Enhancing Motor Learning through Theoretical and Experimental Methods

Acquisition, Retention, and Learning Strategies

By
Vincent Sheng-Wen Huang

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

Baltimore, Maryland
December, 2007

Abstract

The first goal of this thesis is to apply theoretical and experimental methods to investigate the mechanisms of human motor learning. By comparing behavioral results with theoretical predictions, the second goal is to identify components of learning that can be further improved. Specifically, the acquisition, retention, and self-directed strategies of motor learning were studied.

Previous learning studies have shown that spacing learning sessions in time increases the efficacy of verbal skill acquisition. I showed that an increased rate of acquisition in motor adaptation was also observed when the temporal separation of training trials was increased by a few seconds. I hypothesized that each instance of training movement spawns a memory trace of the experience from which the learner continues to sample until another movement was made. I showed that indeed the learner's sensitivity to error (trial-to-trial learning rate) increases with the rest period after a movement with a time constant of about 4 seconds. An alternative theory hypothesized that time passage increases the learner's uncertainty about the task, and the trial-to-trial learning rate should increase with the uncertainty. The fact that the trial-to-trial learning rate was a monotonically increasing function of *post*-movement time contradicted the pre-movement time dependency as predicted the uncertainty-based model.

In learning a motor task, a learner does not only need to estimate the optimal parameters for that task, but also learn the structure of the task itself. However, most motor learning literatures do not vary the task structure. Secondly, learning task structure often takes place at the same time as task parameter estimation. I introduced a novel paradigm that varied the task structure: adaptation to a gradually or a suddenly introduced dynamical perturbation of the same magnitude. By measuring the passive decay of adaptation in naïve and experienced learners, the manifestation of structural learning independent of parameter estimation could be observed. I showed that gradual training promoted retention of adaptation in naïve but not necessarily experienced learners. However, the gradualness alone was sufficient to promote retention in re-training sessions suggesting that a prior of the task structure was acquired at the first exposure and continued to influence people's motor adaptation and retention properties in subsequent learning sessions.

In real life, we are often faced with the problem of having to learn by ourselves. I explored the interactive component of motor learning and the optimality of the strategy people use in selecting the sequence of movements in motor adaptation. I found that the strategy was stereotypical. While people benefit from repeating movements that were previously underperformed, they suffered from a suboptimal strategy to also repeat movements that were already well-performed previously. We showed that this latter result violated the theoretical predictions of an optimal, unbiased active learner who constructs a strategy based on his uncertainty about a task component. Therefore, it may be possible to supplement and enhance human motor learning strategies by a designed, machine-based "coach."

Advisor: Dr. Reza Shadmehr

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Computational Learning Theories and Motor Learning

While it is generally true that practice makes perfect, how we practice, influences the efficiency and level of success in learning to move better. The elucidation of the neural basis of learning, however, remains one of the biggest challenges in neuroscience. The recent advancement in computational neuroscience, fortunately, offers a systematic approach to gain an understanding of our motor learning mechanisms through mathematical models and theoretical predictions. This new knowledge comes with an opportunity to test the theories that could potentially enhance motor learning.

Better motility is a matter of survival. Darwin's theory of evolution tells us that a species adapts to changes of the environment to gain an advantage in survival. Life has evolved to give rise to various ways to cope with the dynamic challenges posed to it. The ability to adapt is the primary tool in the game of survival. One of these adaptations is the development of organs such as the eye and the limb, to passively sense and actively manipulate the environments. In order to connect, interpret and control these interactions, the nervous system was evolved. Together, they allowed our ancestors to be better at survival – so well that their actions seem second-nature. Take an arm reaching for an object, for example. We can unimpressively aim and extend our hand for a cup of coffee, seemingly without too much thought dedication. Moving skillfully – as simple as reaching for an object -- however, requires monumental undertaking of the nervous system. When engineers and computational neuroscientist attempted to simulate biologically inspired robotic arms, they realized that the nervous system must face a plethora of problems. First, the brain must transform sensory information to a motor plan, therefore requiring accurate transformations of co-ordinate systems that are too computationally complex. Second, conduction speed of the biological nervous system is slow, which inherently introduces nonlinearity and instability to the motor system that is loaded with inertial dynamics. Lastly, the environment can change and unpredictable disturbances take place. It's impossible to evolve specialized organs fast enough to deal with minute-to-minute changes in the environment. Therefore, the control program, the nervous system, like life itself, not only needs to adapt as well, it needs to adapt in a timely fashion.

To explain how the nervous system overcomes these difficulties, computational neuroscientists posit that the nervous system, like a computer controlling a robotic skeleton, can theoretically employ several control strategies. One of these strategies is feedback control. The nervous system can simplify the complexity of co-ordinate transformation by only requiring the motor system to minimize the difference between the feedback signal from the proprioceptive and visual sensory system and a pre-defined desired motor plan. An early model that adopted this idea was the equilibrium point hypothesis which postulated that a limb acts like a spring moving toward a final, desired position known as the equilibrium point (Feldman, 1986; Shadmehr and Arbib, 1992; Bizzi, 1993; Giszter et al., 1993). A second strategy is that, since the feedback signal comes after a delay, the nervous system must be able to generate predictions (feed-forward) of the skeletal dynamics based on lagged sensory feedback (Miall, 1995). Being able to generate predictive outputs of dynamics implies that the nervous system must either have a lookup table or an inverse model of the skeletal dynamics (like Newtonian equations of mechanics) with lagged sensory information as input. The combination of inverse and feed-forward controls constitutes the concept of internal model in motor control, in that the nervous system develops a model of expected dynamics of the limb and environment. We now know that the equilibrium point hypothesis is not sufficient to explain motor behaviors because our movement trajectories predictably change depending on task demands (Shadmehr and Mussa-Ivaldi, 1994). The latest development in the control theory of motor system is the incorporation of optimal control(Todorov, 2004). In this control regime, the reference trajectory, or desired trajectory can be generated as a function that minimizes task cost whether it is effort, time, end point error or some other arbitrary variable such as financial incentives associated with the movement. The proposed anatomical sites of motor plan generation of the limb in the nervous system also moved from the spinal cord level as suggested by the equilibrium point hypothesis to higher levels in the central nervous system at the sensory-motor cortex, parietal cortex, basal ganglia, and cerebellum(Nezafat et al., 2001; Diedrichsen et al., 2005).

While these computational theories helped to explain motor control, additional treatments was needed to explain the nervous system's ability to adapt to surprise changes in the kinematics or dynamics of the environment or tool. There has been a great deal of interests invested in how the motor system adapts as revealed by psychophysical experiments involving repetitive movements. Our current understanding of motor learning has benefited from studies of repetitive movements in perturbation-adaptation tasks. In these tasks, the subjects are asked to compensate for a kinematical or dynamical perturbation during movements to move as normally as possible (Shadmehr and Mussa-Ivaldi, 1994; Krakauer et al., 2005; Michel et al., 2007).

Interestingly, learning of the perturbation-adaptation task has features that resemble the physiology of the nervous system. Spatial generalization in joint angle co-ordinates was observed when trained subject moved in a new region of the workspace(Gandolfo et al., 1996). Kinematics errors in one trial affected differentially subsequent movements. The trial-to-trial generalization was the greatest when the subsequent movement was in the same direction, and smallest when it was 90 degrees away. The pattern of spatial generation resembles the tuning curves of some Purkinje cells in the cerebellum (Coltz et al., 1999, 2000). Spatial generalization has been shown to be strong when perturbation is given according to movement speed, and less so when perturbation is given according to the workspace region(Hwang et al., 2003). Therefore, the progressive adaptation of internal model appears to be most sensitive to movement velocity. Biologically inspired model have been developed to demonstrate the feasibility of reproducing the behavioral learning pattern using physiological signals from the proprioceptive apparatus – mostly sensitive to muscle stretch speed -- as inputs(Hwang and Shadmehr, 2005; Hwang et al., 2006b).

1.2 Spaced versus Massed Training

As a result of the perturbation-adaptation studies, computational neuroscientists made strides in advancing the mathematical tools to help us explain motor learning. It is currently thought that an internal expectation (internal model) of the perturbation is formed based on which compensatory motor commands are computed and this internal model is updated when an error is observed. As people repeat these movements through training, they update their internal model on a trial-to-trial basis to better approximate the perturbation and aggressively achieve better compensation. Therefore, it should be possible to measure learning on a trial-to-trial basis, far before adaptation has reached its steady-state (Donchin et al., 2003).

These models of motor adaptation treat sequential motor adaptation trials as discrete learning events taking place in a time-invariant setting. It is not clear whether motor learning is a discrete process because most motor control studies do not vary these temporal variables. In fact, an important property of memory is that learning depends not just on the number, but also the temporal distribution of training trials. This property was first reported by Ebbinghaus on the study of memory(Ebbinghaus, 1964). In his study, he described a training protocol known as *spaced training* in which trials are separated by periods of rest as opposed to a traditional protocol in which trials are given without rests (*massed training*). Counter-intuitively, Ebbinghaus found that he could cut the number of repetitions required to memorize a list of words in half if he spaced the repetitions in time. Spaced training also results in improved performance in other memory tasks. In the motor tasks, trials that were spaced in time produced faster adaptation rates. This implies that at least in the tasks studied, motor error had a time-dependent influence on internal models.

Specifically the temporal structures of training are largely ignored because in these models, time-invariance is assumed in the control and adaptive scheme of the motor system. That is, the motor system always adapts the same way regardless of the distribution of training trials. By incorporating a dependency on the temporal variable in the current computational models as inspired by observations made in the behavioral learning literature, we can make new hypothesis how to increase learning based on a theoretical approach.

Chapter 2 describes the observations that the brain benefited from the passage of time such that a given motor error produced a larger amount of adaptation with increased inter-trial interval. A motor error produced a decaying error-trace from which the nervous system continued to sample and benefit was proposed and the effect of error experienced on one trial on the hidden state of the learner as it was expressed some seconds later in the subsequent trial was quantified. I found that the effect of error grows as a function of time. I estimated that with each trial, the brain forms an error trace that decays with a time constant of about 4 seconds. Lastly, an experiment to directly test the predictions of the error trace model was performed. Yet I still found that only the error trace model could account the empirical data. These results suggested that the brain continuously benefit from a movement experience in the period after the movement and before the next trial.

