 for a representative subject (left column), all subjects (right column). The x-axis, representing hand displacement, was normalized to the maximum extent of each reach. Error bars are SD for single subject, SEM across subjects. The top figure is motor output (forces against channel walls) and velocity for an out-and-back movement to the train target. The forces are the change from forces that the subjects produced in block 0 (null field). Positive velocity indicates a reach to the target. Negative velocity indicates reach back to center **B)** A force index was used to quantify the amount of force produced for reaches to the train target (upward segment, i.e., positive velocities), and for reaches to the probe target (also positive velocities). The index was the regression of the force profile onto the ideal profile resulting from hand velocity. For the train target, the index was calculated for the reach out to maximum extent. For the probe target, the index was calculated for the reach from maximum extent back to the start target. Probe to train ratio is the ratio of the force index for the probe target to the train target, and stands as a proxy for percent generalization. Data are mean±SEM, plotted for the second half of block 0, and all of blocks 1 and 2

However, because all movements to the probe target were in error-clamp, production of these forces was unnecessary: regardless of the forces produced against channel walls for the reach to the probe target, participants experienced no kinematic performance errors, leaving no explicit cues or errors to indicate that this component of behavior, i.e., this generalization, was resulting in task-irrelevant motor commands. As the data in Fig. 2B illustrates, the forces that subjects produced in the probe trials did not show any evidence of reduction, both in terms of absolute value (top part of Fig. 2B), and in terms of relative value (bottom part of Fig. 2B), beyond the first 50 trials of training. Could this inefficient behavior become reduced?

4.3.2 Time away from practice was necessary for the reduction of task-irrelevant forces

After completion of block 2, we separated our subjects into four groups that waited for varying amounts of time until the start of block 3 (Fig. 1B). One group (n=10) waited 24 hours. By the end of training on the first day (block 2), these subjects were able to counteract the field, producing a force index of a_t=0.90±0.02 (mean±SEM, over last 5 error-clamp trials) (Fig. 3A). The participants generalized this learning to the probe target, a_p=0.36±0.04 (mean±SEM, over last 5 error-clamp trials). When the subjects returned at 24hrs (block 3), on their first trial they reached in an error-clamp to the train target. This allowed us to measure recall of the motor memory. We found that on this first trial of Day 2 a_t = 0.33±0.05 (mean±SEM), which represented 37.13% of the value achieved at the end of training on Day 1. Within a few additional trials, after experiencing errors in the field, performance was statistically indistinguishable from performance 24 hours earlier. Indeed, a_t on Day 2 measured over block 3 was not significantly different from a_t at the end of Day 1 (post hoc pair-wise comparison block 2 vs. block 3, p = 0.207, following repeated measures ANOVA with a within-subject factor of mean train force index for blocks 1 to 8, effect of block, $F_{(7,63)}=15.122$, $p < 0.001$). However, the forces to the probe target as measured via a_p showed a significant decrease from Day 1 to Day 2 (post hoc pair-wise comparison block 2 vs. block 3, p = 0.002, following repeated measures ANOVA within-subject factor of mean probe force index for blocks 1 to 8, effect of block, $F_{(7,63)}=4.555$, $p<0.001$). Therefore, on Day 2 the subjects recalled the forces for the train target, but now produced less force for the probe target.

Because we had observed a reduction in the task-irrelevant forces following a 24 hour break, we asked the subjects to return on Days 3 and 4 for further training, wondering whether this trend of reduced forces for the probe target would continue. During these sessions, participants again exhibited recall on the very first trial of each test day (a_t =0.43±0.07 for Day 3, and a_t = 0.57±0.06 for Day 4). There was no significant benefit of time for performance to the train target: a_t as measured over the entire training blocks of Days 3 and 4 were not different from a_t at the end of Day 2 (all post hoc pair-wise comparisons block 4 vs. block 5-8, $p>0.100$). However, the ratio a_p/a_t declined significantly as a function of day of training (effect of day, p = 0.004, $F_{(3,24)}=5.855$, comparison of last 5 values per day). Indeed, we found a within subject 12.8±3.4%

decrease in generalization a_p/a_t from the end of Day 1 to the end of Day 2, a 2.8±3.0% decrease from the end of Day 2 to the end of Day 3, and a 5.4±3.0% decrease from the end of Day 3 to the end of Day 4.

We wondered whether the reduction in the generalization at the first 24 hour break was due to passage of time, which included a normal night of sleep, or simply increased practice. In the 3min group (Fig. 1B), participants (n=10) performed blocks 3 and 4 after a short break following completion of block 2, doubling the amount of training on Day 1. The results of the 3min group are displayed in Figure 3B. In blocks 2 and 3 the train forces for the 3min and 24hr groups were comparable: ANOVA with a within-subject repeated measure of a_t (blocks 2 and 3), and between-subject factor of group, revealed no significant interaction, $F_{(1,18)}=0.41$, p=0.529. However, unlike the 24hr group, the 3min group did not exhibit a decrease in a_p following block 2 (post-hoc comparison of 24hr vs. 3min group, p = 0.023, following a one-way ANOVA for difference in a_p between block 2 and block 3 across all groups, $F_{(3,40)}=4.074$, p = 0.013). Instead, in the 3min group the probe forces in block 3 were not significantly different than in block 2 (post-hoc comparison block 2 vs. block 3, p >0.500 following repeated measures ANOVA within-subject factor of mean a_p in blocks 1 to 8, effect of block, $F_{(7,63)}=8.471$, p=0.003). That is, increased practice was not sufficient to allow for the reduction of the task-irrelevant forces in the probe trials.

After completion of block 4, the 3min group received a 24hr break. Upon return, they exhibited recall of the forces for the train target (Fig. 3B). Remarkably, only after this

24hr passage of time did the 3min group show a reduction in the forces for the probe target (post-hoc pair-wise comparison block 4 vs. block 5, p =0.003). These participants did not experience any further 24 hour set breaks. They did however, practice twice as long on Day 2 as the participants in the 24hr group, performing four blocks of trials. We found that forces produced for the probe target continued to decrease with practice in these sessions. Here, we saw that forces due to generalization continued to slightly decrease with practice in these sessions. However, this drop was only statistically significant when comparing the very beginning of the second test session, to the end of practice (post-hoc comparison block 5 vs block 8, p = 0.041).

In summary, we observed that at 24 hours following initial practice but not 3 minutes, subjects reduced the forces that they produced for the probe target. This suggested that following initial acquisition, time away from practice afforded an advantage that practice alone could not provide: the ability to reduce the task-irrelevant generalization.

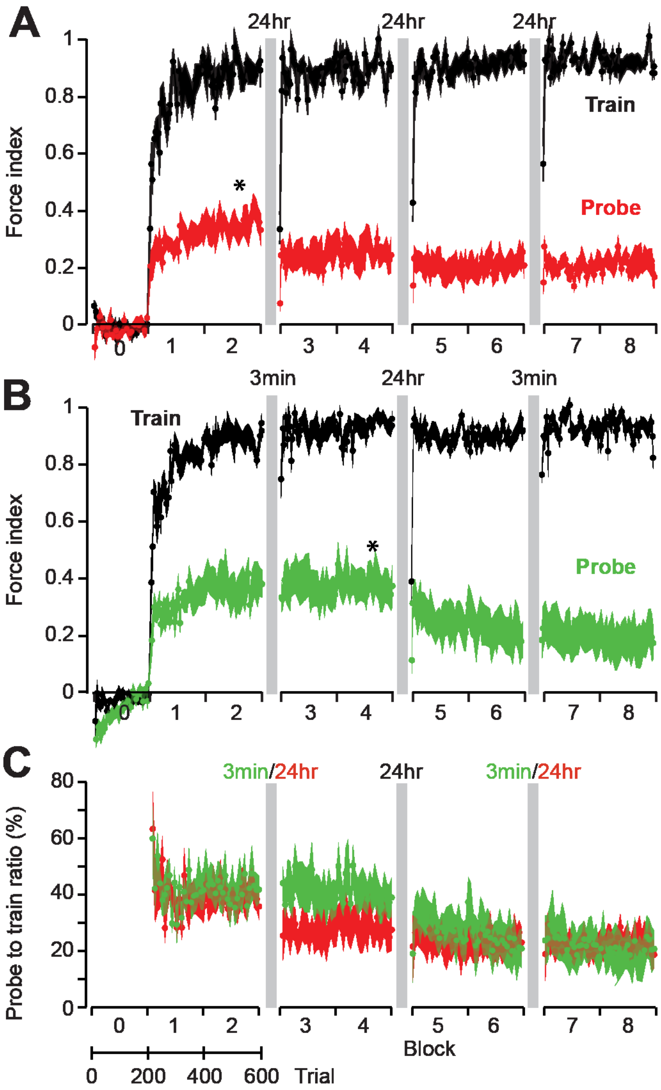


Figure 4.3 Forces produced to probe target reduced with time, not practice. **A)** 24hr group. Training produced generalization. However, when subjects returned at 24hrs, they exhibited reduced forces for the probe target. Asterisk denotes the block at which point all subsequent blocks were statistically significant for probe force index according to post-hoc analysis with a 0.05 significance level. **B)** 3min group. While a 3min break was not sufficient to reduce probe forces, after an overnight passage of time the probe forces declined. Asterisk denotes the block at which point all subsequent blocks were statistically significant for probe force index according to post-hoc analysis with a 0.05 significance level. **C)** Ratio of probe force index to train force index.

4.3.3 Sleep was not necessary for reduction of the task irrelevant forces

The 24hr group experienced time away from the experimental set up and a normal night of sleep. We wondered which factor was critical. Therefore, we recruited two new groups: one group (n=11) waited 30 minutes between blocks 2 and 3, and a second group (n=10) waited 6 hours (30min and 6hr groups in Fig. 1B). In the 30min group, subjects left the experiment room after completion of block 2 and waited in another area in the lab for exactly 30 minutes. Once they returned to the experiment room, we observed that generalization had not been reduced significantly (Fig. 4A, post hoc pair-wise comparison, block 2 vs. block 3, p=0.294, following repeated measures ANOVA with a within-subject factor of mean a_p for blocks 1-8, effect of block, $F_{(7,70)}=9.688$, $p<0.001$). Therefore, 30 minutes away from practice resulted in no evidence of reduction in probe forces. After completion of block 4, the 30min group returned 24 hours later. We now observed a reduction in a_p (post hoc pair-wise comparison block 4 vs. block 5, p=0.001).

In the 6hr group, subjects left the experiment room after completion of block 2 and returned after six hours of wakefulness. Upon their return, we observed a significant drop in a_p (Fig. 4B, post hoc pair-wise comparison block 2 vs. block 3, p=0.028, following repeated measures ANOVA with a within-subject factor of mean a_p in blocks 1-8, effect of block, $F_{(7,63)}=5.083$, $p < 0.001$). We observed no further decrease in the forces to probe target after a 24 hour break between blocks 4 and 5 (post hoc pair-wise comparison block 4 vs. block 5, p=0.827). In fact, the drop following 6 hours was indistinguishable from the group who waited 24 hours after block 2 (post-hoc comparison of 24hr vs. 6hr group, $p = 0.761$, following a one-way ANOVA for difference in probe force index between block 2 and block 3 across all groups, $F_{(3,40)}=4.074$, $p = 0.013$). Therefore, the data from the 6hr group revealed that sleep was not necessary to achieve a reduction in generalization. Rather, 6 hours of time away from practice was sufficient.

On Day 2, both the 30min and the 6hr groups returned to the lab and were re-tested (block 5, Fig. 4 A&B). For the 30min group, on Day 1 there had been no reduction in the generalization forces (as in the 3min group), but now after an overnight period away from practice, they exhibited a reduction (statistics reported above). In contrast, for the 6hr group there had been a reduction in the forces to the probe target in Day 1. Now after this overnight passage of time, they exhibited no further reductions. This observation reiterates the finding that passage of time, and not sleep, was the critical factor that resulted in the ability to detect inefficient behavior and reduce the probe forces. As in the 3min group, the 30min and 6hr participants reached for four blocks on Day 2. Unlike the 3min group, these participants did not show any statistically significant change across these additional blocks ($p>0.100$ for all post-hoc comparisons blocks 5-8 within each group).

Figure 4C summarizes some of these results. We have plotted the change in probe forces as a function of time away from practice (from block 2 to 3). ANOVA with a between subject factor of time yielded a significant effect of group (change in generalization, $F_{(3,40)}=4.074$, $p=0.013$). We found that neither increased practice nor a short amount of rest was sufficient to achieve the reduction (post-hoc comparisons 24hr vs. 3min, p = 0.0.023 and 24hr vs. 30min, p =0.048). Rather, 6 or 24 hours were sufficient (post-hoc comparisons 24hr vs. 6hr, p = 0.761).