1.3 Learning the Structure of a Motor Task

Learning a motor task such as tightening a small screw with a hexagonal socket involves identification of the goal of the task, learning structure of the task, and estimation of the parameters that describe the structure of the task. In this example, tightening the hex screw is the goal and the hex screw can be tightened by two different tools (i.e. two task structures). One may use a regular screw driver creating rotational torques on the screw driver by generating a shear force on the handle with the opposing fingers. Alternatively, one may elect to use a hex wrench (Allen key) and therefore generating the same torques by moving the perpendicular arm of the wrench sideways. With each tool, one has to learn how much force to apply to tighten the screw securely (i.e. task parameter estimation).

In most motor learning literatures, however, the structure of the task is fixed (e.g. use only a screw driver or a hex wrench, but not both) and the goal of the task well defined (Shadmehr and Mussa-Ivaldi, 1994; Todorov et al., 1997; Lackner and Dizio, 1998; Krakauer et al., 1999; Scheidt et al., 2000; Thoroughman and Shadmehr, 2000; Donchin et al., 2003; Lackner and DiZio, 2003; Smith et al., 2006; Kording, 2007). Motor learning in these studies then reduces to an optimal estimation of the parameter values of the task. The aforementioned example, however, clearly shows that this invariability in task structure is not necessarily true in general, even when the goal (e.g. tightening the screw) and the task parameters are fixed (e.g. the necessary amount of torque). Yet we have an amazing ability to identify and adopt different task structures to accomplish various motor goals in our daily life. Even more amazing is our apparent ability to retain the knowledge once the task structure is learned and apply this knowledge when the same task is presented again. Little of this structural learning ability has been studied or characterized. In fact, current motor learning models are essentially variations of parameter estimating models that treat every potential task as having a time-invariant, well-defined, and a priori task structure.

Chapter 3 introduces a novel force adaptation paradigm to study the effect of variability in task structure on the retention of motor adaptation in a force adaptation task. Subjects were trained to counteract a perturbation force to make straight reaching movements. The task structure was varied, however, in that subjects were exposed to the perturbation either gradually over many trials or suddenly. As expected, experimental results showed that subjects responded differently to the two task structures in that their retained adaptation decayed passively at different rates after the perturbation was removed. In a subsequent retraining block, the gradualness of the first training alone was sufficient to promote higher retention of the adaptation regardless of the schedule of the retraining and despite of an extended intervening washout block intervening the two training blocks. Thus, I demonstrated that people indeed learns different task structures in an adaptation task and the knowledge of the task structure persists and influences how people react to further training of similar task structures.

1.4 Self-directed Strategy in Motor Learning

Another defining feature of human motor learning is that learning is most often accomplished autonomously. The autonomy requires an active role of the learner: the learner must initiate and direct his or her interactions with the task environment to seek out the things that help learning. The field of machine learning has realized this as important element of knowledge acquisition and has characterized optimal learning algorithms that seek out the most informative training examples at any point in time(Cohn et al., 1996). In contrast, the behavioral sciences have largely ignored this active aspect of learning. Most learning studies have a preset sequence of learning examples; that is, they treat the learner as passive systems.

In Chapter 4, I provided the first systematic and quantitative description of active learning and its influence on performance in biological motor control. I asked the questions what factors guide human active learning strategies and whether human active learning follows an uncertainty-based algorithm as proposed as an optimal strategy in machine learning. In a force adaptation task, I found that human strategies were strongly and robustly influenced by the last movement error. Learners tended to repeat a task component in which they experienced a large error, a strategy that improves learning performance. However, even when participants performed perfectly on a component, they did not avoid repeating that same component. This behavior violated a strong prediction of the uncertainty-based optimal machine learning algorithms, and was indeed suboptimal as it led to poorer learning outcomes. Thus these results demonstrated the limitation of human active learning. Given the importance of interaction and active learning in every-day life, the novel systematic description of factors driving active choice during the learning process would make an important contribution to our knowledge on human and animal learning. Furthermore, the comparison to machine learning algorithms indicates the possibility of designing machine-based “coaches” that adaptively chooses new training examples to improve human motor learning in instances such as rehabilitation training.

CHAPTER 2

THE EFFECTS OF INTER-TRIAL INTERVALS ON MOTOR LEARNING

2.1 INTRODUCTION

Learning depends not just on number of repeated exposures, but also the temporal distribution of the exposures. Consider the distinction between *massed* and *spaced* training first coined by Ebbinghaus on the study of memory (Ebbinghaus, 1964). In massed training, the trials take place in close temporal proximity. In spaced training, trials are separated by periods of rest. Ebbinghaus found that spacing training sets over time was more effective in allowing him to memorize a list of nonsense words to criterion with less practice than massing them in a single set. Spaced training also improves rates of learning in other tasks (Han et al., 1998; Commins et al., 2003; Aboukhalil et al., 2004; Savion-Lemieux and Penhune, 2005). For example, two recent studies examined this effect in the context of reach adaptation. Bock et al. (Bock et al., 2005) trained subjects to point in a novel visual feedback environment and observed that the rate of adaptation was faster if the trials were separated by 5 to 40s than by 1s. Similarly, Francis (Francis, 2005) noted that learning to control a novel tool (reaching while holding a robot arm in a force field) was faster if the trials were separated by 5-20s than 0.5s. These results are not accounted for by current computational models of motor learning where adaptation is driven only by motor error (Scheidt et al., 2001; Donchin et al., 2003; Wainscott et al., 2005) because these models treat the spaced and massed training paradigms identically. Rather, the results suggest that trial-to-trial effect of motor error depends on the inter-trial interval between the movements.

More recent computational models of motor learning suggest that motor error engages multiple adaptive processes of different timescales: one process strongly responds to error but has poor retention, and another has poor sensitivity to error but has better retention (Smith et al., 2006). This model and its Bayesian variant (Kording, 2007) emphasize that passage of time is an important variable that influences content of memory. Can such models help explain the massed vs. spaced training effects?

Here we begin with an experiment that confirms the previous findings that reach adaptation in force fields is indeed faster (i.e. require fewer trials) when the inter-trial interval (ITI) is increased from 4 to 14 seconds. We demonstrate that the multiple-timescale model under a Bayesian formulation can account for this result. Since time passage increases the Bayesian learner's uncertainty about his environment, and increased uncertainty promotes his incentive to adapt, longer ITIs lead to less number of trials required for adaptation. The model predicts that the learner will learn more from a trial that immediately *follows* a long delay than one that immediately follows a short delay.

To test this prediction, we performed another experiment where ITIs were randomly distributed. However, we found results inconsistent with the predictions of the Bayesian model. People learned more from a trial that *preceded* a long delay, not a trial that *followed* a long delay. Therefore, the improved adaptation rates in spaced trials were unlikely to be a result of increased uncertainty. Rather, our results suggest that movement errors produced a trace that continued to benefit the learner during the ITI. It appears that this trace has a time constant of about 4 seconds.

2.2 METHODS

We reconsidered a well-studied reach adaptation task (Shadmehr and Mussa-Ivaldi, 1994) and asked whether spaced training benefited rates of adaptation. Subjects held the handle of a two-joint, planar manipulandum equipped with torque motors, rotary encoders and force transducers and reached to visual targets. Subject's upper arm was supported by a sling restricting their movements to the same horizontal plane as the manipulandum. During "field trials," the torque motors of the manipulandum perturbed the movement of the subjects by a viscous curl-force field:

$$V = \begin{bmatrix} 0 & 13 \\ -13 & 0 \end{bmatrix} \quad ; \quad F = V\dot{x} \tag{2.1}$$

In Eq. (2.1), force F is in Newton, and hand velocity $\dot{x}$ is in m/s. No forces were applied during "catch trials" and trials in the null training set.

2.2.1 Behavioral Training

All procedures were approved by the Johns Hopkins Medicine Institutional Review Board. Subjects gave consent prior to their participation in the study. All subjects were healthy and right-handed. They were naïve to the purpose of the experiment, and had never participated in any experiment with our device before. Participants were seated in front of a flat-screen monitor situated at eye level (Figure 2.1). With their right hand they grasped the handle at the end of the manipulandum to navigate a cursor (a white dot) on the screen. Subjects were trained to maintain the cursor at the center of the screen as indicated by a yellow crosshair until a 1 cm2 green target box appeared which also served as the "go" signal. They were also told to "aim for the target as it appears and try making a straight point-to-point movement to the target in a fast and smooth fashion." The yellow crosshair vanished as soon as subject started to move.

The green target box could appear at any one of the eight locations spaced on an invisible circle of 10 cm radius centered at the crosshair. The sequence of the target locations was determined pseudo-randomly with all eight possible target locations visited with equal probability. After completion of the movement, the green target box turned magenta when the tangential peak speed exceeded 0.55 m/s or cyan when the speed faltered below 0.20 m/s respectively. If the movement duration surpassed 0.57 sec or fell under 0.43 sec, the box turned red or blue. If the movement profile met the above parameters, it was considered ideal and the target box exploded. Distinctive audio feedback was also given for magenta, blue and exploding target boxes. The manipulandum subsequently returned the subject’s hand to the origin at the center of the screen.