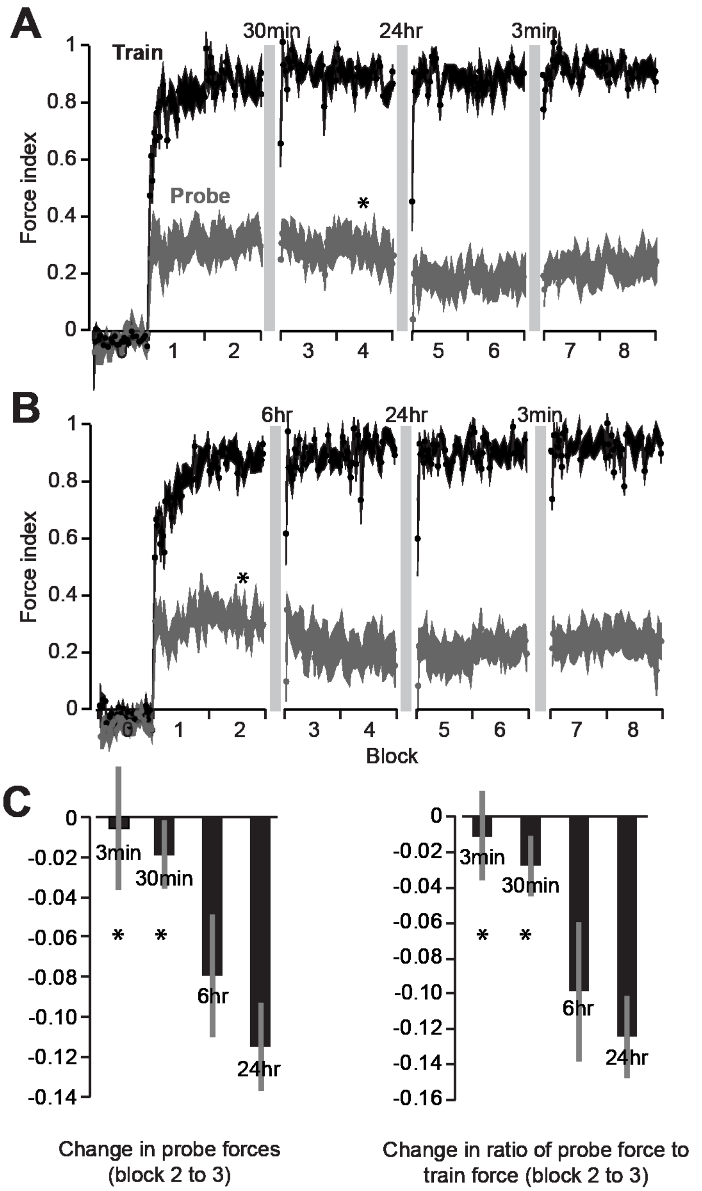


Figure 4.4 Sleep is not necessary for the reduction of probe forces. A) 30min group. While 30 minutes was not sufficient to reduce the probe forces, after an overnight passage of time these forces declined. Asterisk denotes the block at which point all subsequent blocks were statistically significant for probe force index according to post-hoc analysis with a 0.05 significance level. **B)** 6hr group. After 6 hours of time away from practice, probe forces were reduced. **C)** Summary of data as a function of time. Change in the probe forces was measured by computing the difference between the probe forces in blocks 2 and 3 for each subject. The same procedure was followed for computing the change in ratio of probe to train forces. Data are mean±SEM. Asterisk denotes significant difference from the changes observed in the 24hr group at the 0.05 significance level.

4.3.4 Confounding variables

Two confounding variables are time of initial acquisition and time of the follow up test sessions, which may affect properties of motor memory through circadian rhythms. We approached this problem by first considering whether time of acquisition (initial training) affected the magnitude of the forces that subjects produced for the train and probe targets. We found no evidence that time of acquisition served as a significant factor in predicting the forces at the end of training in block 2 (a_t, t = -0.721, p = 0.474; a_p, t = 1.478, p = 0.146). Next, we considered time of testing (recall) and asked whether the reduction in the forces produced for the probe target was affected by this variable. We found that time of testing did not serve as a significant factor in predicting the reduction in a_p (t = -1.141, p = 0.261). Indeed, only the time away from initial practice proved to be a significant predictor of the reduction in a_p (t = -2.435, p = 0.020). Therefore, time of acquisition and time of testing appeared to have little or no effects on the amount of forces that the subjects produced toward either target. The relevant factor was time away from practice.

4.3.5 Practice, but not passage of time, improved the task-relevant forces

Finally, we focused on performance for the train target and asked whether time away from practice produced any advantages in performance, over and beyond benefits from practice alone. The data for various measures of performance to the train target are shown in Figure 5. Using the average force index a_t of each training block, we performed a repeated measures ANOVA with a between subject factor of time and found improvements across practice (effect of block, $F_{(7,259)}$=24.616, p<0.001), but no significant effect of group (group*block interaction $F_{(21,259)}$=1.310, p=0.168). This implied that increased practice alone was sufficient to improve task-relevant forces. We performed the same analysis using maximum displacement on the reach out to the train target as the variable of interest, and found similar results. Again, we saw a significant difference across blocks of practice (effect of block $F_{(7,259)}$ = 62.319, p <0.001), but no interaction between group and block ($F_{(21,259)}$=1.384, p = 0.126). When we considered consistency of movements (a measure of correlation between neighboring trials), the results continued this trend: consistency improved with practice (effect of block $F_{(7,259)}$=83.644, p < 0.001), but no additional benefits were found due to passage of time (group*block interaction, $F_{(21,259)}$=0.897, 0.618). Finally, we considered reaction time and found a monotonic decrease with practice for the training target (effect of block $F_{(7,259)}$=7.622, p < 0.001). However, there were no additional benefits to the changes in reaction time due to passage of time (group*block interaction, $F_{(21,259)}$=0.763, p=0.763).

A decrease in reaction time was also observed for reaches to the probe target across training, indicating that the improvement in this performance measure was not direction specific (effect of block, $F_{(7,259)}$=14.568, p<0.001). As we had observed for the train target, there were no additional benefits due to passage of time for the reaction time of reaches to the probe target (group*block interaction $F_{(21,259)}$=0.925, p = 0.559). Noticeably, the reaction time for the probe target was slightly longer than for reaches to the train target. This is presumably due to the infrequent nature of the probe trials.

In summary, practice resulted in an increase of task-relevant forces, a reduction in hand displacement, a reduction of movement variability, and a reduction of reaction times, all of which improved performance. However, the groups that experienced passage of time in between the blocks of practice did not show any benefit in these variables over and above what practice alone provided. Together, it appears that practice (and not time) enhanced the task-relevant component of motor memory, whereas time (and not practice) reduced the task-irrelevant component.

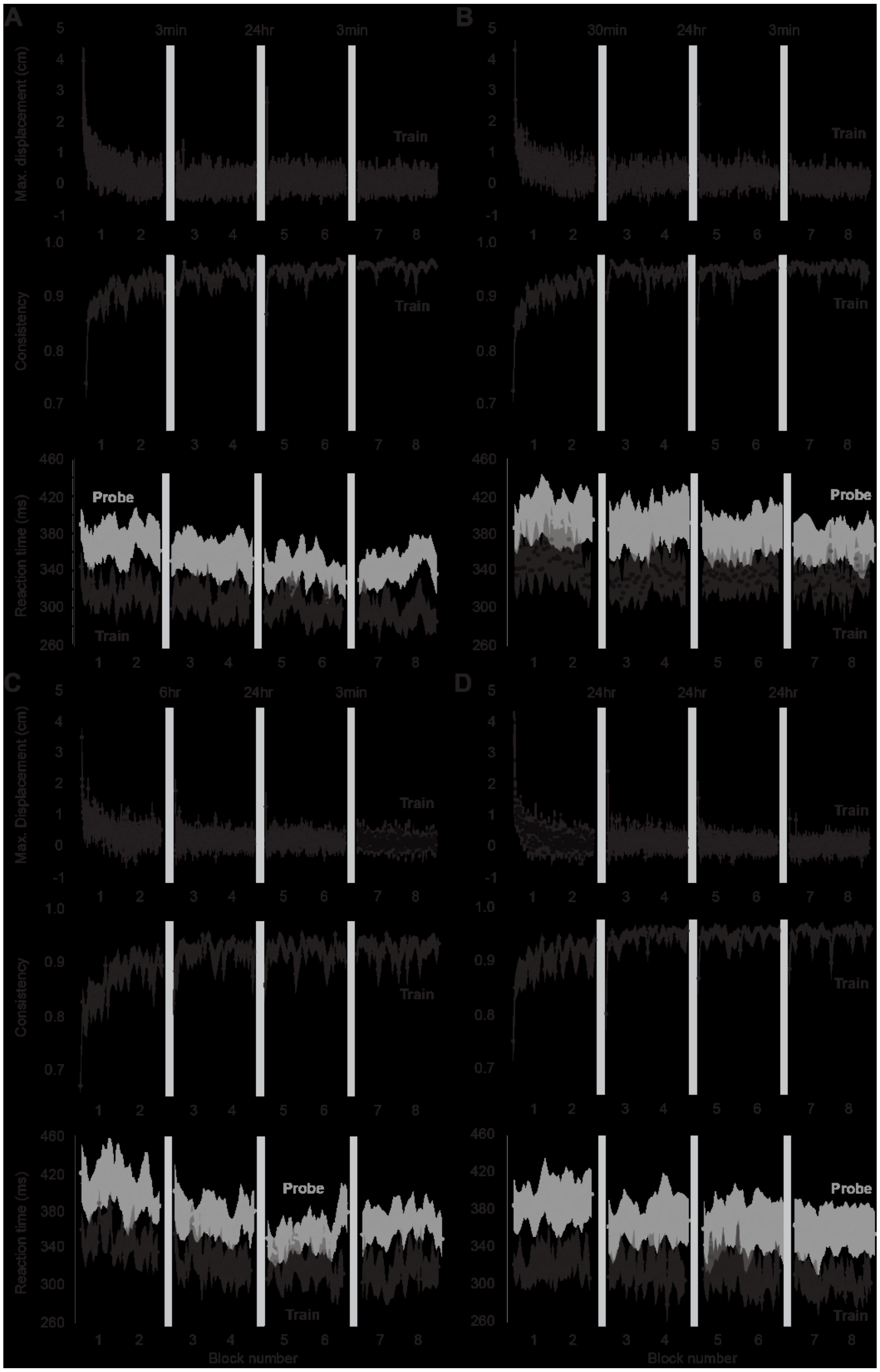


Figure 4.5 Changes in speed and accuracy of movements. **A)** Data for the 3 minute group. The top plot shows maximum displacement of the hand during the reach out to the train target. Middle plot displays consistency of the reaching movements for the train target. Consistency refers to average correlation between pairs of reach trajectories in bins of 7 trials. The bottom plot displays reaction time (RT), i.e., the time from appearance of the target to start of the reach for the train and probe targets. For the train target, bins of seven trials were used to calculate the reaction time, and bins of five trials were used for the probe target. For all measures, the data are across subject mean+/-SEM. Each block is 192 trials, divided into 24

probe targets and 168 train targets. **B)** Results for 30 minute group. **C)** Results for 6 hour group. **D)** Results for the 24 hour group.

4.3.6 Task irrelevant forces were minimized if they had kinematic consequences

The task-irrelevant forces that our participants produced in their reaching movements toward the probe target did not have kinematic consequences, as these movements were always in error-clamp. In order to reduce these forces, time away from practice was necessary. What happens if the production of task-irrelevant forces coincides with kinematic consequences?

Examples of such task-irrelevant forces are those that subjects produced during the return phase of their reach to the train target (Fig. 2A). Using the same force index, we quantified these forces during error-clamp trials and have plotted the results in Fig. 6. We found that participants produced very little force throughout the entire experiment. Indeed, forces on the reach back were not significantly different from zero by the end of training in block 3 for all groups ($p>0.100$ for all t-test).