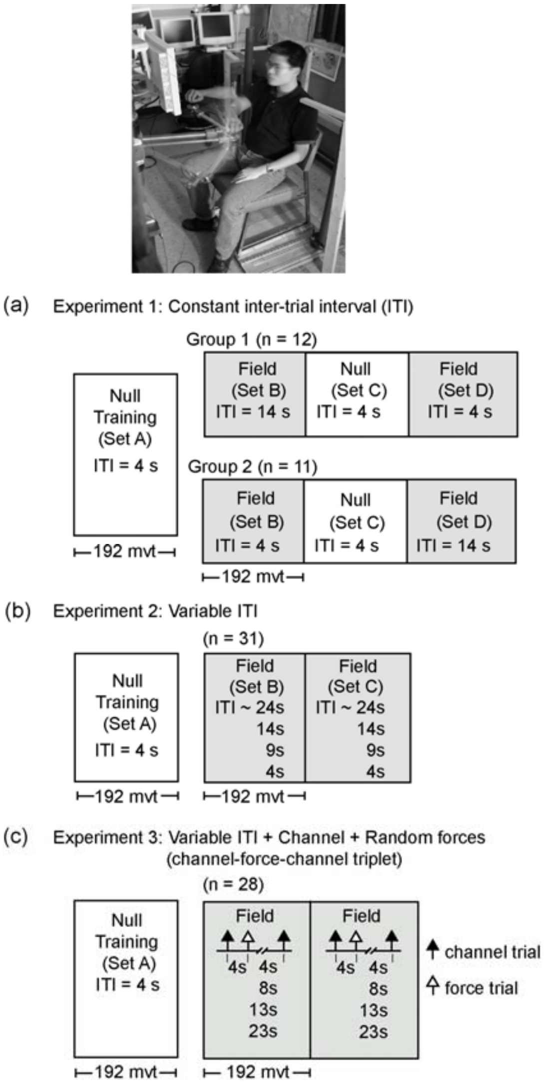

Figure 2.1 Experiment design. **a.** Experiment 1 protocol. Inter-trial intervals (ITI) were constant within a movement set (Sets A, B, C, D). Shaded blocks indicate that the force field was on for all but catch movement trials in that set. **b.** Experiment 2 protocol. ITIs were randomly varied during field training sets. **c.** Experiment 3 protocol. Field trials (solid triangle) were sandwiched between pairs of channel trials (open triangle) and we varied the time between the field trial and the subsequent channel trial.

2.2.2 Performance Measure

Movement errors were measured as the signed perpendicular displacements (PD) of the reach at peak speed with respect to a straight line to the target. We grouped 32 movements into one movement bin and averaged the errors in field and catch trials to arrive at $\overline{PD}_{field}$ and $\overline{PD}_{catch}$. Next, we computed a learning index (LI) (Criscimagna-Hemminger et al., 2003; Hwang et al., 2003):

$$LI = \frac{\overline{PD}_{catch}}{\overline{PD}_{catch} - \overline{PD}_{field}} \tag{2.2}$$

The denominator of this expression is a measure of limb compliance; smaller stiffness gives rise to larger differences in errors between catch and field trials. The numerator is a measure of change in motor output with respect to null trials. Therefore, the ratio is a measure of learning, normalized with respect to limb compliance. As subjects learn to predict the forces, their movement errors in field trials decrease while errors in catch trials increase. Thus, we would expect the learning index to grow and plateau as the subjects adapt. Complete adaptation would result in a learning index of one. However, catch trials cause unlearning (Thoroughman and Shadmehr, 2000) and prevent the index from reaching one. The ratio of field trials to catch trials (i.e. 5/6) dictates the theoretical limit of learning index, which is 0.83. In our experience, the highest actual learning performance is slightly below this number (Criscimagna-Hemminger et al., 2003). Movements were excluded if they did not meet the following criteria: maximal tangential speed was between 0.20 and 0.55 m/s, movement duration was between 0.3 and 1.2 s, and the total movement trajectory length was less than 20 cm.

2.2.3 Experiment 1: constant ITI

In this experiment we sought to replicate the results of Francis (Francis, 2005) and Bock et al. (Bock et al., 2005). The experiment was divided into four sets (Figure 2.1a), each containing 192 reach trials: set A (null baseline), set B (field), set C (null washout), and set D (field). In sets B and D, the field was randomly removed in 1/6th of the trials (catch trials). Short breaks of five minutes were given between sets. Subjects (n=24) were randomly assigned to two counter-balanced groups. One individual did not follow instructions and was excluded from data analysis. Both groups performed the same sequence of movements in each of the four sets with one difference: in set B of the first group, after the hand returned to the center location the target presentation was delayed by 10s while target presentation in all other sets was delayed by only 0.5s (sets A, C, D).

In the second group, target presentation was delayed by 10s in set D and by 0.5s in all other sets (sets A, B, C). We defined ITI to be the time between the onsets of consecutive reaching trials and it included all delays and movement time. Thus, the presentation delays resulted in mean ITI of either 4 or 14s.

Because we found that longer ITI produced significantly faster rates of adaptation, we considered two general models that might account for the data.

2.2.4 Model 1 (Bayesian multi-rate model): Motor adaptation as optimal inference at multiple rates

One way to account for time-dependent changes in motor performance is to envision that the learner is a Bayesian estimator that assumes that motor error arises from multiple causes: some perturbations go away quickly but tend to be highly variable (fast system, e.g., fatigue), while other perturbations tend to go away slowly and tend to be less variable (slow system e.g., disease). We recently formalized this idea and demonstrated that a multiple-system model can account for a large body of data in saccade and reach adaptation (Kording, 2007). The principal idea in this model is that movement error results in a credit assignment problem for the nervous system. To solve a credit assignment problem, we need to determine how to vary the contributions of the two systems to a common task: what is the timescale of perturbation that is most likely responsible for the current error? Is the perturbation likely to go away quickly, or is it likely to be sustained? If it is likely to be sustained -- as with spaced perturbations -- the learner should increase his error sensitivity for slower timescales. To show how such a model will be affected by ITI, suppose that the learner assumes that the perturbations (for example, force f imposed on the limb) are caused by a linear combination of two sources, each with its own states, timescale and noise properties:

$$
\begin{aligned}
f^{(n)} &= z_1^{(n)} + z_2^{(n)} + \varepsilon_0^{(n)} & \varepsilon_0 &\in N\left(0,\sigma_0^2\right) \\
z_1^{(n)} &= a_1 z_1^{(n-1)} + \varepsilon_1^{(n-1)} & \varepsilon_1 &\in N\left(0,\sigma_1^2\right) \\
z_2^{(n)} &= a_2 z_2^{(n-1)} + \varepsilon_2^{(n-1)} & \varepsilon_2 &\in N\left(0,\sigma_2^2\right)
\end{aligned}
\tag{2.3}
$$

In the above equations at each iteration *n*, the learner's state (represented by variable *z*) is a reflection of two underlying sources: one system (represented by state z_1) for the fast perturbations and another (represented by state z_2) for the slow perturbations. ε_0 is the noise in our sensors that measure the perturbation, and ε_1 and ε_2 are noises associated with the fast and slow perturbations (we assume that $a_1 < a_2 < 1$ and $\sigma_1 > \sigma_2$). The learner's "knowledge" on each trial is a sum of contribution of the fast and slow systems with associated noises. It is convenient to rewrite Eq. (2.3) in vector format as:

$$
\begin{aligned}
f^{(n)} &= \mathbf{x}^T \mathbf{z}^{(n)} + \varepsilon_0^{(n)} & \varepsilon_0 &\in N\left(0,\sigma_0^2\right) \\
\mathbf{z}^{(n)} &= A\mathbf{z}^{(n-1)} + \boldsymbol{\varepsilon}^{(n-1)} & \boldsymbol{\varepsilon} &\in N\left(0,Q\right)
\end{aligned}
\tag{2.4}
$$

In Eq. (2.4), $\mathbf{x}^T = \begin{bmatrix}1 & 1\end{bmatrix}$ and *A* and *Q* are diagonal matrices with components described in Eq. (2.3). On trial *n*, given that the learner has observed the last $n-1$ trials, it will have a prior estimate $\hat{\mathbf{z}}^{(n|n-1)}$ and a predicted perturbation $\hat{f}^{(n)} = \mathbf{x}^T \hat{\mathbf{z}}^{(n|n-1)}$. The optimal way that it can distribute the error $f^{(n)} - \hat{f}^{(n)}$ to each of the two potential sources is described by the gain $\mathbf{k}^{(n)}$ (the Kalman gain). The mean of the posterior estimate will become:

$$
\begin{aligned}
\hat{\mathbf{z}}^{(n|n)} &= \hat{\mathbf{z}}^{(n|n-1)} + \mathbf{k}^{(n)}\left(f^{(n)} - \mathbf{x}^T \hat{\mathbf{z}}^{(n|n-1)}\right) \\
\mathbf{k}^{(n)} &= \frac{P^{(n|n-1)}\mathbf{x}}{\mathbf{x}^T P^{(n|n-1)}\mathbf{x} + \sigma_0^2}
\end{aligned}
\tag{2.5}
$$

In Eq. (2.5), $P^{(n|n-1)}$ is the prior uncertainty matrix, describing the variance covariance of each component of **z**.

Now if during an iteration of the model, the learner makes a movement, it will make an observation, and therefore the posterior uncertainties will change:

$$P^{(n|n)} = \left(I - \mathbf{k}^{(n)}\mathbf{x}^T\right)P^{(n|n-1)} \tag{2.6}$$

During model iterations where the learner is not allowed to make observations, the posterior estimate will not change:

$$\begin{aligned} \hat{\mathbf{z}}^{(n|n)} &= \hat{\mathbf{z}}^{(n|n-1)} \\ P^{(n|n)} &= P^{(n|n-1)} \end{aligned} \tag{2.7}$$

However, regardless of whether the learner makes an observation or not, at the next iteration the prior uncertainty will change:

$$P^{(n+1|n)} = AP^{(n|n)}A^T + Q \tag{2.8}$$

We see that uncertainty decreases when an observation is made (Eq. (2.6)), but can potentially increase between iterations (Eq. (2.8)) due to Q. We assume that the learner updates the uncertainty with each iteration. Assuming that each iteration takes a constant amount of time, the longer ITIs allow more iterations than shorter ITIs. Consequently, longer ITIs (i.e. after more iterations) can produce increased uncertainty. This in turn produces higher sensitivity to error (Eq. (2.5)), which can result in an increased rate of trial-to-trial adaptation. As a consequence, this model predicts that longer ITIs have the potential to produce faster rates of adaptation.

2.2.5 Model 2 (Error trace model): Adaptation as continuous integration of motor error

Model 1 assumes that adaptation in response to error is an instantaneous process that completes by the next iteration. A different way to view adaptation is to imagine that it is a process that is initiated with the experience of error, but continues as long as the error memory trace is available. Let us assume that the error trace is an exponentially decaying function of the form $\mathbf{y}^{(n)}\frac{1}{r}\exp\left(-\tau/r\right)$ where $\mathbf{y}^{(n)}$ is the error experienced at trial n at time t_{n+1}, τ is the time variable and r is the time constant for the temporal decay. Suppose that the learner continuously learns from the error trace that was initiated at the time of trial n until the new error trace at trial $n+1$ interrupts this process, the learner's state at t_{n+1} (immediately before experiencing the error in trail $n+1$) becomes:

$$\mathbf{z}(t_{n+1})=\mathbf{z}(t_n)+\mathbf{y}^{(n)}\int_{t_n}^{t_{n+1}}\frac{1}{r}\exp\left(\frac{t_n-\tau}{r}\right)d\tau \tag{2.9}$$

The integral in Eq. (2.9) can be simplified:

$$\mathbf{z}(t_{n+1})=\mathbf{z}(t_n)+\mathbf{y}^{(n)}\left(1-\exp\left(\frac{t_n-t_{n+1}}{r}\right)\right) \tag{2.10}$$

In this model of adaptation, a movement is like a point process that resets the previous error trace and replaces it with the most recently acquired sample. As ITIs increase, the contribution of error, i.e., the exponential term in Eq. (2.10), increases. As a result, the system learns more from an error when the trials are spaced in time.

2.2.6 Predictions of the two models

Models 1 and 2 both predict that longer ITIs will affect rates of adaptation, but their mechanisms are different. In Model 1, when a trial is followed by a long delay, parameter uncertainties can increase, which in turn increases the sensitivity to error in the trial that *follows* the long delay. In other words, longer ITIs signal the learner to pay attention. In Model 2, when a trial is followed by a long delay, the error that was experienced in the preceding trial is integrated over a longer time interval. In other words, longer ITIs allow more samplings from the error. This increases the sensitivity to error in the trial that *preceded* the long delay. Therefore, the key experiment is one that measures the learner's sensitivity to error in trial *n*, as a function of the time that either preceded or followed that trial.

2.2.7 Experiment 2: variable ITI

This experiment (Figure 2.1b) consisted of three sets (each with 192 trials) separated by short breaks: set A (null), set B (field), and set C (field). During the null training set, presentations of targets were delayed by 0.5 seconds as in Experiment 1. During the field sets, 1/6 of the trials were catch trials. However, unlike Experiment 1, the delay before target presentation was pseudo-randomly selected to be 0.5, 5, 10, or 20s. An equal number of different delays for a given movement were presented in each set.. Since an ITI also included the time for movements (that were remained relatively constant for each participant), imposing this 'go' signal delay resulted in mean ITIs of 4, 9, 14, and 24s.

We recruited a new group of naive subjects for this experiment (n=31). To assess sensitivity to error and the influence of ITI, we fitted the trial-by-trial movement errors to a state space model (Wainscott et al., 2005). The state space model included a hidden state for each of the 8 directions of movement. We estimated a trial-to-trial generalization function from the direction in which error was experienced to all possible directions. This generalization function was a measure of sensitivity to error. The models predicted that changing the ITI would affect this sensitivity. In particular, Model 1 predicted that the sensitivity to error experienced in trial *n* should increase as a function of the delay that preceded that trial. Model 2 predicted that the sensitivity to error experienced in trial *n* should increase as a function of the delay that followed that trial.

2.2.8 Experiment 3: variable ITI with channel trials and random forces

The results of Experiment 2 were consistent with predictions of Model 2 but not Model 1. To test the assumptions of Model 2 more directly, we performed a final experiment. In this experiment, we measured state of the motor system directly through the use of "channel trials" (Scheidt et al., 2000; Hwang et al., 2006b; Smith et al., 2006). In a channel trial, the robot restricts the hand's motion along a straight line to the target. While it prevents errors during the reach, it allows us to measure how much force the subject expected to experience for that trial. This expectation is equal to the force that the subject produces against the channel wall. Previous work had found that during reach adaptation, these forces gradually approximate the force field that the robot produces during free movements (Hwang et al. 2006). Our idea here was to use channel trials to measure the change in the expected force as a function of error in the previous reach trial.

Experiment 3 (Figure 2.1c) consisted of three sets (each 192 trials), separated by short breaks: set A (null baseline), set B (random field), and set C (random field). We recruited a new group of naïve subjects (n=28). Unlike Experiments 1 and 2, however, there were only two targets: either up or down with respect to the center position. In the field sets subject were given, in a pseudo-random order, a clockwise curl-force field (Eq. (2.1), +**V**, 3/8 of the movements), a counter-clockwise curl-force field (-**V**, 3/8 of the movements), or a "channel trial" in which movements perpendicular to the target direction were prevented by a stiff one-dimensional spring/damper (2 kN/m, 45 Ns/m). The seamless production of the channel force was based on the hand position in the center start box and unperceivable to participants unless they purposely tried to move in the perpendicular direction. Such movements were not observed, and participants reported that they were not aware of this force pattern at all.

We were interested in the change in the force exerted by the subject against the channel walls when two channel trials were separated by a trial in which there was a movement error. There were 34 such channel-force-channel trial triplets dispersed pseudo-randomly in set B and 35 in set C. In the triplet, between the first and the second movement, the ITI was kept at a constant 4 s. Between the second and the third movement, we varied the ITI in a random fashion identical to the Experiment 2 design. Thus, for each triplet we looked at the difference in the force output in the first and third movements (both were channel trials) as a function of the time passed since the error experienced in the second movement. Model 2 predicts that when one experiences a reach error in trial n, on trial n+1 one will produce a force against the channel walls that is proportional to this error, and this sensitivity will grow as a function of the time between trial n and n+1.

2.2.9 Bootstrap methods for estimating the confidence limits

We followed the procedures described in our previous publications (Donchin et al., 2003) which followed procedures laid out by Efton and Tibshirani (Efron and Tibshirani, 1993) to estimate the standard errors (Figure 2.4) in the model parameters in the analysis of experiment 2. We re-sampled data from the 31 subjects with replacement, and estimated the parameters from the averaged sample. We iterated this procedure 200 times, and the standard deviation of the 200 estimations of the parameters yielded an estimation of the standard error of the parameters.

2.3 RESULTS

2.3.1 Experiment 1: spaced training resulted in faster adaptation

In experiment 1, we sought to replicate the results of Francis (Francis, 2005) and Bock et al (Bock et al., 2005). We quantified the effect of massed versus spaced training in a within-subject design (Figure 2.1a). Subjects were randomly assigned to train with either a long (13.60 ± 0.13s, mean±SD) or short delay (3.88±0.14s) before each trial. We found that they showed better adaptation in the longer ITI set – compensation in field trials was stronger and after-effects in catch trials were larger (Figure 2.2a-c). We measured performance with a learning index in each of the six 32-movement bins in each set. Repeated-measure three-factor ANOVA analysis (ITI, set order, and movement bin) on the learning index showed no significant effect due to set order ($F(1,21)=0.020$, $p=0.89$). We therefore combined data from the two groups and considered the effect of ITI on the learning (Figure 2.2d). We found a significant interaction effect between ITI and movement bin ($F(5, 105)=4.07$, $p=0.002$). Post-hoc analysis showed significantly better performance in bins 2, 3, and 4 within subjects (paired t-test, $p<0.05$ for each bin) for the long ITI set. Therefore, adaptation rate was enhanced with the longer ITI.

We checked whether the gains in performance with longer ITI might have been due to a fatigue-like process. In our task, forces that counter the perturbing field are about one third of the forces that move the arm toward the target (Bhushan and Shadmehr, 1999). Therefore, if the limb fatigues with short ITIs, forces that move the limb toward the target should show a positive correlation with respect to the time passed since last trial. As a proxy for this force, we looked at the magnitude of peak velocity vectors parallel to the direction of the target – a measure highly correlated with fatigue (de Haan et al., 1989; Jaric et al., 1997). If the limb fatigues with smaller ITIs, movements should become slower with smaller ITIs. Peak velocity was not different in the two ITI groups (two-tailed t-test, t = -0.22, D.F. = 44, p = 0.83). In addition to the velocity measure, we would expect the fatigued group to show a lower performance due to inability to express learning. However, the two groups attained similar performance toward the end of training session (Figure 2.2d). Together these results suggested that differences in performance were unlikely due to a potential for fatigue in the short ITI group.