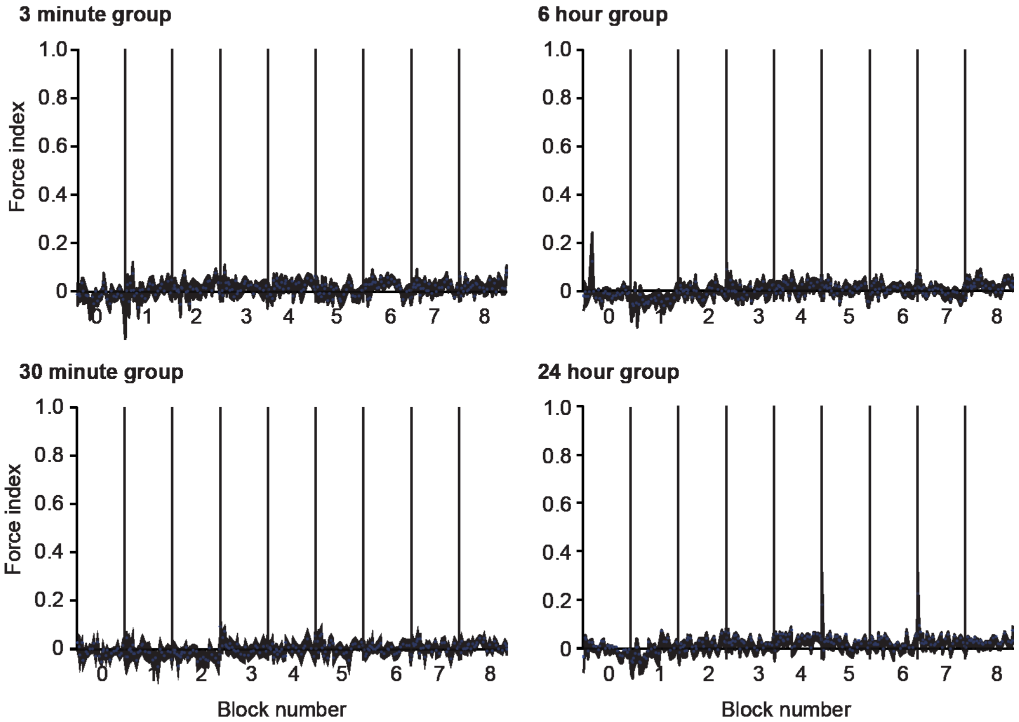


Figure 4.6 Downward segment of reach. The forces that were produced during the downward segment of the reach for the train target, as quantified via a force index (Eq. 1). The forces were generally near zero and did not differ between groups. Error bars represent between subject SEM.

4.4 Discussion

We designed a task in which with practice, subjects learned to produce task-relevant motor commands, countering a perturbation as they reached to a target. This practice also resulted in generalization, which we measured during reaches to a probe target. Though generalization is often considered a hallmark of learning, in our task these motor commands were task-irrelevant as all reaches to the probe target were in error-clamp, making it so that any forces that were generalized to the probe target were unnecessary. Importantly, production of the task-irrelevant forces did not result in kinematic errors, making it so that the only indication of this inefficiency was an intrinsic measure of effort.

We found that practice alone was not sufficient to reduce the task-irrelevant forces: the more participants practiced reaching to the train target, learning to produce task-relevant forces, the more they produced task-irrelevant forces as they reached to the probe target. However, at 24 hours after end of initial practice, but not 3 or 30 minutes, subjects were able to both maintain production of the task-relevant forces for the train target, as well as reduce the task-irrelevant forces for the probe target. Sleep was not necessary to achieve this feat, as the reductions were also observed following 6 hours of time away from practice. Therefore, whereas practice appeared critical in allowing for the improvement in the task-relevant motor commands, it was time away from practice that appeared critical in allowing for the reduction in the task-irrelevant commands.

It seems likely that passage of time away from practice alters motor memory, affording certain measures of stability to its representation: a six hour time window corresponds with some previous studies of force-field learning which found it to be a critical amount of time necessary to reduce interference from competing tasks (Shadmehr and Brashers-Krug, 1997;Overduin et al., 2006) (but see (Caithness et al., 2004)), and the time-span over which the neuronal basis of the memory was significantly altered (Shadmehr and Holcomb, 1997). Similarly, in sequence learning tasks it was found that 6-8 hours of time provided protection from interference from a second sequence (Walker et al., 2003;Korman et al., 2007). Here we found an additional dimension to the benefits of passage of time: time away from practice made it possible to represent the memory more efficiently through reduction of the component that was energetically wasteful.

From a computational perspective, generalization is thought to be a behavioral assay of the width of the receptive fields of the neurons that participate in the process of learning (Shadmehr, 2004;Hwang and Shadmehr, 2005). One way to view reduction of generalization is via a reduction in the width of these receptive fields, which may occur through weakening of learning-induced potentiated synapses. That is, it is possible that initial practice in the task produces wide spread changes in synaptic strength, resulting in wide generalization, and then with passage of time away from practice, some of the changed synapses are returned to near baseline, especially if the generalization produced task-irrelevant behavior. In this hypothesis, the motor memory is made more efficient through a return of the learning-induced synaptic changes toward baseline for those components of the memory that were associated with task-irrelevant motor commands.

What may be the mechanism with which this is achieved? Recent *in vivo* imaging techniques have allowed for longitudinal tracking of the formation and elimination of specific dendritic spines in mice. Two studies have found that training of a new motor skill in mice leads to increases in the rate of dendritic spine formation (Xu et al., 2009;Yang et al., 2009), with new spines forming within an hour of initial training (Xu et al., 2009). However, with passage of time and continued practice in additional training sessions, the rate of elimination of spines also increases, bringing the net synaptic changes near baseline. Indeed, it has been hypothesized that a critical function of sleep may be to prune memories that are acquired during wakeful hours (Tononi and Cirelli, 2006), an idea that is supported by data in juvenile mice (Maret et al., 2011) and fruit flies (Bushey et al., 2011).

Here, we found that instead of sleep, 6 hours of time was sufficient to reduce a large part of the task-irrelevant motor commands, while maintaining the task-relevant portion. This independence from sleep appears consistent with other results in motor adaptation tasks. In a number of adaptation tasks, it has been shown that sleep does not add significantly to the benefits of time away from practice (Debas et al., 2010;Donchin et al., 2002;Doyon et al., 2009), and is not necessary for memory consolidation (Shadmehr and Brashers-Krug, 1997;Overduin et al., 2006). Possibly the previously reported wide-scale synaptic changes are responsible for the improved motor efficiency observed here, but unlike other tasks, motor adaptation may not require sleep in order for this process of synaptic pruning to be initiated.

We interpret our observation of decreased generalization over time as an improvement in the efficiency of the participants' movements, as these forces were an energetically costly byproduct of learning. Though these forces are small in magnitude, they are not completely negligible and incur a metabolic cost (Huang et al., 2012). In our task, there were no kinematic errors or differences in performance feedback to alert the brain that these forces were unnecessary. We posit that the energetic cost of generalization served as a cue to alter these motor commands.

A concern in our study may be regarding the tool that we relied on to measure the task-irrelevant motor commands: error-clamps. A potential limitation is that error-clamps are not innocuous probes of motor memory (Vaswani and Shadmehr, 2013). When they are given as a continuous block of trials, the changes that take place in motor commands appear to be partly due to an ability of the nervous system to detect a contextual switch from train trials (in which errors take place) to error-clamp trials (in which errors are eliminated). Here, the reaches to the probe target were always in error-clamp, eliminating the potential problems associated with changes in context.

Our results suggest that only the first break of 6 or 24 hours following initial learning contributed significantly to reduced generalization. A subset of our participants who performed the 24hr group paradigm experienced two additional set breaks of 24 hours after the first crucial break. Despite a trend towards reduction, there was no statistically significant decrease in generalization across these additional rest periods. The remaining participants did not experience any further set breaks of critical duration, but instead experienced additional practice; blocks 5-8 were all performed in the same session. Though the participants in the 3min group did display a slight decrease across this long practice session, the subjects in the 30min and 6hr group did not show a change in forces due to generalization. Therefore, we did not find strong evidence that additional set breaks of critical duration or additional practice after the first set break could further reduce the inefficient motor behavior. Though neither of these interventions appeared to improve performance, future experiments could benefit from systematically controlling these two variables, and investigating the possible benefits of combining practice and additional long break durations.

With future work, we hope to understand whether generalization that results in task-irrelevant force production could be completely eliminated. Though here we observed benefits from passage of time, at the end of the 4-day experiment subjects still produced significant task-irrelevant forces. Perhaps additional passage of time, for example 48 or 72 hours, could provide a larger drop in generalization. However, we believe that passage of time may not be the only critical factor, but that the coupling of the train targets and probe targets may play a role. Reaching to the probe target provides the subject with the necessary experience to identify that this effort is unnecessary, but additional practice to the train target may act to counter this behavior. Therefore, altering the probability of the probe target with respect to the train target may affect the patterns of generalization. Finally, we note that the largest drops in generalization occurred after large durations of rest, which led to more forgetting and a larger initial error in subsequent training sessions. As training progressed, there was less forgetting across rest breaks, and generalization was not decreased further. It is possible that greater re-learning at the start of each block, which could be achieved by introducing null blocks or catch trials into our paradigm, could also serve to reduce the task-irrelevant generalization.

In summary, we studied a motor task in which learning produced a memory that included both task-relevant and task-irrelevant motor commands. With practice, the task relevant motor commands improved, and this improvement did not benefit from time away from practice. However, with time away from practice the production of the task-irrelevant motor commands was reduced, resulting in a more efficient control of movements. The crucial factor necessary for this reduction was time away from practice, as additional training was not sufficient to improve efficiency. Therefore, our work illustrates a previously unknown property of motor memory: practice (and not time) improves the task-relevant component of motor memory, whereas time (and not practice) makes the task-irrelevant component eligible for reduction, a process that makes the motor memory more efficient.

Chapter 5: Reward dependent modulation of movement variability[3]

5.1 Introduction

Movement variability is often considered an unwanted and unavoidable byproduct of noise in the nervous system. However, behavioral evidence suggests that variability serves a critical role in motor learning (Wu et al., 2014;Tumer and Brainard, 2007). Variability may benefit learning because carefully controlled fluctuations in motor output can serve as a form of exploration, allowing the animal to find a better solution for achieving a goal. Indeed variability fluctuates in order to manage the pressures of achieving success and attaining reward. For example, humans increase their movement variability during periods of low success or minimal feedback, which is thought to reflect an active search for a rewarding outcome (Galea et al., 2013;Izawa and Shadmehr, 2011b). Similarly, monkeys exhibit increased variability of eye velocity, latency, and amplitude for saccades that are made to targets that are not paired with reward, as compared to targets that are paired with reward (Takikawa et al., 2002). When variability in a lever pressing task is rewarded instead of repetition, pigeons can produce highly variable lever sequences similar to those produced by a random number generator (Page and Neuringer, 1985).

---

[3] The work in this chapter has previously published in the Journal of Neuroscience: Pekny SE, Izawa J and Shadmehr R (2015) Reward dependent modulation of movement variability J Neurosci 35(9):4015-24

Deciding whether to repeat a movement or vary one's actions depends on the ability to predict future occurrences of punishment or reward. The difference between the actual and expected outcome is known as a reward prediction error, which relies on dopamine dependent processes (Schultz et al., 1997). It is therefore not surprising that variability, especially in terms of goal directed exploration, has been linked to dopamine and the basal ganglia. In songbirds the source of variability in song production is believed to be in brain structures homologous to the mammalian basal ganglia (Olveczky et al., 2005;Kao et al., 2005). Activating striatal D1 and D2 receptors in mice alters the decision process to stay with or switch from the current behavior in order to obtain reward (Tai et al., 2012). During periods of low variability, administration of a D2 agonist increases variability in rats (Pesek-Cotton et al., 2011). In humans, a D2 antagonist abolishes the increase in variability observed during periods of low reward (Galea et al., 2013).

Given this potential link between control of movement variability and the basal ganglia, we hypothesized that patients with basal ganglia dysfunction would have difficulty controlling their motor variability in response to reward prediction errors. Indeed patients diagnosed with Parkinson's disease (PD) are known to have difficulties in certain cognitive learning tasks which depend on trial and error feedback (Knowlton et al., 1996), with some evidence suggesting a specific learning deficit based on negative reward prediction errors (Frank et al., 2004;Frank et al., 2007;Bodi et al., 2009). Here, we considered a reaching task and provided subjects with binary feedback about the success of the reach. We manipulated the probability of reward and quantified the resulting changes in variability in healthy and PD populations.

5.2 Methods

A total of n=26 volunteers participated in our study. Among them were n=9 mildly affected volunteers diagnosed with Parkinson's disease (63±6.9 years old, including 5 females and 4 males), n=8 healthy age-matched controls (65±8.1 years old, including 4 females and 4 males), and n=9 healthy young controls (25±5.6 years old, mean ± SD, including 7 females and 2 males). All participants provided consent by signing a form approved by the Johns Hopkins University School of Medicine Institutional Review Board.

PD patients

All Parkinson's disease (PD) patients were free of dementia as assessed by a Mini-mental Status Examination (Folstein et al., 1975) on which all subjects scored better than 28. Clinical severity was measured by using the Unified Parkinson's Disease Rating Scale (UPDRS) (Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease, 2003), the results of which are provided in Table 1. All subjects were free of musculoskeletal disease and had no neurological disease other than PD, as confirmed by a neurologist. All subjects were taking dopamine agonist medications at the time of testing.

5.2.1 Behavioral Task

The experimental task was similar to those described in a previous experiment (Izawa and Shadmehr, 2011b). Subjects made reaching movements in the horizontal plane holding onto the handle of a two-joint robotic manipulandum (Figure 1A). An opaque screen was placed above the subject's arm, upon which a video projector painted the scene. At the start of each trial, the robot positioned the participant's hand in a start box. Once the hand was at the start box, a target of 6° width located 10 cm from start was displayed. Participants were instructed to make quick, shooting movements so that the robotic handle passed through the target. Success was indicated via an animated target explosion when the participant's hand passed through an experimentally controlled rewarding target region. Movements were also required to have a reaction time of less than 0.6s and movement duration of less than 1s to be successful. Following a successful reach, a point was added to the participant's score, which was displayed throughout the experiment. This target explosion and point were the reward given in our task. Participants were compensated for their time, and the total payment was not based on task score.

All participants first performed a familiarization block of 50 trials in which full visual feedback of the movement was provided via a projected cursor (5x5mm) representing hand position. These movements were performed in a 0ROT condition, in which participants were rewarded if they passed within a region of ±4° centered at 0°, the target center (rewarding target region is highlighted in gray in Figures 1A&1B). A clockwise rotation is defined as positive. Following this initial block of training, cursor feedback was shut off and participants did not receive visual information about the handle position for the remainder of the experiment. The only performance feedback that participants continued to receive was regarding success or failure of the trial. After the visual cursor feedback was removed, participants performed another block of 50 trials in the 0ROT condition. Participants next experienced a block of 100 trials in which unbeknownst to the subject, the rewarding target region was shifted and centered at +4° from midline, referred to as the +4ROT condition (reaches were now rewarded if they fell between 0° and +8°, as illustrated in Figure 1A&1B). This block of trials was followed by two blocks of 50 trials where the rewarding target region was adjusted again, and returned back to the 0ROT condition. Subjects then performed a block of 100 trials in a -4ROT condition (reaches now rewarded if they fell between -8° and 0°). Two blocks of 50 trials in the 0ROT condition followed this perturbation.

For the remainder of the experiment the participants performed two blocks of 200 trials in the 0ROT condition, but the probability of reward was now controlled. For example, in the 40% reward condition, if the movement placed the cursor in the rewarding target area, the probability of reward was 40%. This reward probability was altered and left constant for 25 sequential trials. Participants experienced each reward condition of 40%, 60%, 80% and 100% over two separate 25 trial intervals, as shown in Figure 1B.

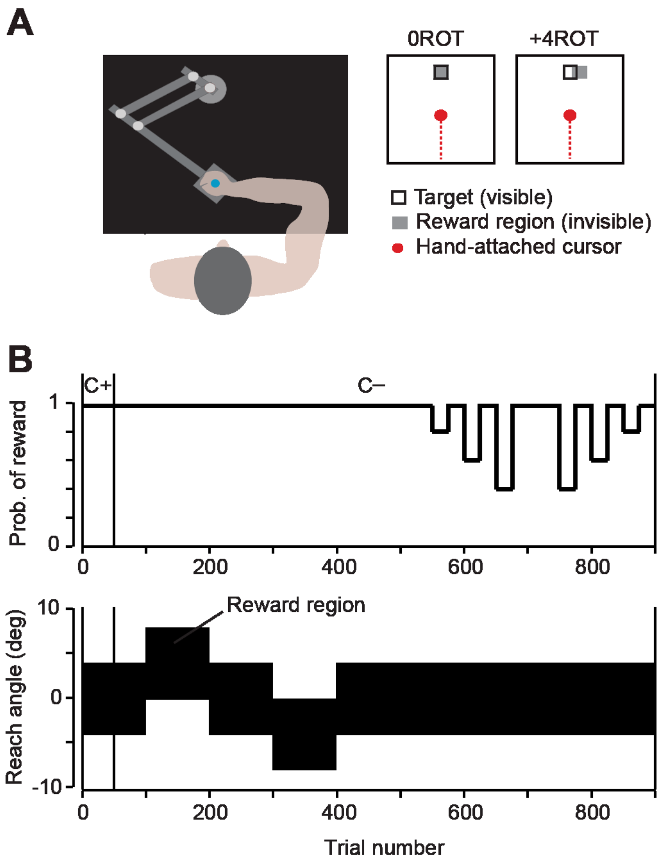


Figure 5.1 Experimental task and set up. **(A)** Participants held on to the handle of a robotic manipulandum and made shooting movements through a target located 10 cm away from start. A screen was placed over the subject's hand in order to obscure the participant's view of their arm and the robotic handle. **(B)** During the first 50 trials of the experiment, a cursor was projected onto the screen in order to provide the subject with visual feedback of their movement. After these 50 trials, the cursor was shut off. For the remaining trials, the only feedback the participant received was a brief target explosion and a point added to their score following a successful trial. The gray area indicates the region through which the participant needed to reach in order to receive successful feedback. In the second half of the experiment, this region was held constant, and the probability of receiving a reward if the hand crossed through this area was then manipulated.

5.2.2 Data Analysis

Movement Analysis

Hand position and velocity were recorded at the robotic handle at 100Hz and analyzed offline with MATLAB R2009b. The main variable for performance was the reach angle of the participant's movement. First, a reach endpoint was defined as the point at which the participant's hand crossed a circle with radius 10 cm centered at the start position. A reach angle was calculated for each movement as the angle between the hand path from start to reach endpoint and the line connecting start to target center.

Statistical Analysis

Statistical analysis was accomplished using IBM SPSS Version 19. All one-way ANOVA were tested for the assumption of homogeneity of variance, using the Levene's F Test for equality of variance. For those one-way ANOVA tests in which this assumption is violated, the Brown-Forsythe statistic is reported. In these cases, the Games-Howell post hoc test was then used. For cases in which the assumption of homogeneity of variance has been met, the Tukey (HSD) test is used for post-hoc analysis.

5.3 Results

5.3.1 Baseline reach variability was comparable between groups

Participants began the experiment with a familiarization block in which full visual feedback was provided via a cursor (C+ trials, Figure 1B). These reaches were performed in the 0ROT condition in which the invisible reach reward region (±4°, gray region in Figure 1A) was centered on the visible target (black box in Figure 1A). We observed no statistically significant differences among groups in the number of successful trials ($F_{(2,13.211)}=2.203$, $p=0.149$, one-way ANOVA for total reward in last 25 trials), reach variability (trial-to-trial change in reach direction, $F_{(2,25)}=0.300$, $p=0.743$, one-way ANOVA for average absolute difference in reach angle in last 25 trials), or reach peak velocity ($F_{(2,25)}=0.877$, p = 0.0430, average maximum velocity in last 25 trials). We did observe a significant effect of group for reach reaction time ($F_{(2,25)}=4.399$, p = 0.024, one-way ANOVA for average reaction time in last 25 trials). Here the PD patients had a longer reaction time than the young controls (p = 0.019), though there was no significant difference between the two control groups (p = 0.425) or between the PD patients and the older controls (p = 0.266).

Following this baseline block, cursor feedback was removed (C- trials, Figure 1B) and participants performed another block of 50 trials in the 0ROT condition. We again found no statistically significant differences across groups in terms the number of successful trials ($F_{(2,25)}=0.967$, p=0.395, one-way ANOVA for reward in last 25 trials), reach variability (trial-to-trial change in reach direction, $F_{(2,25)}=0.677$, p = 0.517, one-way ANOVA for average absolute difference in reach angle in last 25 trials) or in reach peak velocity ($F_{(2,25)}=1.578$,p =0.228, average maximum velocity in last 25 trials). We did however, observe a statistical difference in reach reaction time ($F_{(2,12.471)}=7.996$, p = 0.006, average reaction time in last 25 trials), with the patients with PD exhibiting a longer reaction time than the young controls (p = 0.020). Importantly, there was again no statistical difference between the PD patients and the aged controls (p = 0.71) or between the two control groups (p = 0.327). Therefore, the patients were able to meet the physical demands of the task, even with the absence of visual feedback. Additionally, there was no evidence of baseline differences in trial-to-trial reach variability or success rate across the groups.

5.3.2 Reach variability increased following an unrewarded trial

We next shifted the reward region covertly with respect to the target, requiring participants to alter their reach direction in order to continue receiving reward (trials 100-200, Figure 1B). Because no cursor feedback was available in these and all subsequent trials, the only information provided was the successful acquisition of reward (R+) or failure (R-) at the end of each trial.

For the first perturbation (+4ROT condition), the reward region was shifted by +4° for 100 trials (reaches were rewarded if they fell between 0° and +8°, as illustrated in Figure 1B). This block of training was followed by 100 trials of washout in which the reward region was returned to the 0ROT condition. After this first washout period, participants experienced 100 trials in the -4ROT condition followed by another 100 trials of washout in the 0ROT condition.

Reach angles are plotted in Figure 2A for a typical subject from each group. (These three participants were selected for display because they achieved similar scores during this block of trials, receiving reward on 88.0, 88.8 and 89.4% of the 500 trials for the volunteer in the young control, aged control, and PD groups respectively.) A first glance, the data in Figure 2A suggest that each subject varied their reach in order to find the reward zone. However, to more closely analyze the data, we quantified how much the reach angle was changed from trial-to-trial as a function of whether the initial trial was rewarded (R+) or not (R-). In this analysis, we measured change in reach angle u from trial n to trial $n+1$, and represented this change as:

$$\Delta u^{(n)} = \left| u^{(n+1)} - u^{(n)} \right| \tag{1}$$

We next quantified the change in reach angles following each trial *n*, in which the reach was rewarded and computed a conditional probability distribution $p\left(\Delta u \middle| R+\right)$ for each subject (green colored distribution, Figure 2B). Similarly, we quantified the change in reach angles given that trial n was not rewarded $p\left(\Delta u \middle| R-\right)$ (red colored distribution, Figure 2B).

In comparing behavior of the healthy subjects following R+ and R- trials, we found that the distributions appeared broader following R- trials. That is, following an unrewarded trial both of these representative healthy subjects changed their reach angle by a larger amount than following a rewarded trial. This implies that if the current motor commands were unrewarded, these two healthy subjects changed their motor commands on the next trial by a fairly large amount, perhaps searching for a better solution. In the representative PD subject, the distribution following an R+ trial appeared similar to the two healthy subjects (green distributions, Figure 2B). However, the distribution for $p\left(\Delta u \middle| R-\right)$ appeared narrows than normal. This implies that for the PD subject, there was less change in the reach angles following an unrewarded trial than in the healthy subjects.

To compare the behavior of all subjects following a rewarded or unrewarded trial, we fit a folded normal distribution to the change in reach angle for each subject for the two reward conditions, using the equation:

$$f(\Delta u;\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(-\Delta u-\mu)^2}{2\sigma^2}\right)+\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(\Delta u-\mu)^2}{2\sigma^2}\right)\quad \Delta u\geq 0 \qquad (2)$$

In Eq. (2), the mean μ and standard deviation σ values were calculated for each individual participant's range of reach angle differences following an R+ and an R- trial. These fits are plotted with the heavy line in Figure 2B for the three representative subjects.

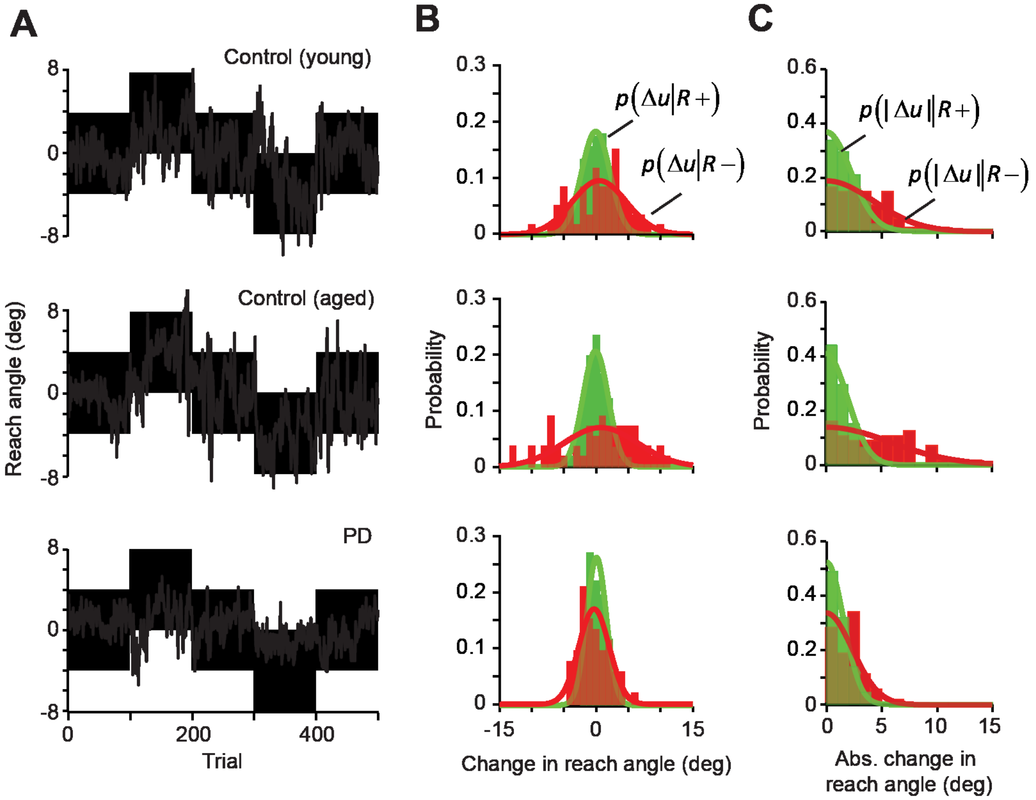


Figure 5.2 Participants respond differentially to positive vs. negative task outcomes. **(A)** Reach angles of representative subjects. The gray area indicates the region where the reach would be rewarded. **(B)** Absolute change in reach angle for representative subjects. Plotted in green are the absolute change in reach angle between trial n and n+1 for each subject following a successful trial (R+). The changes in reach angle following an unsuccessful trial (R-) are plotted in red. The fit of these absolute changes to a folded normal distribution are plotted over the histogram of each individual's data.