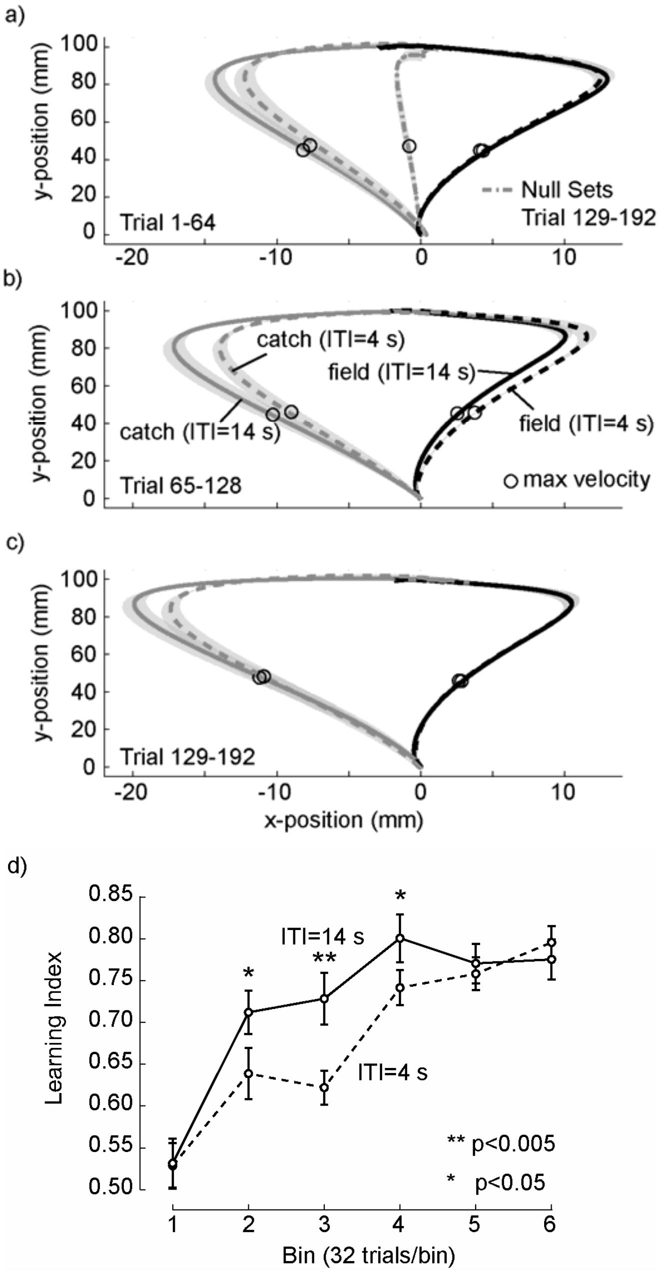

Figure 2.2 Mean movement paths averaged across subjects, directions, and training sets in Experiment 1. Shadings indicate the standard error of the mean trajectory averaged across subjects. **a**. Mean movement paths in the early stage of training (field and catch trials). **b.** Mean movement paths in the middle stage. **c**. Mean movement paths in the late stage. **d**. Comparison of learning performance between short and long ITI movement sets. Each sequence of movement data was grouped into 32-movement bins and a learning index was computed for each bin. Learning indices were averaged across subjects and sets of the same ITI. Within-subject performances were significantly different for bins 2, 3, and 4 (paired t-test, $p<0.05$ for each within-subject comparison). The presence of catch trials limits the theoretical ceiling of learning indices to 0.83. Error-bars are SEM.

2.3.2 Existing Models and predictions

To account for the observation that performance during motor learning exhibited time-dependent changes such as savings and spontaneous recovery, our group had previously proposed a deterministic multi-rate model that suggests that motor output is a sum of at least two systems: a fast adapting system that rapidly forgets, and a slow adapting system that has good retention (Smith et al., 2006). In this model, the learner's state (represented by variable *z*) is a reflection of two underlying systems: one system (represented by state z_1) is highly sensitive to errors and changes rapidly, but has a limited capacity and will tend to quickly forget. Another system (represented by state z_2) has low sensitivity to error and changes slowly, but has large capacity and will tend to remember. The learner's "knowledge" on each trial is a sum of contribution of the fast and slow systems: $z = z_1 + z_2$. In this model, following an error $y^{(n)}$ in trial *n*, the state of the learner $\{z_1, z_2\}$ changes deterministically as follows:

$$\begin{aligned} z_1^{(n+1)} &= a_1 z_1^{(n)} + b_1 y^{(n)} \qquad &\text{(fast system)} \\ z_2^{(n+1)} &= a_2 z_2^{(n)} + b_2 y^{(n)} \qquad &\text{(slow system)} \end{aligned} \tag{2.11}$$

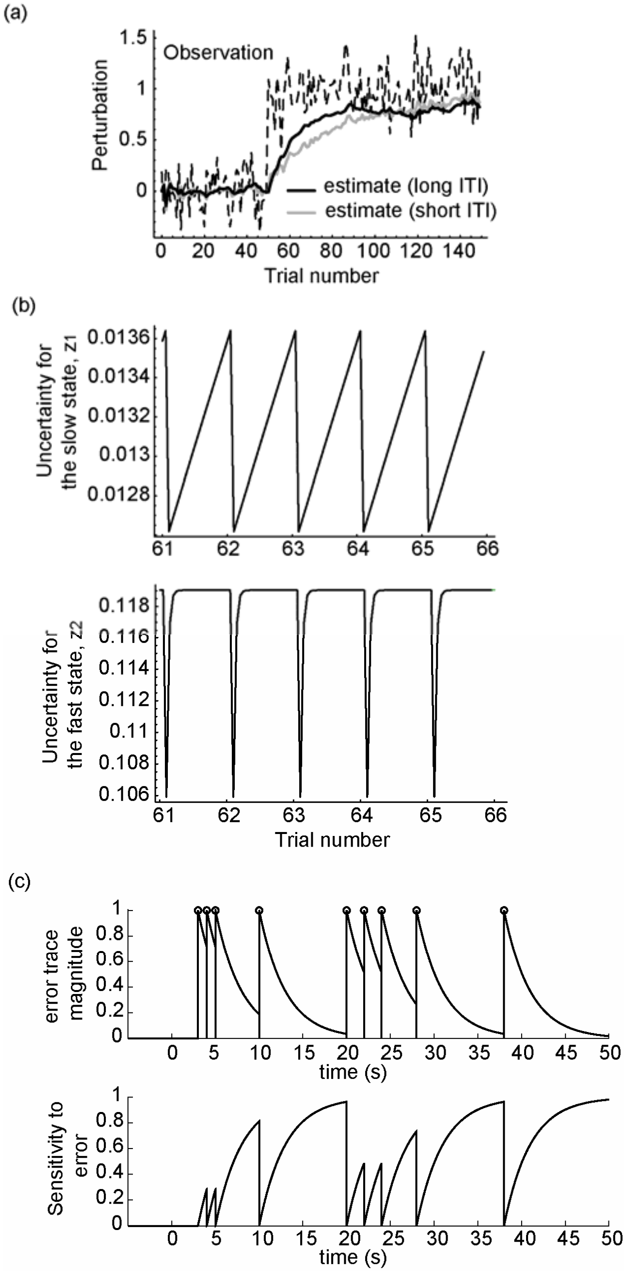

Figure 2.3 Effect of ITI on the Bayesian and error trace models. **a.** Simulation of the Bayesian multi-rate model. The thin black line represents the perturbation observed by the learner, and the thick lines are the learner's estimate of the perturbation when observations are made with either a long (black) or a short ITI (gray). To produce these plots, Eqs. (2.5)-(2.8) were simulated iteratively with observation noise $\sigma_0^2 = 0.04$, state noises $\sigma_1^2 = 0.00008$ and $\sigma_2^2 = 0.1$, and forgetting rates $a_1 = 0.999$ and $a_2 = 0.4$. We assumed that each iteration took a constant amount of time and defined a "trial" as an observation. For the short ITI, the learner made an observation on every other iteration of the simulation. For the long ITI, the learner made an observation on every 20th iteration of the simulation. Longer ITI produced a faster rate of adaptation. **b.** Evolution of parameter uncertainty in the long ITI scenario for 6 consecutive trials. The learner estimates the perturbation as a sum of two parameters in Eq. (2.3), z_1 (slow state) and z_2 (fast state). There is an uncertainty (or variance) associated with each parameter. Increased uncertainty due to a longer ITI prior to an observation (i.e. a "trial") will result in increased sensitivity to error in the *following* observation and thus, faster rate of adaptation. When the learner makes an observation, the uncertainties decrease sharply. Between trials, the learner cannot make observations but estimation of uncertainties change. The uncertainty for the fast state rapidly increases and saturates. In contrast, the uncertainty for the slow state gradually increases. **c.** Illustration of the error trace model. For simplicity, each observation / trial (circles) generates an exponentially decaying trace. An old error trace is replaced as soon as a new movement is made. A long interval *after* an observation allows more time to integrate the error trace and therefore increases the sensitivity to the *preceding* error.

In Eq. (2.11) we have $a_1 < a_2 < 1$ and $b_2 < b_1 < 1$. In this equation, passage of time affects the learner through the variables a_1 and a_2. Since these variables are less than one, passage of time always degrades memory. Therefore, the model cannot explain the finding of improved rates of learning in spaced training. However, a variation of this model casts the timescales in a Bayesian framework (Eq. (2.3)). In this framework, the learner keeps a measure of uncertainty about its knowledge. Importantly, the passage of time between trials affects this uncertainty. As its uncertainty escalates, the learner's incentive to learn increases. Figure 2.3a shows a simulation of this method of learning. We simulated 150 trials in two conditions: with a short ITI (two model iterations between observations), or a long ITI (twenty model iterations between observations). The longer ITI produced faster rates of adaptation. To see the reason for this, it is instructive to examine the parameter uncertainties in the long ITI scenario (Figure 2.3b). With each observation (i.e. trial), the learner acquires information, and therefore its uncertainty declines. During the interval between observations, the uncertainties increase. Yet, the rate of this increase is different for the fast and slow states. The faster state has higher noise, and so its uncertainty rapidly increases until it reaches an asymptote. The slower state has less noise, and so its uncertainty slowly increases during the ITI. Therefore, the longer ITI disproportionately affects the uncertainty of the slow state. Increased uncertainty means an increased sensitivity to error that the learner experiences in the subsequent trial (Eq. (2.5)). As a result, the Bayesian learner adapts faster with a longer ITI.

The Bayesian model explains that the learner adapts faster in the longer ITI design because his uncertainty grows with passage of time. Therefore, this model predicts that the sensitivity to error in trial n will become larger with increased time between trials n-1 and n. Experiment 2 was designed to test this prediction.

2.3.3 The link between trial-to-trial sensitivity to error, generalization, and overall learning rate

In the Bayesian model, Eq. (2.5)-(2.8) describe how the state of the learner changes on a trial-to-trial basis. We showed that in simulation, an increased trial-to-trial sensitivity to error sufficiently expedites overall learning rate (Figure 2.3a). The trial-to-trial sensitivity to error, however, is not uniform across all movement directions. Gandolfo et al demonstrated that the amount of trial-to-trial sensitivity depends on the angular disparity between movement directions (Gandolfo et al., 1996). That is, errors in one direction affect the states of the learner in other directions – a phenomenon termed generalization. Therefore, if increased ITIs increase trial-to-trial learning and in turn overall learning rate, the effect will produce a modulation of generalization as a function of ITIs.