Figure 3A plots the average fit across all participants in each group, and Figure 3B plots the mean of the distributions for each group in the R+ and R- conditions. We found both a significant effect of condition (R+ vs R-) and a significant interaction between condition and group (effect of condition, $F_{(1,23)}=193.806$, $p<0.001$ and condition by group interaction $F_{(2,23)}=6.231$, $p = 0.007$) following a repeated measures ANOVA with reward condition as the within-subject measure and groups as the between-subject measure. We observed that across all groups, the mean change in reach angle was greater following an unrewarded trial than a rewarded trial ($p < 0.001$). A post-hoc test in which we analyzed each condition individually revealed that following an R+ trial, the mean change in reach angle was not different among the PD, aged, and young groups (one-way ANOVA on average absolute change in reach angle following an R+ trial, $F_{(2,25)}=0.512$, $p=0.606$). However, this change was smaller than normal in the PD group following an R- trial (one-way ANOVA on average absolute change in reach angle following unsuccessful trial, $F_{(2,25)}=3.681$, $p= 0.041$).

Was this policy of changing the reach angle useful in acquisition of reward? We found that the PD patients had a lower average score (number of rewarded trials) at the end of the +4ROT and -4ROT conditions (post-hoc comparisons, $p=0.038$ and $p=0.007$ against age-matched and young controls respectively, following a significant one-way ANOVA on total reward in last 25 trials from both blocks, $F_{(2,16.098)}=6.638$, $p=0.008$). Therefore, when the only feedback was a binary reward signal, it appeared that actively searching the reach direction space after an unsuccessful trial was a good strategy for acquiring more reward.

In summary, when a motor command was rewarded (R+ trials), the trial-to-trial change following this command was similar in all three groups. When a motor command was not rewarded (R- trials), there was a larger trial-to-trial change following this command than after R+ trials. However, after an R- trial both healthy groups altered their motor commands on the next trial more than the PD subjects. This hinted that sensitivity to an R- trial was lower in PD than in controls. To test for this systematically, we performed a second experiment.

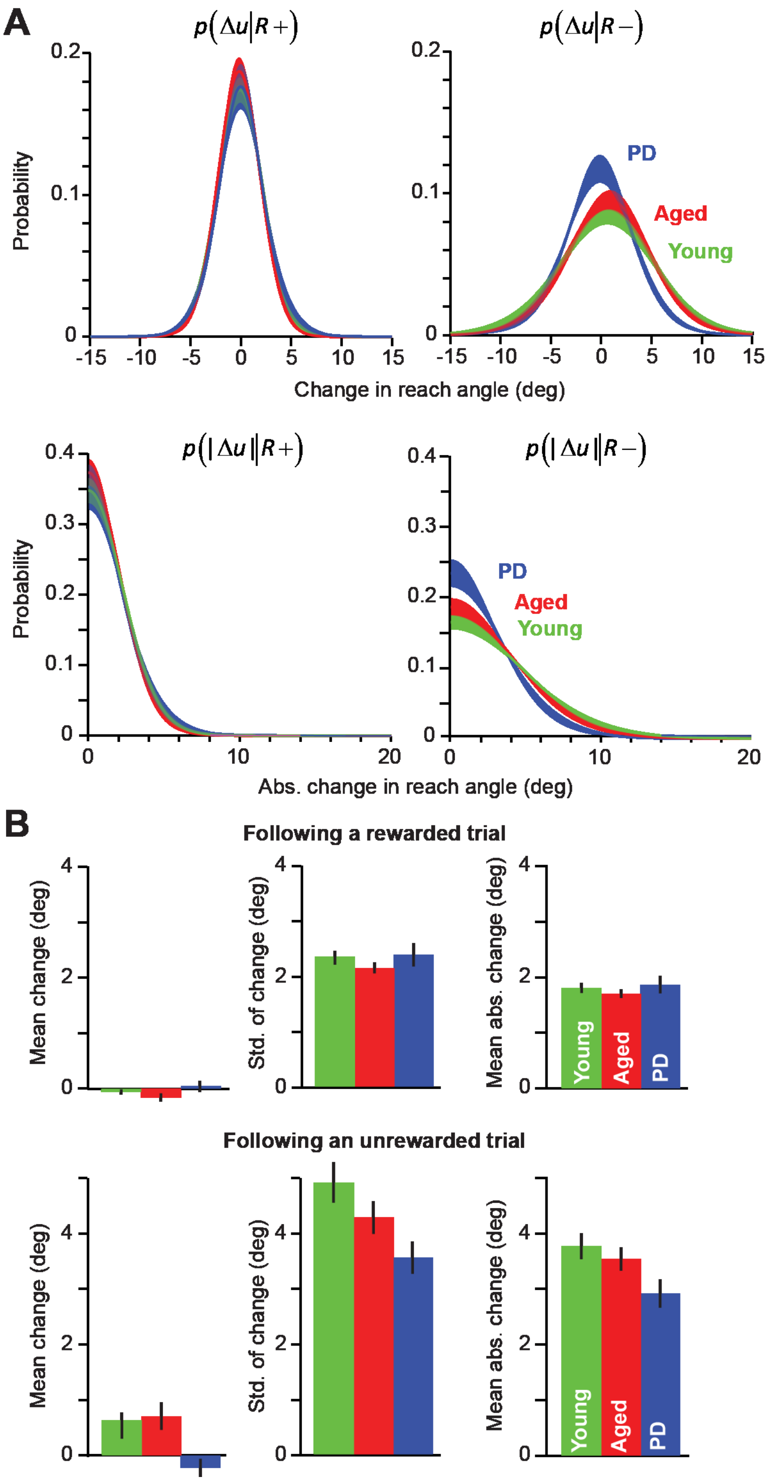


Figure 5.3 PD participants are less sensitive to negative outcomes. **(A)** Group average for the individual subject fits to a folded normal distribution for the absolute change in reach angle, following a successful (R+) or unsuccessful trial (R-). Shaded area indicates SEM. **(B)** The mean absolute change in reach angle for each group, following the presence or absence of reward (R+ or R- trials). Error bars are SEM.

5.3.3 A relationship between probability of reward and reach variability

To directly test whether the absence of reward resulted in increased trial-to-trial change in the reach angles, we performed an experiment in which we controlled the probability of reward on each trial (final 400 trials of Figure 1B). In this experiment all the trials were in the 0ROT condition but we regulated the probability of reward: if the subject's reach placed the unseen cursor in the reward region, reward was provided at a probability of 40, 60, 80 or 100% for bins of 25 trials, as shown in Figure 1B.

For each subject we computed the average Δu in each probability condition and found that in the two healthy groups Δu was largest when the probability of reward was lowest (40%), and then Δu gradually declined as the probability of reward increased (Figure 4A). Therefore, in healthy people we found that a lower probability of reward coincided with larger changes in reach angle. However, this trend was not present in the PD group. Rather, the PD patients appeared to exhibit approximately the same level of Δu across all reward probabilities. A repeated measures ANOVA with reward probability as the within-subject measure and groups as the between-subject measure found a significant group by reward interaction ($F_{(6,69)}=4.699, p<0.001$). A post-hoc test in which we analyzed each group individually revealed a significant effect of reward probability for both control groups ($F_{(3,31)}=6.51$, $p = 0.002$, for aged and $F_{(3,35)}=9.01$, $p<0.001$ for young, one-way ANOVA on reward probability). In contrast, in the patients we found no significant effect of reward probability on Δu ($F_{(3,35)}=0.281$, $p = 0.838$, one-way ANOVA on reward condition). Notably, the three groups had similar performances during the highly rewarded condition ($F_{(2,25)} = 0.797$, $p = 0.462$ for one-way ANOVA on 100% reward condition alone), indicating that the PD patients were not impaired at the task overall, but instead did not adjust their level of active search in response to a lack of reward.

Figure 4A had displayed Δu for each reward probability that were experimentally imposed. However, this probability does not necessarily equal the probability of reward that the participants actually acquired during the experiment. To examine this question, in Figure 4B we have plot the change in reach angle over the actual reward probability achieved for each subject. In order to compare the change in reach angle vs. probability of reward relationship, for each subject we applied a linear regression and estimated the slope (Figure 4C). We again found that the PD patients had a significantly smaller slope than the two control groups (post-hoc $p = 0.031$ vs. young, and $p = 0.003$ vs. aged, following a significant effect of group, one-way ANOVA, $p = 0.003$, $F_{(2,25)}=7.686$).

These results suggested that in the two healthy groups of volunteers the probability of reward significantly modulated the change in motor output: as reward probability decreased, the trial-to-trial change in reach angles increased. However, in PD reward probability was not a significant modulator of reach variability.

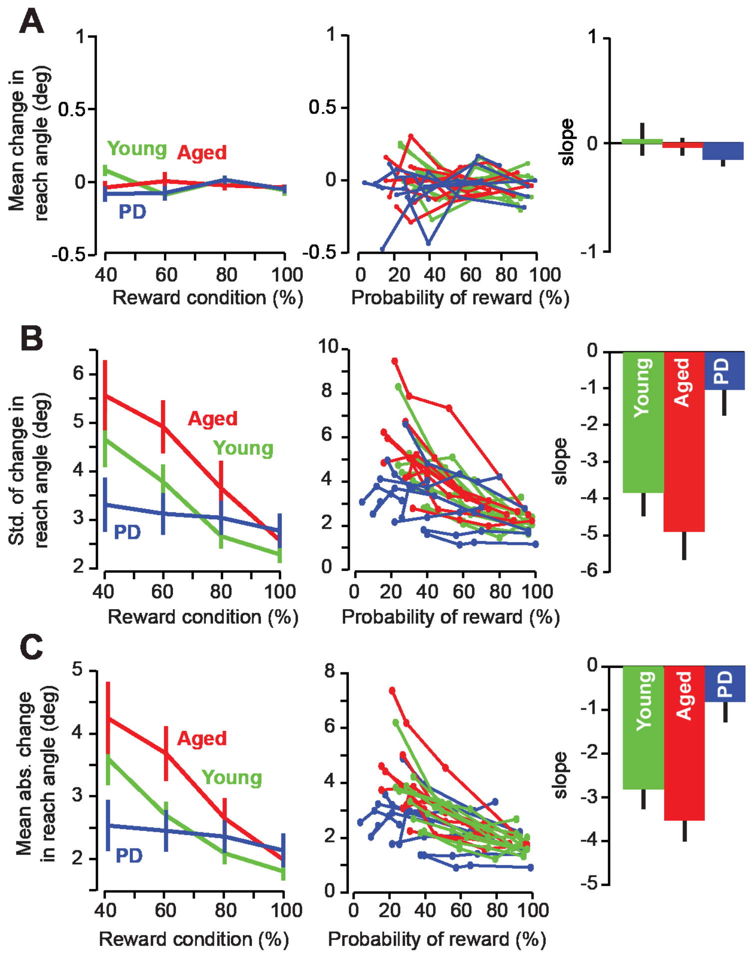


Figure 5.4 Reward based modulation of variability is impaired in PD. **(A)** Group data, indicating average absolute change in reach angle during each reward condition. Error bars are SEM. **(B)** Individual subject data indicating the average absolute change in reach angle, and actual amount of reward received during each reward condition. **(C)** Group data indicating the average slope, which was fit to each individual subject's active search plotted in (B). Error bars are SEM.

5.3.4 Measuring sensitivity to history of reward

To mathematically describe the relationship between Δu and reward, we extended our trial-to-trial analysis to include the history of reward. In Figure 5A we have plotted Δu as a function of the history of reward (for this analysis, we have included data from all trials, 1-900, from the two parts of the experiment). The history of reward is represented by variables $R(n)$, $R(n-1)$, and $R(n-2)$, indicating whether the subject was successful in trials n, n-1, and n-2, respectively. [In Fig. 5A, history of reward is a binary vector, ordered from left to right from all 3 trials successful to no trial successful.] We found that across groups, Δu was largest when the last 3 trials had been unsuccessful. Indeed a repeated measure ANOVA with reward as the within-subject measure and group as the between-subject measure produced a significant effect of reward history on Δu ($p < 0.001$, $F_{(7,161)}=45.073$), as well as an interaction between group and reward ($p < 0.001$, $F_{(14,161)}=3.453$), suggesting that sensitivity to reward was smaller in PD than in the other two groups.