2.3.4 Experiment 2: sensitivity to error in trial n increased with time between trials n and n+1

In Experiment 2, the delay between movements was drawn from a multinomial random variable. The ITIs were 4, 9, 14 and 24 seconds (respectively, 4.14±0.17s, 8.63± 0.14s, 13.63± 0.16s, or 23.70± 0.12s (mean±SD)). We were interested in measuring the learner's sensitivity to error in each trial as a function of the ITI. To estimate this sensitivity, we used a state-space approach (Donchin et al., 2003). For each subjects we measured the error $\mathbf{y}^{*(n)}$ in each trial n, force $\mathbf{f}^{(n)}$ produced by the robot at that velocity, and target direction $L^{(n)}$. As the movement directions were identical between subjects, we averaged the movement errors and forces across subjects and arrived at a single sequence.

To determine how the error in each trial affected the movement that the subject made in the subsequent trial, we fitted this sequence to a hidden state space model (Donchin et al., 2003; Wainscott et al., 2005). The hidden states represented the knowledge of the learner about the perturbation in each direction of movements. Upon experience of an error in a given direction, we estimated how this error was generalized to other directions. We also estimated how this generalization was modulated by the time spent between the trials. The sequence of movement errors were fitted to the following dynamical system:

$$\begin{aligned} \mathbf{y}^{(n)} &= D\mathbf{f}^{(n)} - L^{(n)}\mathbf{z}^{(n)} \\ \mathbf{z}^{(n+1)} &= a\left(t_{n+1} - t_n\right)\mathbf{z}^{(n)} + BL^{(n)T}\mathbf{y}^{(n)}k\left(t_n - t_{n-1}\right) \end{aligned} \tag{2.12}$$

In this model, **z** represents a vector of hidden states (learner's knowledge about the perturbations for each of the 8 possible target directions). By fitting the observed variables ($\mathbf{y}^{*(n)}, \mathbf{f}^{(n)}$, and $L^{(n)}$) to Eq. (2.12), we estimated the unknown parameters D (the arm's compliance), B (the generalization function), a (a time-dependent function that describes the deterioration of the state during the time between trials), and k (a time-dependent function that modulates the generalization function as a function of time between trials). The Bayesian model predicted that k would be a monotonically increasing function of Δ, where $\Delta = t_n - t_{n-1}$.

Using non-linear optimization (the *lsqnonlin* function in Matlab with default settings), we fitted Eq. (2.12) to the measured data ($F(35,349)=33.95$, $r^2=0.7730$, $p< 0.0001$). The resulting generalization function B (with standard errors of the mean estimated through a bootstrap procedure) is plotted in Figure 2.4a. This function had its peak at 0^o, and decreased with angular distance in a pattern similar to those recorded in other studies where ITIs were kept constant (Donchin et al., 2003; Wainscott et al., 2005). We found that the function a remained extremely close to 1 (Figure 2.4b), suggesting that there was little or no forgetting during the seconds that passed between trials. Arm compliance D was consistent with previous measurements of limb compliance (Figure 2.4c)(Mussa-Ivaldi et al., 1985). However, contrary to the predictions of the Bayesian model, we found that *k* did not monotonically increase with the time that preceded the trial (Figure 2.4d, gray line).

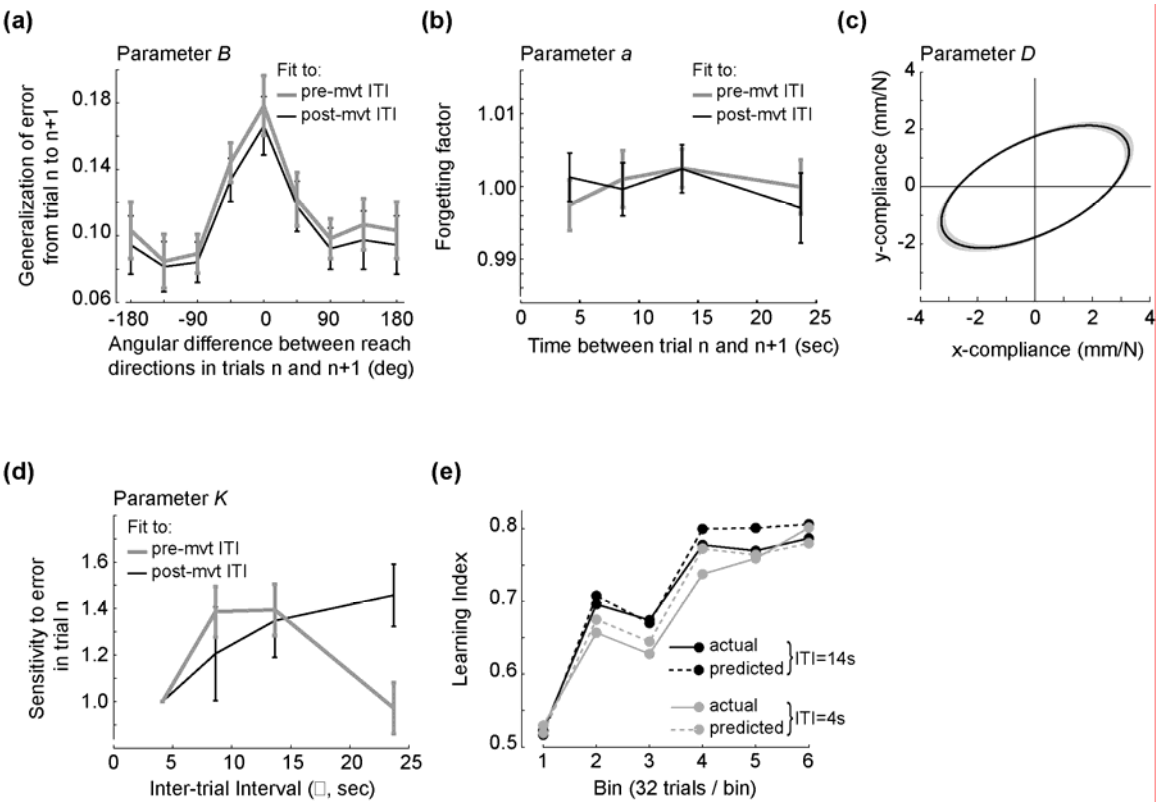

Figure 2.4 Sensitivity to error as estimated from Experiment 2. Data from Experiment 2 were fitted to Eq. (2.12) and Eq. (2.13). For Eq. (2.12), it is the *pre*-movement ITI that affects the sensitivity to error. For Eq. (2.13), it is the *post*-movement ITI that affects the sensitivity to error. The Bayesian model predicts that sensitivity to error should increase with pre-movement ITI. The trace model predicts that sensitivity should increase with post-movement ITI. Error bars are bootstrap estimates of SEM. **a.** The generalization function *B*. This function quantifies the fraction of error experienced in trial *n* that affected motor output in trial *n*+1. The function is parameterized with respect to the angular distance between trials *n* and *n*+1. Error-bars are bootstrap estimates of the SEM (200 re-samples). There were no significant differences in estimates of *B* for the two models. **b.** Parameter *a*. This parameter estimates the effect of time passage on the decay of memory. A value of 1 indicates no decay. **c.** Parameter *D*. This parameter estimates limb compliance. The estimates for the two models were nearly identical. **d.** Parameter *k*. This parameter estimates the sensitivity to error in trial *n*. For the pre-movement ITI model, the plot shows $k(\Delta)$, where $\Delta = t_n - t_{n-1}$, or $\Delta = t_{n+1} - t_n$. To estimate $k(\Delta)$, we arbitrary set $k = 1$ for $\Delta = 4s$ and found the remaining values with respect to this value. **e.** Cross validation of the error trace model. We fitted Eq. (2.13) to the data in Experiment 2, resulting in the parameters shown in parts a-d, and then used these parameters to predict performance in Experiment 1.

2.3.5 The Error Trace Model

Because the results of our fit were inconsistent with the Bayesian model, we considered an alternative model. It is possible that in spaced training, adaptation rate is faster because errors produce a memory trace that decays with time but that the learner continues to benefit from the trace during the ITI period (Figure 2.3c), effectively continuing to learn from the error trace. In this model, the time that is of importance is the period that follows a trial, not the time that precedes it (Eq. (2.10)). The error trace model predicts that sensitivity to error in trial *n* should increase as a function of the time between trials *n* and *n*+1. (In contrast, the Bayesian model predicted that the sensitivity to error in trial *n* should increase a function of time between trials *n*-1 and *n*.) To test for this prediction, we slightly modified Eq. (2.12) to represent this idea, and then fitted Eq. (2.13) to the same data observed in Experiment 2:

$$\begin{aligned} \mathbf{y}^{(n)} &= D\mathbf{f}^{(n)} - L^{(n)}\mathbf{z}^{(n)} \\ \mathbf{z}^{(n+1)} &= a\left(t_{n+1} - t_n\right)\mathbf{z}^{(n)} + BL^{(n)T}\mathbf{y}^{(n)}k\left(t_{n+1} - t_n\right) \end{aligned} \tag{2.13}$$

The results produced a highly significant fit ($r^2 = 0.7733$, $F(35, 349) = 34.01$, $p<0.0001$). While there were little or no changes in the model parameters *B*, *a*, or *D* (Figure 2.4a-c), the important parameter *k* changed significantly: it now became a function that monotonically increased with ITI (Figure 2.4d). That is, results of experiment 2 suggested that the sensitivity to error in trial *n* increased as a function of post-movement ITI, not pre-movement ITI. This result was consistent with the error trace model but not the Bayesian model.