To quantify this relationship across each participant, we formulated a state-space model to relate Δu to the history of the past rewards.

$$\Delta u(n) = \alpha_0\left(1 - R(n)\right) + \alpha_1\left(1 - R(n-1)\right) + \alpha_2\left(1 - R(n-2)\right) + \varepsilon \tag{3}$$

In the above equation, the change in reach angle on trial n is written as a function of reward history in the last few trials. The term α_0 represents sensitivity to reward prediction error in the current trial, and ε is the variability in the movements that cannot be explained by reward prediction errors. A large α_0 would indicate that following an unsuccessful trial there would be a large change in reach angle. We fit the above equation to the data from each participant and have plotted the resulting parameter values in Figure 5B. We found that sensitivity was largest to the reward prediction error in the current trial, and then declined with trial history. A repeated measures ANOVA with sensitivity as the within-subject measure and group as the between-subject measure found a significant effect of sensitivity ($F_{(2,46)} = 90.222$, $p<0.001$), and a significant group by sensitivity interaction ($F_{(4,46)}=4.363$, $p = 0.004$). Sensitivity prediction error α_0 was significantly smaller in PD than in the two groups ($p = 0.032$ for young, $p = 0.014$ for aged). There were no differences between the sensitivity values of the control groups ($p = 0.657$). Although the PD patients did respond to negative reward prediction error, as evidenced by a non-zero α_0 value (z-test, $p < 0.001$), they appeared to be less affected by this feedback and did not adjust their motor commands to the same extent as the controls. PD participants were also less affected by reward prediction errors from trials further in the past. The α values for these participants quickly decreased to the point where the α_2 was not significantly different than 0 (z-test, $p = 0.407$).

Once exploration successfully leads to reward, the best strategy is to maintain this performance. This behavior is captured by the ε value, which determines the change in reach angle following a series of rewarded trials. The PD patients had a similar ε value as the control groups (no effect of group on a one-way ANOVA, $F_{(2,25)}=2.642$, $p = 0.093$), indicating that these participants had the same amount of natural variability in maintaining their reach direction following a rewarded trial. This fact is further supported by the similar trial-to-trial reach angle changes that were observed across groups during the initial baseline blocks and in the 100% reward condition, in which many trials were rewarded and motor exploration was unnecessary.

These results show that trial-to-trial changes in motor output are partially driven by the history of reward. These changes are most sensitive to reward prediction error in the current trial, and have smaller sensitivity to prediction errors in the previous trials. In

PD, trial-to-trial variability is comparable to healthy controls following rewarded trials, but exhibits smaller than normal sensitivity to trials that are not rewarded.

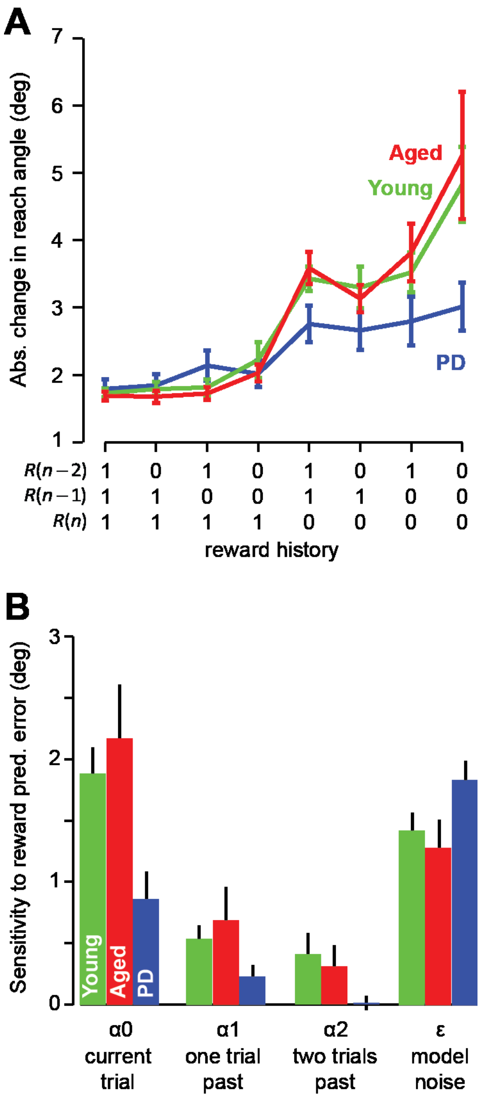


Figure 5.5 A state-space model. **(A)** Group data for the absolute change in reach angle between trials n and n+1, as a function of the reward history for the three most recent trials, (R(n), R(n-1), R(n-2)). Error bars are SEM. (**B**) A state-space model was used to determine each subject's change in reach angle as a function of the past history of reward. Data indicate group average, error bars are SEM.

5.4 Discussion

We examined the hypothesis that trial-to-trial changes in motor commands are an active search process driven by history of reward. Our hypothesis made two key predictions. First, trial-to-trial changes in motor commands should be driven by the success or failure of each reach: subjects should stay with their current motor commands if the trial was successful, but change if it was not. As a result, the search for a rewarding outcome should lead to large between trial changes in performance during periods of low reward probability, but small changes during periods of high reward probability. Second, because in this formulation variability in motor commands is partly driven by the mismatch between expected and achieved success, if a disease affects the ability of the brain to encode reward prediction errors, then the trial-to-trial changes in motor commands in response to failure will also be affected. As a result, the search for a rewarding outcome may be impaired in people diagnosed with Parkinson's disease (PD). That is, the ability to increase variability in response to reward prediction error should be impaired in PD.

To test our hypothesis we used a reaching task in which the only performance feedback was success or failure of the trial. We first confirmed that in healthy control subjects, larger inter-trial changes in reach angle were observed during periods of low reward probability. To estimate the relationship between change in reach angle and history of success, we used a model similar to one employed by Bayer & Glimcher (Bayer and Glimcher, 2005) to capture the response of midbrain dopamine neurons to a reward signal in monkeys. We found that the control participants were highly sensitive to the outcome of the most recent trial. Patients with Parkinson's disease also showed the greatest dependence on the outcome of the most recent trial, but exhibited smaller trial to trial changes in motor commands overall. Compared to the control participants, the PD subjects had similar levels of variability during periods of high reward probability, but a much smaller change in trial to trial reach angle following unrewarded trials. This indicated that participants with PD were impaired at modulating their variability, but only in response to unrewarding outcomes.

Here we saw that the PD subjects on their usual schedule of medication exhibited an impaired ability to control reach variability, but only after experiencing a negative outcome. By our measures the impairment in the PD subjects was not in learning per se, but in the failure to increase their performance variability in order to search the task space and achieve reward. It has been considered that the major source of the human motor variability is in the peripheral motor organs, including motor neurons (Jones et al., 2002). However, variability exists in neural activity during motor planning in the premotor cortex which might also contribute to variability in movement execution (Churchland et al., 2006). It is therefore reasonable that the motor variability observed in a typical motor control experiment is composed of both the peripheral and central motor noise. Since there is no major dopaminergic modulation of motor neurons via the basal ganglia, PD might reduce the sensitivity of the central motor noise to the negative rewards. Importantly, the baseline motor variability of PD was comparable with that of the age-matched, suggesting that the peripheral motor noise was not influenced by PD.

The impaired ability of the PD patients to respond differentially to positive vs. negative outcomes has also been observed in other tasks which depend on trial and error learning. Importantly, there is evidence that when the patients are on medication (as was the case in our study), they are specifically impaired at learning from negative reward prediction errors (Frank et al., 2004;Frank et al., 2007;Bodi et al., 2009). Though learning deficits have been reported in many associative learning tasks, this has not proven to be ubiquitous. The heterogeneity of results is believed to be due to differences in task demands, clinical severity of the disease, and importantly, the presence or absence of medication (Shohamy et al., 2006;Cools et al., 2001;Rutledge et al., 2009;Frank et al., 2004;Frank et al., 2007;Bodi et al., 2009). Dopaminergic agents have been found to alter the ability to learn during feedback based tasks in healthy controls (Pizzagalli et al., 2008;Pessiglione et al., 2006). The role of these agents in altering the ability to learn from reward prediction error is thought to result from the interruption of the phasic bursts and pauses of dopaminergic cell firings, which code for positive and negative reward prediction errors respectively. Unfortunately, PD patients experience significant discomfort during periods of withdrawal from their usual medication schedule, and therefore we did not collect data from an experimental group of PD patients in the off medication state. Selecting an action to maximize the probability of reward on the current and future trials is associated with striatal activity (Samejima et al., 2005;Tanaka et al., 2004), and is also susceptible to manipulations of dopaminergic activity. Dopamine levels are thought to play a role in the selection of low-level parameters of movement, such as the velocity, acceleration, and latency of an action (Mazzoni et al., 2007;Galea et al., 2013;Niv et al., 2007). A recent study found that altering D1 and D2 activity in mice could alter the decision to switch or stay with a given behavior, and that this decision was influenced by reward history (Tai et al., 2012). The use of D2 antagonists and agonists have been found to increase or decrease behavioral variability respectively, specifically following negative outcomes (Galea et al., 2013;Pesek-Cotton et al., 2011). In our task the PD patients were less likely to increase their performance variability following a negative reward prediction error, indicating a reluctance to switch from their current action despite the lack of success. This is similar to other studies that have found that PD patients may settle on a less than optimal solution in order to complete a task (Vakil et al., 2014;Shohamy et al., 2004), which in this instance is persistence instead of exploration.

Despite impairments in learning from reward prediction errors, PD patients have exhibited normal behavior in many standard motor learning tasks. PD patients are able to perform comparably to controls on motor skill and mirror inversion tasks while on medication (Paquet et al., 2008;Agostino et al., 1996). Several studies have shown that PD patients are also able to adapt to visuomotor rotations as well as control participations, though consolidation of this learning is impaired in those with the disease (Marinelli et al., 2009;Bedard and Sanes, 2011;Leow et al., 2012;Leow et al., 2013). This intact performance during motor learning in PD is presumably due to their ability to recruit other learning processes that do not depend on reward prediction errors, such as learning from sensory prediction errors which may depend on the cerebellum (Izawa et al., 2012). Indeed earlier we found distinct signatures of learning from sensory vs. reward prediction errors, and predicted that the basal ganglia structures were responsible for altering movements in response to reward (Izawa and Shadmehr, 2011b). Similarly, Huang et. al found that motor adaptation employs several learning processes, with the reinforcement of successful actions proving to be a critical component in order to achieve savings (Huang et al., 2011b). Based on this result, the authors predicted that in PD patients, impairments in with reinforcement learning may be the cause of the lack of consolidation seen in the previously mentioned visuomotor adaptation tasks. Though we did not test this hypothesis directly, the impairment of PD subjects to respond accordingly to negative reward prediction errors in our reaching task would support this idea.

In conclusion, we find that in response to negative reward prediction errors, healthy adults changed their motor commands in order to search the task space for a more rewarding solution. As a consequence, during periods of low reward probability healthy subjects increased their performance variability. This process appears to depend on basal ganglia structures, as PD participants were less responsive to a negative outcome, despite an intact ability to perform the task and to maintain normal performance following a positive outcome.

Subject ID	Age	Handedness	Sex	Disease duration (yrs)	motor UPDRS	total UPDRS
PD1	77	R	F	4	10	18
PD2	51	R	F	0.25	7	12
PD3	65	R	M	4	8	23
PD4	61	R	F	5	9	14
PD5	65	R	M	1	5	23
PD6	68	R	M	2	6	6
PD7	61	R	M	1	4	6
PD8	61	R	M	4	4	7
PD9	62	R	F	0.5	9	13

Table 1. Clinical characteristics of the volunteers with Parkinson's disease.

Chapter 6: Conclusions

The use of our robotic manipulandum provided us with the unique opportunity to dissect how people first learn to physically interact with a novel tool. Throughout these chapters, we trained volunteers to make reaching movements using our robotic manipulandum and then applied perturbations to the hand, to study how participants adapted to these imposed environments. The goal of this thesis work was to understand what cues were important for the acquisition of multiple motor memories, what aspects of training were retained, and how we could recall these motor commands.

In Chapter 2, we started by introducing a new method to analyze the components of motor memory. Through the use of a long block of error-clamp trials, we were able to analyze the force bias in these trials, and conclude that a portion of all force field training was retained in a single experimental session. This indicated that when participants successfully adapted to one environment on our novel tool, they were able to retain and protect this memory when they were introduced to another environment on the same device. Reinforcement of actions appeared to play a critical role in this retention, perhaps by serving as a cue that the environment had changed.

Through our experiments in Chapter 2, we also discovered that reinforcement plays a critical role in the recall of motor memories. When participants make a successful movement and achieve their goal, the best policy is to try and maintain these motor commands. As explored in Chapter 5, this leads to low motor variability during periods of high reward probability. However, when reinforcement is withheld, the best policy is to alter ones actions and search for more rewarding motor commands. In the experiments of Chapter 2, withholding reinforcement served as a performance cue to alert the participant that the dynamics of the tool had changed, encouraging participants to change their motor commands or recall previously successful actions. In Chapter 5, we quantified behavior in a different way and found periods of low reward probability lead to periods of high reach variability in healthy controls. Presumably this process of action selection has some dependency on the basal ganglia, as PD patients were unable to modulate their reach variability specifically in response to periods of low reward probability.