Was the sensitivity to error significantly greater at longer ITIs? To check for this, we asked whether the slope of the function $k\left(\Delta\right)$ in Eq. (2.13) was significantly greater than zero. We first computed $k\left(\Delta\right)$ from each of the bootstrapped groups and then estimated the slope by fitting a straight line. The p-value was estimated by counting the number of non-positive slopes and then dividing by the total number of bootstrapped samples. We found a significantly positive slope ($p < 0.005$). On average, the subjects' generalization was 45% higher during those trials with the longest ITI with respect to those with the shortest ITI.

What was the time constant of the error trace? To estimate this time constant, we replaced the term $k\left(\Delta\right)$ in Eq. (2.13) with the exponential function in Eq. (2.10) and refitted the system of equations to the measured data. We again found a highly significant fit ($r^2 = 0.77$, $F(29, 355) = 40.8$, $p<0.0001$). The fit estimated an error trace time constant of about 4 seconds ($r = 3.77 \pm 0.6$ sec).

In summary, we found that improved performance in the spaced trials was due to increased error sensitivity as a function of the period that followed the movement, as would be predicted by the error trace model, and not the period that preceded the movement, as would be predicted by the Bayesian model.

2.3.6 Cross validation of the error trace model

To validate the error trace model, we performed two tests. First, we asked whether the specific parameter values found in the random ITI of experiment 2 could explain the performances in the constant ITI of Experiment 1. Second, we performed experiment 3 to specifically measure the change in motor output as a function of ITI.

In our first test, we asked whether the model of Eq. (2.13) could predict the specific shape of the learning function that we had measured in the short and long-delay conditions of Figure 2.2d. We ran the dynamic system of Eq. (2.13) on the target sequence $L^{(n)}$ and force sequence $\mathbf{f}^{(n)}$ of experiment 1, generating a sequence of errors in field and catch trials. We then computed a learning index on this sequence of movement errors in the same way that we had computed the performance of our subjects

(Eq. (2.2)). For simplicity, predicted and actual performance was computed on a single sequence of movements by first averaging data across subjects; the results were similar to that shown in Figure 2.2d. Figure 2.4e shows performance of the predicted and the actual data from experiment 1. There was an excellent correspondence between the predicted performance and the measured data.

2.3.7 Experiment 3: Predictions of the error trace model

A crucial prediction of the error trace model is that when one experiences a motor error, the longer one waits, the larger will be the effect of this error on the motor output in the next trial. To test this prediction, we employed force channels, first introduced by Scheidt et al. (Scheidt et al., 2000). In a "channel" trial, the hand is guided and restricted along a straight line to the target. The motor output that is relevant to the task is the force that the subject produces against the walls of this channel. In a triplet of "channel-field-channel" trials, one can measure the change in motor output between the two channel trials as a function of the error in the intervening field trial. The prediction of the error trace model is that this change should monotonically increase with the time period between the field trial (when error was experienced) and the second channel trial.

In a channel trial *n*, let us label the force perpendicular to the direction of motion produced at maximum velocity as $u^{(n)}$. Assume $u^{(n)} = \hat{V}^{(n)}\dot{x}^{(n)}$ where $\hat{V}^{(n)}$ is the subject's best guess at the constant V in Eq. (2.1), and that $\hat{V}^{(n+1)} = \hat{V}^{(n)}$ since subject did not experience an error in trial *n* (as it was a channel trial). Now in trial *n*+1 (field trial), the subject experienced a movement with error $y^{(n+1)}$, and this results in adaptation:

$$\hat{V}^{(n+2)} = \hat{V}^{(n)} + s(\Delta)y^{(n+1)} \tag{2.14}$$

In Eq. (2.15), the learner's trial-to-trial sensitivity to error is labeled with variable s. This sensitivity is a function of time between trial n+1 and n+2, that is: $\Delta = t_{n+2} - t_{n+1}$. Thus, for any channel-force-channel triplet, we have:

$$\begin{aligned} u^{(n+2)} &= \hat{V}^{(n+2)} \dot{x}^{(n+2)} \\ &= u^{(n)} \frac{\dot{x}^{(n+2)}}{\dot{x}^{(n)}} + s(\Delta) y^{(n+1)} \dot{x}^{(n+2)} \end{aligned} \tag{2.15}$$

We fitted this equation to each triplet in experiment 3 and estimated sensitivity to the error experienced in movement *n*+1 as a function of the time interval between movements *n*+1 and *n*+2 (Figure 2.5). (We kept the time interval between movements *n* and *n*+1 constant in this experiment). The ITIs between movement *n*+1 to *n*+2 were 4, 8, 13 and 23 seconds (or more precisely, 3.66 ± 0.05 s, 8.23 ± 0.06 s, 13.30 ± 0.05 s, and 23.37 ± 0.07 s (mean±SD)). There was a significant effect of the time interval between movements *n*+1 and *n*+2 (one-way ANOVA, F(3, 81) = 3.156, p = 0.029), and one-tailed t-tests revealed that the sensitivity was significantly higher for longer ITIs than shorter ones (Figure 2.5). It should be noted that the relationship between regressed error sensitivity and ITI appeared to be linear rather than exponential. It is not clear whether the linearity was a result of the regression or the randomness of the perturbation given. However, consistent with the general predictions of the error trace model, we observed that when one experienced an error, one learned more from that error if one waited longer before the next trial.

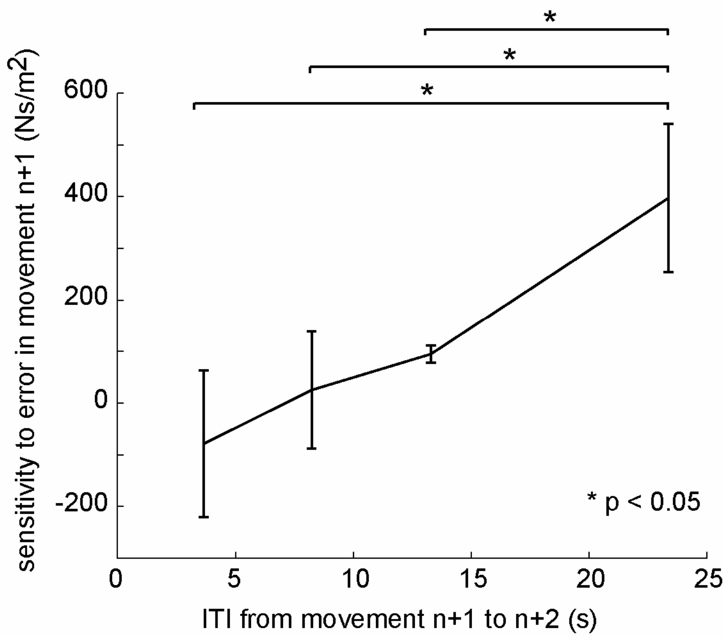

Figure 2.5 Results of Experiment 3. Sensitivity to random errors in trial n+1 as a function of time between trial n+1 and n+2. (The time interval between trials n and n+1 were kept constant at 4 s.) The parameter for sensitivity s was estimated via the change in the motor output u in channel trials n and n+2 (Eq. (2.15)). To arrive at this figure, Eq. (2.15) was fitted to the data for each subject, resulting in an estimate of $s(\Delta)$ for each subject. The values shown are the mean±SEM.

2.4 DISCUSSION

Two previous reports had demonstrated that reach adaptation required fewer training trials when trials were spaced in time (Bock et al., 2005; Francis, 2005); a finding that we reproduced in experiment 1. (In terms of absolute time, long-ITI sessions took longer to complete.) These data demonstrate that the motor system is affected by not only motor error, but also time. How does time affect the way the brain learns from motor error?

We considered two models of motor adaptation that are sensitive to passage of time: a Bayesian multi-rate model (Eq. (2.5)) and an error trace model (Eq. (2.10)). Our previous work suggested that motor errors result in an adaptive response in at least two "systems": a fast system that rapidly learns but has poor retention, and a slow system that is less sensitive to error but hardly forgets (Smith et al., 2006). Unfortunately, this model could not explain the results of experiment 1. A Bayesian variant of this model recasts it in a probabilistic framework (Kording, 2007). It hypothesizes that spaced training leads to improved rates of adaptation because during the time between trials $n-1$ and n, the brain becomes uncertain about its internal model. The increased uncertainty results in increased sensitivity to subsequent motor errors. The model predicts that if one could measure error sensitivity on each trial, one would find that the sensitivity to error in trial n increases as a function of the time between trials n-1 and n. In experiment 2, we tested and found results that were inconsistent with this prediction: the sensitivity to error in trial n monotonically increased with the time period that *followed* that trial, not the period that *preceded* that trial. Therefore, spaced training improved rates of performance not because time delay made the learner more sensitive to the error in the next movement, but because it made the learner learn more from the error in the last movement.

To account for this result, we proposed that the error might be represented by a trace that exponentially declined with time, effectively allowing the nervous system to learn from the trace for as long as it was available. Results of experiment 2 suggested that the error trace had a time constant of 4 sec. To test the model more directly, we performed a final experiment where movements experienced random errors, but were sandwiched between "channel" trials from which we could measure change in motor output from error. Consistent with the trace model, we found that the sensitivity to error experienced in trial n increased with the delay between trials n and n+1.

We assumed that the error trace was "reset" by the next trial. We also considered the possibility that the error trace lingered beyond the immediate next trial. If the time constant of such a "lingering" trace is more than a few seconds, one expects that the generalization function will be independent of the ITI. This was inconsistent with our results in the second experiment.

We believe cognitive strategies had minimal effect in our data. We examined the effect of conscious policy versus an implicit memory in a recent report and found that there was a high probability that subjects would become conscious of the force field pattern when the number of targets was small (3) and the forces were consistent (Hwang et al., 2006a). Furthermore, in that experiment we did find that subjects who became conscious of the force pattern had a small but significant boost on learning versus those who did not. In the current experiment, we designed our procedures to minimize this effect. First, we had 8 targets rather than 3. Second, we performed a control experiment where forces were random. Importantly, we observed that the effect of ITI in the random experiment was consistent with the ITI effect in the constant force field experiment. Furthermore, we asked subjects what they were thinking during the inter-trial intervals in post-experimental questionnaires (experiment 3). No subject was aware of the force pattern (as it was random) and only one subject in 28 answered that they were attempting to come up with a cognitive strategy. Since participants were required to manually keep the cursor centered during the wait, the vast majority answered that they were tying to focus on centering the cursor.