In Chapter 3, we found that once participants interacted with our robotic manipulandum, they generalized this learning to a similar robotic manipulandum. We attempted to train participants to learn an association between a given dynamical environment, and a specific robotic tool. However, we found that even with repeated training sessions, participants did not learn this pairing. This is similar to the difficulties we find in everyday life, when we switch between similar tools. For example, when I borrow someone else's computer, I notice that I make many errors when I first start to type with the unfamiliar computer, as I have grown accustomed to the key locations on my own computer. Even though I know I am not on my own computer and can see differences between the two tools, use of a similar computer nevertheless recalls my motor memories acquired on the computer I use most frequently.

Finally, we found that motor memories were not resistant to change. With passage of time, we found that inefficiencies in motor commands were reduced. This phenomena was not observed with additional training. One confound in this observation, was that with time participants also had more forgetting. Therefore at recall, we observed larger errors. Current experiments in the lab are following up on this observation, to determine if repeatedly relearning the motor commands also leads to improved efficiency, or if this is only achieved through passage of time.

Through this work, we were able to identify some of the cues that allow for the retention and recall of motor memories. We also determined what modifications occurs to motor memory over time. Through future work, these findings can be extended beyond the laboratory setting. For example, knowing what cues are necessary for acquisition and recall of a motor memory, we could potentially manipulate these cues to instead erase unwanted motor memories. Or if we could determine what links training to a given tool, we could try and reverse this pairing, which would be useful to rehabilitation strategies. Hopefully the work of this thesis serves as the groundwork for these types of experiments.

References

Addou T, Krouchev N, Kalaska JF (2011) Colored context cues can facilitate the ability to learn and to switch between multiple dynamical force fields. J Neurophysiol 106:163-183.

Agostino R, Sanes JN, Hallett M (1996) Motor skill learning in Parkinson's disease. J Neurol Sci 139:218-226.

Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 47:129-141.

Bedard P, Sanes JN (2011) Basal ganglia-dependent processes in recalling learned visual-motor adaptations. Exp Brain Res.

Bock O, Schneider S, Bloomberg J (2001) Conditions for interference versus facilitation during sequential sensorimotor adaptation. Exp Brain Res 138:359-365.

Bodi N, Keri S, Nagy H, Moustafa A, Myers CE, Daw N, Dibo G, Takats A, Bereczki D, Gluck MA (2009) Reward-learning and the novelty-seeking personality: a between- and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. Brain 132:2385-2395.

Bouton ME (2002) Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. Biol Psychiatry 52:976-986.

Brashers-Krug T, Shadmehr R, Bizzi E (1996) Consolidation in human motor memory. Nature 382:252-255.

Bushey D, Tononi G, Cirelli C (2011) Sleep and synaptic homeostasis: structural evidence in Drosophila. Science 332:1576-1581.

Caithness G, Osu R, Bays P, Chase H, Klassen J, Kawato M, Wolpert DM, Flanagan JR (2004) Failure to consolidate the consolidation theory of learning for sensorimotor adaptation tasks. J Neurosci 24:8662-8671.

Chen-Harris H, Joiner WM, Ethier V, Zee DS, Shadmehr R (2008) Adaptive control of saccades via internal feedback. J Neurosci 28:2804-2813.

Churchland MM, Afshar A, Shenoy KV (2006) A central source of movement variability. Neuron 52:1085-1096.

Cools R, Barker RA, Sahakian BJ, Robbins TW (2001) Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. Cereb Cortex 11:1136-1143.

Cos I, Belanger N, Cisek P (2011) The influence of predicted arm biomechanics on decision making. J Neurophysiol 105:3022-3033.

Cothros N, Wong J, Gribble PL (2008) Distinct haptic cues do not reduce interference when learning to reach in multiple force fields. PLoS One 3:e1990.

Cothros N, Wong J, Gribble PL (2009) Visual cues signaling object grasp reduce interference in motor learning. J Neurophysiol 102:2112-2120.

Cothros N, Wong JD, Gribble PL (2006) Are there distinct neural representations of object and limb dynamics? Exp Brain Res 173:689-697.

Criscimagna-Hemminger SE, Shadmehr R (2008) Consolidation patterns of human motor memory. J Neurosci 28:9610-9618.

Darainy M, Ostry DJ (2008) Muscle cocontraction following dynamics learning. Exp Brain Res 190:153-163.

Debas K, Carrier J, Orban P, Barakat M, Lungu O, Vandewalle G, Hadj TA, Bellec P, Karni A, Ungerleider LG, Benali H, Doyon J (2010) Brain plasticity related to the consolidation of motor sequence learning and motor adaptation. Proc Natl Acad Sci U S A 107:17839-17844.

Diedrichsen J, White O, Newman D, Lally N (2010) Use-dependent and error-based learning of motor behaviors. J Neurosci 30:5159-5166.

Donchin O, Sawaki L, Madupu G, Cohen LG, Shadmehr R (2002) Mechanisms influencing acquisition and recall of motor memories. J Neurophysiol 88:2114-2123.

Doya K, Samejima K, Katagiri K, Kawato M (2002) Multiple model-based reinforcement learning. Neural Comput 14:1347-1369.

Doyon J, Korman M, Morin A, Dostie V, Hadj TA, Benali H, Karni A, Ungerleider LG, Carrier J (2009) Contribution of night and day sleep vs. simple passage of time to the consolidation of motor sequence and visuomotor adaptation learning. Exp Brain Res 195:15-26.

Ethier V, Zee DS, Shadmehr R (2008) Spontaneous recovery of motor memory during saccade adaptation. J Neurophysiol 99:2577-2583.

Flanagan JR, Bittner JP, Johansson RS (2008) Experience can change distinct size-weight priors engaged in lifting objects and judging their weights. Curr Biol 18:1742-1747.

Flook JP, McGonigle BO (1977) Serial adaptation to conflicting prismatic rearrangement effects in monkey and man. Perception 6:15-29.

Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12:189-198.

Frank MJ, Samanta J, Moustafa AA, Sherman SJ (2007) Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. Science 318:1309-1312.

Frank MJ, Seeberger LC, O'reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science 306:1940-1943.

Franklin DW, Osu R, Burdet E, Kawato M, Milner TE (2003) Adaptation to stable and unstable dynamics achieved by combined impedance control and inverse dynamics model. J Neurophysiol 90:3270-3282.

Galea JM, Ruge D, Buijink A, Bestmann S, Rothwell JC (2013) Punishment-induced behavioral and neurophysiological variability reveals dopamine-dependent selection of kinematic movement parameters. J Neurosci 33:3981-3988.

Galea JM, Sami SA, Albert NB, Miall RC (2010) Secondary tasks impair adaptation to step- and gradual-visual displacements. Exp Brain Res 202:473-484.

Gandolfo F, Mussa-Ivaldi FA, Bizzi E (1996) Motor learning by field approximation. Proc Natl Acad Sci U S A 93:3843-3846.

Haruno M, Wolpert DM, Kawato M (2001) Mosaic model for sensorimotor learning and control. Neural Comput 13:2201-2220.

Hirashima M, Nozaki D (2012) Distinct motor plans form and retrieve distinct motor memories for physically identical movements. Curr Biol 22:432-436.

Howard IS, Wolpert DM, Franklin DW (2013) The effect of contextual cues on the encoding of motor memories. J Neurophysiol 109:2632-2644.

Huang HJ, Kram R, Ahmed AA (2012) Reduction of metabolic cost during motor learning of arm reaching dynamics. J Neurosci 32:2182-2190.

Huang VS, Haith A, Mazzoni P, Krakauer JW (2011a) Rethinking motor learning and savings in adaptation paradigms: model-free memory for successful actions combines with internal models. Neuron 70:787-801.

Huang VS, Haith A, Mazzoni P, Krakauer JW (2011b) Rethinking motor learning and savings in adaptation paradigms: model-free memory for successful actions combines with internal models. Neuron 70:787-801.

Huang VS, Shadmehr R (2009) Persistence of motor memories reflects statistics of the learning event. J Neurophysiol 102:931-940.

Hwang EJ, Shadmehr R (2005) Internal models of limb dynamics and the encoding of limb state. J Neural Eng 2:S266-S278.

Hwang EJ, Smith MA, Shadmehr R (2006) Dissociable effects of the implicit and explicit memory systems on learning control of reaching. Exp Brain Res 173:425-437.

Izawa J, Criscimagna-Hemminger SE, Shadmehr R (2012) Cerebellar contributions to reach adaptation and learning sensory consequences of action. J Neurosci 32:4230-4239.

Izawa J, Rane T, Donchin O, Shadmehr R (2008) Motor adaptation as a process of reoptimization. J Neurosci 28:2883-2891.

Izawa J, Shadmehr R (2011a) Learning from sensory and reward prediction errors during motor adaptation. PLoS Comput Biol 7:e1002012.

Izawa J, Shadmehr R (2011b) Learning from sensory and reward prediction errors during motor adaptation. PLoS Comput Biol 7:e1002012.

Jacobs R, Jordan M, Nowlan S, Hinton G (1991) Adaptive mixture of local experts. Neural Computation 3:79-87.

Jeannerod M (1984) The timing of natural prehension movements. J Mot Behav 16:235-254.

Jenmalm P, Johansson RS (1997) Visual and somatosensory information about object shape control manipulative fingertip forces. J Neurosci 17:4486-4499.

Joiner WM, Smith MA (2008) Long-term retention explained by a model of short-term learning in the adaptive control of reaching. J Neurophysiol 100:2948-2955.

Jones KE, Hamilton AF, Wolpert DM (2002) Sources of signal-dependent noise during isometric force production. J Neurophysiol 88:1533-1544.

Kao MH, Doupe AJ, Brainard MS (2005) Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. Nature 433:638-643.

Karniel A, Mussa-Ivaldi FA (2002) Does the motor control system use multiple models and context switching to cope with a variable environment? Exp Brain Res 143:520-524.

Keisler A, Shadmehr R (2010) A shared resource between declarative memory and motor memory. J Neurosci 30:14817-14823.

Kistemaker DA, Wong JD, Gribble PL (2010) The central nervous system does not minimize energy cost in arm movements. J Neurophysiol 104:2985-2994.

Kluzik J, Diedrichsen J, Shadmehr R, Bastian AJ (2008) Reach adaptation: what determines whether we learn an internal model of the tool or adapt the model of our arm? J Neurophysiol 100:1455-1464.

Knowlton BJ, Mangels JA, Squire LR (1996) A neostriatal habit learning system in humans. Science 273:1399-1402.

Korman M, Doyon J, Doljansky J, Carrier J, Dagan Y, Karni A (2007) Daytime sleep condenses the time course of motor memory consolidation. Nat Neurosci 10:1206-1213.

Koshland GF, Galloway JC, Nevoret-Bell CJ (2000) Control of the wrist in three-joint arm movements to multiple directions in the horizontal plane. J Neurophysiol 83:3188-3195.

Krakauer JW, Ghez C, Ghilardi MF (2005) Adaptation to visuomotor transformations: consolidation, interference, and forgetting. J Neurosci 25:473-478.

Krakauer JW, Ghilardi MF, Ghez C (1999) Independent learning of internal models for kinematic and dynamic control of reaching. Nat Neurosci 2:1026-1031.

Krouchev NI, Kalaska JF (2003) Context-dependent anticipation of different task dynamics: rapid recall of appropriate motor skills using visual cues. J Neurophysiol 89:1165-1175.

Lee JY, Schweighofer N (2009) Dual adaptation supports a parallel architecture of motor memory. J Neurosci 29:10396-10404.

Leow LA, de RA, Loftus AM, Hammond G (2013) Different mechanisms contributing to savings and anterograde interference are impaired in Parkinson's disease. Front Hum Neurosci 7:55.

Leow LA, Loftus AM, Hammond GR (2012) Impaired savings despite intact initial learning of motor adaptation in Parkinson's disease. Exp Brain Res 218:295-304.

Lewis D, Smith PN, McAllister D (1952) Retroactive Facilitations and Interference in Performance on the Modified Two-Hand Coordination Task. J Exp Psychol 44:247-260.

Maret S, Faraguna U, Nelson AB, Cirelli C, Tononi G (2011) Sleep and waking modulate spine turnover in the adolescent mouse cortex. Nat Neurosci 14:1418-1420.

Marinelli L, Crupi D, Di RA, Bove M, Eidelberg D, Abbruzzese G, Ghilardi MF (2009) Learning and consolidation of visuo-motor adaptation in Parkinson's disease. Parkinsonism Relat Disord 15:6-11.

Martin TA, Keating JG, Goodkin HP, Bastian AJ, Thach WT (1996) Throwing while looking through prisms. II. Specificity and storage of multiple gaze-throw calibrations. Brain 119 (Pt 4):1199-1211.

Mazur JE (1995) Conditioned reinforcement and choice with delayed and uncertain primary reinforcers. J Exp Anal Behav 63:139-150.

Mazzoni P, Hristova A, Krakauer JW (2007) Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. J Neurosci 27:7105-7116.

McGonigle BO, Flook J (1978) Long-term retention of single and multistate prismatic adaptation by humans. Nature 272:364-366.

Medina JF, Garcia KS, Mauk MD (2001) A mechanism for savings in the cerebellum. J Neurosci 21:4081-4089.

Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease (2003) The Unified Parkinson's Disease Rating Scale (UPDRS): status and recommendations. Movement disorders: official journal of the Movement Disorder Society 18:738.

Nader K, Schafe GE, Le Doux JE (2000) Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. Nature 406:722-726.

Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: opportunity costs and the control of response vigor. Psychopharmacology (Berl) 191:507-520.

Olveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. PLoS Biol 3:e153.

Orban de Xivry JJ, Criscimagna-Hemminger SE, Shadmehr R (2011) Contributions of the motor cortex to the adaptive control of reaching depend on the perturbation schedule. Cereb Cortex 21:1475-1484.

Overduin SA, Richardson AG, Lane CE, Bizzi E, Press DZ (2006) Intermittent practice facilitates stable motor memories. J Neurosci 26:11888-11892.

Page S, Neuringer A (1985) Variability is an operant. Journal of Experimental Psychology: Animal Behavior Processes429.

Paquet F, Bedard MA, Levesque M, Tremblay PL, Lemay M, Blanchet PJ, Scherzer P, Chouinard S, Filion J (2008) Sensorimotor adaptation in Parkinson's disease: evidence for a dopamine dependent remapping disturbance. Exp Brain Res 185:227-236.

Pasquereau B, Turner RS (2013) Limited encoding of effort by dopamine neurons in a cost-benefit trade-off task. J Neurosci 33:8288-8300.

Patton JL, Stoykov ME, Kovic M, Mussa-Ivaldi FA (2006) Evaluation of robotic training forces that either enhance or reduce error in chronic hemiparetic stroke survivors. Exp Brain Res 168:368-383.

Paz R, Boraud T, Natan C, Bergman H, Vaadia E (2003) Preparatory activity in motor cortex reflects learning of local visuomotor skills. Nat Neurosci 6:882-890.

Pekny SE, Criscimagna-Hemminger SE, Shadmehr R (2011) Protection and expression of human motor memories. J Neurosci 31:13829-13839.

Pesek-Cotton EF, Johnson JE, Newland MC (2011) Reinforcing behavioral variability: an analysis of dopamine-receptor subtypes and intermittent reinforcement. Pharmacol Biochem Behav 97:551-559.

Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006) Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442:1042-1045.

Pizzagalli DA, Evins AE, Schetter EC, Frank MJ, Pajtas PE, Santesso DL, Culhane M (2008) Single dose of a dopamine agonist impairs reinforcement learning in humans: behavioral evidence from a laboratory-based measure of reward responsiveness. Psychopharmacology (Berl) 196:221-232.

Reisman DS, Wityk R, Silver K, Bastian AJ (2009) Split-belt treadmill adaptation transfers to overground walking in persons poststroke. Neurorehabil Neural Repair 23:735-744.

Rossetti Y, Rode G, Pisella L, Farne A, Li L, Boisson D, Perenin MT (1998) Prism adaptation to a rightward optical deviation rehabilitates left hemispatial neglect. Nature 395:166-169.

Rutledge RB, Lazzaro SC, Lau B, Myers CE, Gluck MA, Glimcher PW (2009) Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. J Neurosci 29:15104-15114.

Salimpour Y, Shadmehr R (2014) Motor costs and the coordination of the two arms. J Neurosci 34:1806-1818.

Samejima K, Ueda Y, Doya K, Kimura M (2005) Representation of action-specific reward values in the striatum. Science 310:1337-1340.

Scheidt RA, Reinkensmeyer DJ, Conditt MA, Rymer WZ, Mussa-Ivaldi FA (2000) Persistence of motor adaptation during constrained, multi-joint, arm movements. J Neurophysiol 84:853-862.

Schiller D, Monfils MH, Raio CM, Johnson DC, Ledoux JE, Phelps EA (2010) Preventing the return of fear in humans using reconsolidation update mechanisms. Nature 463:49-53.

Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275:1593-1599.

Shadmehr R (2004) Generalization as a behavioral window to the neural mechanisms of learning internal models. Hum Mov Sci 23:543-568.

Shadmehr R, Brashers-Krug T (1997) Functional stages in the formation of human long-term motor memory. J Neurosci 17:409-419.

Shadmehr R, Holcomb HH (1997) Neural correlates of motor memory consolidation. Science 277:821-825.

Shadmehr R, Krakauer JW (2008) A computational neuroanatomy for motor control. Exp Brain Res 185:359-381.

Shadmehr R, Mussa-Ivaldi FA (1994a) Adaptive representation of dynamics during learning of a motor task. J Neurosci 14:3208-3224.

Shadmehr R, Mussa-Ivaldi FA (1994b) Adaptive representation of dynamics during learning of a motor task. J Neurosci 14:3208-3224.

Shohamy D, Myers CE, Geghman KD, Sage J, Gluck MA (2006) L-dopa impairs learning, but spares generalization, in Parkinson's disease. Neuropsychologia 44:774-784.

Shohamy D, Myers CE, Onlaor S, Gluck MA (2004) Role of the basal ganglia in category learning: how do patients with Parkinson's disease learn? Behav Neurosci 118:676-686.

Sing GC, Smith MA (2010) Reduction in learning rates associated with anterograde interference results from interactions between different timescales in motor adaptation. PLoS Comput Biol 6.

Smith MA, Ghazizadeh A, Shadmehr R (2006) Interacting adaptive processes with different timescales underlie short-term motor learning. PLoS Biol 4:e179.

Stollhoff N, Menzel R, Eisenhardt D (2005) Spontaneous recovery from extinction depends on the reconsolidation of the acquisition memory in an appetitive learning paradigm in the honeybee (Apis mellifera). J Neurosci 25:4485-4492.

Synofzik M, Lindner A, Thier P (2008) The cerebellum updates predictions about the visual consequences of one's behavior. Curr Biol 18:814-818.

Tai LH, Lee AM, Benavidez N, Bonci A, Wilbrecht L (2012) Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. Nat Neurosci 15:1281-1289.

Takikawa Y, Kawagoe R, Itoh H, Nakahara H, Hikosaka O (2002) Modulation of saccadic eye movements by predicted reward outcome. Exp Brain Res 142:284-291.

Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S (2004) Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. Nat Neurosci 7:887-893.

Thoroughman KA, Shadmehr R (1999) Electromyographic correlates of learning an internal model of reaching movements. J Neurosci 19:8573-8588.

Tononi G, Cirelli C (2006) Sleep function and synaptic homeostasis. Sleep Med Rev 10:49-62.

Torres-Oviedo G, Bastian AJ (2010) Seeing is believing: effects of visual contextual cues on learning and transfer of locomotor adaptation. J Neurosci 30:17015-17022.

Tumer EC, Brainard MS (2007) Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. Nature 450:1240-1244.

Uno Y, Kawato M, Suzuki R (1989) Formation and control of optimal trajectory in human multijoint arm movement. Minimum torque-change model. Biol Cybern 61:89-101.

Vakil E, Hassin-Baer S, Karni A (2014) A deficit in optimizing task solution but robust and well-retained speed and accuracy gains in complex skill acquisition in Parkinsons disease: multi-session training on the Tower of Hanoi Puzzle. Neuropsychologia 57:12-19.

Vaswani PA, Shadmehr R (2013) Decay of motor memories in the absence of error. J Neurosci 33:7700-7709.

Verstynen T, Sabes PN (2011) Adaptive priors emerge from associative learning in motor planning networks. J Neurosci in press.

Wada Y, Kawabata Y, Kotosaka S, Yamamoto K, Kitazawa S, Kawato M (2003) Acquisition and contextual switching of multiple internal models for different viscous force fields. Neurosci Res 46:319-331.

Walker MP, Brakefield T, Hobson JA, Stickgold R (2003) Dissociable stages of human memory consolidation and reconsolidation. Nature 425:616-620.

Wang W, Dounskaia N (2012) Load emphasizes muscle effort minimization during selection of arm movement direction. J Neuroeng Rehabil 9:70.

Wolpert DM, Kawato M (1998) Multiple paired forward and inverse models for motor control. Neural Networks 11:1317-1329.

Wu HG, Miyamoto YR, Gonzalez Castro LN, Olveczky BP, Smith MA (2014) Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. Nat Neurosci 17:312-321.

Xu T, Yu X, Perlik AJ, Tobin WF, Zweig JA, Tennant K, Jones T, Zuo Y (2009) Rapid formation and selective stabilization of synapses for enduring motor memories. Nature 462:915-919.

Xue YX, Luo YX, Wu P, Shi HS, Xue LF, Chen C, Zhu WL, Ding ZB, Bao YP, Shi J, Epstein DH, Shaham Y, Lu L (2012) A memory retrieval-extinction procedure to prevent drug craving and relapse. Science 336:241-245.

Yang G, Pan F, Gan WB (2009) Stably maintained dendritic spines are associated with lifelong memories. Nature 462:920-924.

Zarahn E, Weston GD, Liang J, Mazzoni P, Krakauer JW (2008) Explaining savings for visuomotor adaptation: linear time-invariant state-space models are not sufficient. J Neurophysiol 100:2537-2548.

Curriculum Vitae

SARAH E. PEKNY

Sep205@gmail.com

EDUCATION

2009-2015	Johns Hopkins University, School of Medicine Candidate for Ph.D. in Biomedical Engineering, March 2015
2005-2009	The Pennsylvania State University, School of Engineering B.S. in Bioengineering

PEER-REVIEWED PUBLICATIONS

SE Pekny, J Izawa, and R Shadmehr (2015) Reward dependent modulation of movement variability. Journal of Neuroscience (In Press)

SE Pekny and R Shadmehr (2015) Optimizing effort: time away from practice increases efficiency of motor memory. Journal of Neurophysiology 113(2):445-54.

J Izawa, **SE Pekny**, MK Marko, C Haswell, R Shadmehr, and SH Mostofsky (2012) Motor learning relies on integrated sensory inputs in ADHD, but over-selectively on proprioception in Autism spectrum conditions. Autism Research 5:124-136.

SE Pekny, SE Criscimagna-Hemminger, and R Shadmehr (2011) Protection and expression of human motor memories. Journal of Neuroscience 31:13829-13839.

JJ Orban de Xivry, MK Marko, **SE Pekny**, D Pastor, J Izawa, P Celnik, and R Shadmehr (2011) Stimulation of the human motor cortex alters generalization patterns of motor learning. Journal of Neuroscience 31:7102-7110.

ABSTRACTS & RESEARCH PRESENTATIONS

SE Pekny and R Shadmehr (2014). Rebuilding of motor memory: Increased efficiency at recall after time away from practice. Society for Neuroscience Meeting, Washington, DC

SE Pekny and R Shadmehr (2012) Long-term practice alters generalization properties of motor. Society for Neuroscience Meeting, New Orleans, LA

MD Harran, **SE Pekny**, and R Shadmehr (2012) Cerebellar degeneration alters the generalization of motor memories. Society for Neuroscience Meeting, New Orleans, LA

AM Haith, **SE Pekny**, R Shadmehr, JW Krakauer (2011) Evidence for model-free learning during force field adaptation. Society for Neuroscience Meeting, Washington, DC

SE Pekny, SE Criscimagna-Hemminger and R Shadmehr (2011) Reward prediction error produces spontaneous recovery of a motor memory. Society for Neuroscience Meeting, Washington, DC

MD Harran, **SE Pekny**, AM Haith, and R Shadmehr (2011) The nature of motor memory in people with a damaged cerebellum. Society for Neuroscience Meeting, Washington, DC

SE Pekny, SE Criscimagna-Hemminger, and R Shadmehr (2010) Disengagement and protection of motor memories. Society for Neuroscience meeting, San Diego, CA

JJ Orban de Xivry, **SE Pekny**, MK Marko, D Pastor, J Izawa, P Celnik, and R Shadmehr (2010) Generalization patterns of motor learning depend on M1 plasticity. Society for Neuroscience meeting, San Diego, CA

SE Pekny, KJ Paralikar, RS Clement(2008) Understanding the Role of Electrode Impedance Variability in Impacting the Chronic Recording Ability of the Device. Biomedical Engineering Society Annual Fall Meeting. St. Louis, MO

JT Matsui, KJ Paralikar, **SE Pekny**, M Polons, AJ Barber, RJ Clement (2007) Exploring Links in Recording Performance, Impedance and Histology in Chronic Intracortical Implants. Biomedical Engineering Society Annual Fall Meeting. Los Angeles, CA

TEACHING ASSISTANT POSITIONS

2013 Learning Theory, Professor Reza Shadmehr
2013 Systems Bioengineering II: Neuroscience for Engineers, Professor Xiaoqin Wang