The idea that adaptation might take place during the time between trials is a common theme among computational models of learning where events produce an eligibility trace for synaptic plasticity. For example, Sutton and Barto (Sutton and Barto, 1981) suggested that a stimulus or error signal that excites a neuron may produce an eligibility trace on the neuron's synapses that acts as a low-pass filter of that input. When the input is removed, the trace declines exponentially in time. As long as the error and stimulus traces are available, their coincidence results in modification of the synapse associated with the stimulus. If each new error or stimulus cancels the trace of the previous input, such models predict that the effect of a given error should grow with the ITI between trials. The results of experiments 2 and 3 are consistent with this framework.

For the reaching task considered here, a candidate area where such computations may be performed is the cerebellum (Smith and Shadmehr, 2005). An eligibility trace (or in our terms, sensitivity to error) may be represented by the concentrations of second-order messenger chemicals. Several studies suggest that parallel fiber activity is responsible for the graduate rise of this trace (Kettner et al., 1997; Raymond and Lisberger, 1998). Similar proposals have been suggested for timed learning in delayed conditioning of eye blinks (Fiala et al., 1996); parallel fiber activity leads to increased phosphorylation of receptors over time and in turn, reduces Purkinje cell firing during the interval between the sustaining conditioned stimulus (CS) and the onset of unconditioned stimulus (US). Interestingly, it has been observed that delayed conditioned response can be learned if CS precedes the US by up to 4 seconds (Gormezano, 1966) – suggesting that such persistent phosphorylation of Purkinje cell receptors would have similar time course put forth here in the error trace model.

From a neurobiological perspective, synaptic changes that are produced by spaced training produce memories that are dependent on protein synthesis (Tully et al., 1994; Maldonado et al., 1997; Josselyn et al., 2001; Locatelli et al., 2002; Scharf et al., 2002; Comas et al., 2004). Indeed, cellular response in animal models to spaced stimuli may prime additional memory traces and give rise to resistance to memory interference in different temporal phases (Isabel et al., 2004). However, such cellular processes generally occur on a much longer time scale than what we examined here. The structure of our model might imply that the error trace is kept in some kind of buffer that continues to benefit the learner. Any mechanism with which the influence of the error on the next trial can grow as a function of ITI will produce the same result. What might be the neural basis of such a mechanism?

Neurons that are stimulated with longer ITIs produce larger LTP (Scharf et al., 2002) and are more resistant to de-potentiation. Staddon et al. (Staddon and Higa, 1996; Staddon et al., 2002) and Fusi et al. (Fusi et al., 2005) proposed a cascade model of synaptic plasticity that can account for this. For example, in the model of Fusi et al. (Fusi et al., 2005), a given synaptic strength is supported by a synaptic state that may be shallow or deep in its cascade. The probability of transition in the synaptic state depends on the depth of that state: the deeper the state, the more resistant it is to change. If we imagine that it takes time for the internal state of the synapse to transition from one depth in its cascade to another, and that the time needed increases with the depth of the state, then events that can cause synaptic change are more effective when they come spaced in time. At short ITI, only those synapses change that have a shallow internal state. With increased ITI, one engages not only the shallow state synapses, but also has a higher likelihood of causing state change in the deeper state synapses. Neural models of adaptation that rely on such synapses should exhibit the ITI dependent patterns of generalization that we found here.

There are a number of limitations to our error trace model of motor adaptation. By itself, the model cannot account for much of the rich body of data that was recently highlighted in Smith et al. (2006). For example, if there is an error trace, at this point we do not know how that trace affects the fast and slow systems that were inferred from that study. To account for that data, one idea is to combine the error trace model with the Bayesian model (Kording, 2007) so that the effect of an observed error is a memory trace that decays in time. It has been suggested that uncertainty of a task variable is encoded in the lateral intraparietal area (Platt and Glimcher, 1997; Schall and Thompson, 1999) and acetylcholine and norepinephrine have been suggested to play crucial roles in forming the context-dependent priors during learning (Yu and Dayan, 2005; Dayan and Yu, 2006). Our data does suggest that this uncertainty is not changing during the delay period between trials.

CHAPTER 3

LEARNING OF THE TASK STRUCTURE AS REVEALED BY MOTOR RETENTION

3.1 INTRODUCTION

The ability to modify our motor outputs allows us to interact with the environment quickly and efficiently. This ability includes adaptive responses to changes in the environment (e.g. swimming in water), changes in the medium between the environment and our body (e.g. using tools), and changes in our body (e.g. natural growth). Many studies have established that the brain constructs an internal model -- an approximation of the state of the reality (i.e. the environment, the medium, and the body) -- upon which motor actions are planned and executed. Any errors caused by the discrepancy between the internal model and reality then are used as the main driving signal for the brain to modify its model.

Adapting to these changes, however, is a complex task. Changes can take place at a different frequency and pace, depending on the structure of task. Natural growth, for example, slowly changes the inertial properties of our limbs and trunk. On the other hand, tool use such as handling a hammer can change the inertial characteristics of the end actions almost instantly. Within a particular task structure, changes can also take place at a different magnitude. For example in handling a hammer, the hammer can be a light or heavy. Therefore, there are two objectives the learner needs to accomplish. The first is learning the task structure (e.g. adapting a gradual or sudden disturbance). The second is estimating the task parameter (e.g. the weight of the hammer).

In most computational motor control literature, the task structure is kept constant over time. In these studies, the adaptation problem simplifies to the optimal estimation of the task parameter. Various computational models have been proposed to explain the trial-to-trial adaptation to perturbation in this regard with some success (Thoroughman and Shadmehr, 2000; Scheidt et al., 2001; Donchin et al., 2003; Smith et al., 2006; Huang and Shadmehr, 2007). These models are state-space models that assume time-invariance in the task structure. One of these models is the time-invariant two-state model mentioned previously in Chapter 2 (Smith et al., 2006). The two-state model comprises two time-invariant sub-controllers that estimate the expected amount of perturbation in their corresponding states with different trial-to-trial rates: a fast sub-controller that learns from errors quickly and forgets quickly, and a second, slow sub-controller that learns from errors slowly and forgets slowly. The contributions of the states of the two sub-controllers are combined as the basis (internal model) to produce the motor output. In this model, the trial-to-trial learning and forgetting rates are independent constants. As a result, the model is not sensitive to the structure of the task. Indeed, we will show that the two-state model cannot account for the structural learning sensitive data described in the current study.

While a substantial number of literature have addressed the problem of optimal estimation of the task parameter when the task structure is fixed, little is known on whether or how the human motor system responds to a variable task structure within a controlled experimental session. In the current study, we introduced a novel paradigm that fixes the structure of a force adaptation task within a movement training block, but varies the structure across two training blocks separated by a washout block. To avoid confounds, we looked at the passive retention of the adaptation after training. By varying the gradualness at which participants were exposed to changes in their limb dynamics, we investigated the effects of varying training schedules (i.e. task structures) on motor adaptation and passive retention.

Given a training block of a single task structure, the time-invariant two-state model predicts that a sudden experience of errors would drive the fast sub-controller more. Given another training block of another single task structure, a gradual experience would drive the slow sub-controller more. Thus, adaptation may show a fast or slow decay depending on the temporal history of errors within a block. Since a single task structure is used in a single block, the two-state model's insensitivity to task structure is masked. However, motor learning is often done in multiple sessions with intervening time not specifically used for training and it is not clear how motor adaptation is retained after multiple training sessions with intervening washout periods. If our motor system, for instance, behaves invariably with time such as suggested by the time-invariant two-state model, the motor system does not change the rate of adaptation and retention based on prior training experiences. In such case, the retention response should be the same after first training and retraining if the washout period is sufficiently long. This is because in a time-invariant model, the sub-controllers are always sensitive to errors; therefore the learner would actively de-adapt when the source of the errors are removed such as during washout. In contrast, prior experience with training may help motor retention by priming the motor learning mechanisms. In such scenario, the motor system changes the rate of adaptation and retention based on prior experience with training. Therefore, retention behaviors in subsequent training behaviors may depend on how the first training was given.

We considered an aim-and-punch task where human subjects made rapid hand movements toward one of two targets while holding a robotic handle that produced a viscous perturbation pushing the hand orthogonal to the motion. We measured the decay of their adaptive response following matched performance. We also measured the response following a wash-out and retraining block. We found that the first gradual training did induce greater retention as predicted by the two-state model. However, the current motor learning model could not account for the behavioral data after retraining. Importantly, we showed that the gradualness of the first exposure to perturbation alone was sufficient to promote retention of future adaptations. Together, these observations suggested that the human motor retention process receives a top-down control based on the history of training, and the structure of the learning process is not time-invariant as previously assumed.

3.2 METHODS

3.2.1 The aim-and-punch game

We used an aim-and-punch game to examine the influence of training schedules, and in turn, error histories on the properties of motor memory. The goal of the game was to become as proficient as possible at striking through one of two targets with a rapid center-out movement. Participants held the handle of a robotic manipulandum (Hwang et al.). Game graphics were projected onto an opaque, horizontal screen directly on top of the hand and manipulandum. The hand position of the participants was displayed as a small square cursor (5 mm x 5 mm) on the screen at all times and as the game proceeded, a sequence of targets (5 mm x 5 mm) were displayed. In all trials, the closer the hand cursor came to passing through the target, the greater the number of reward points the participant received. Four critical accuracy levels were established: 5.16°, 4.49°, 3.61°, and 2.48°. For each additional accuracy level achieved, the movement was award one additional point in that trial for up to a maximum of 4 points. Participants received point-based financial incentives.

3.2.2 Participants

All participants were right-hand-dominant and used their right hand to perform the movements. Procedures and protocols were approved by the Johns Hopkins Medicine

Institutional Review Board and participants gave their written consent prior to the experiments.

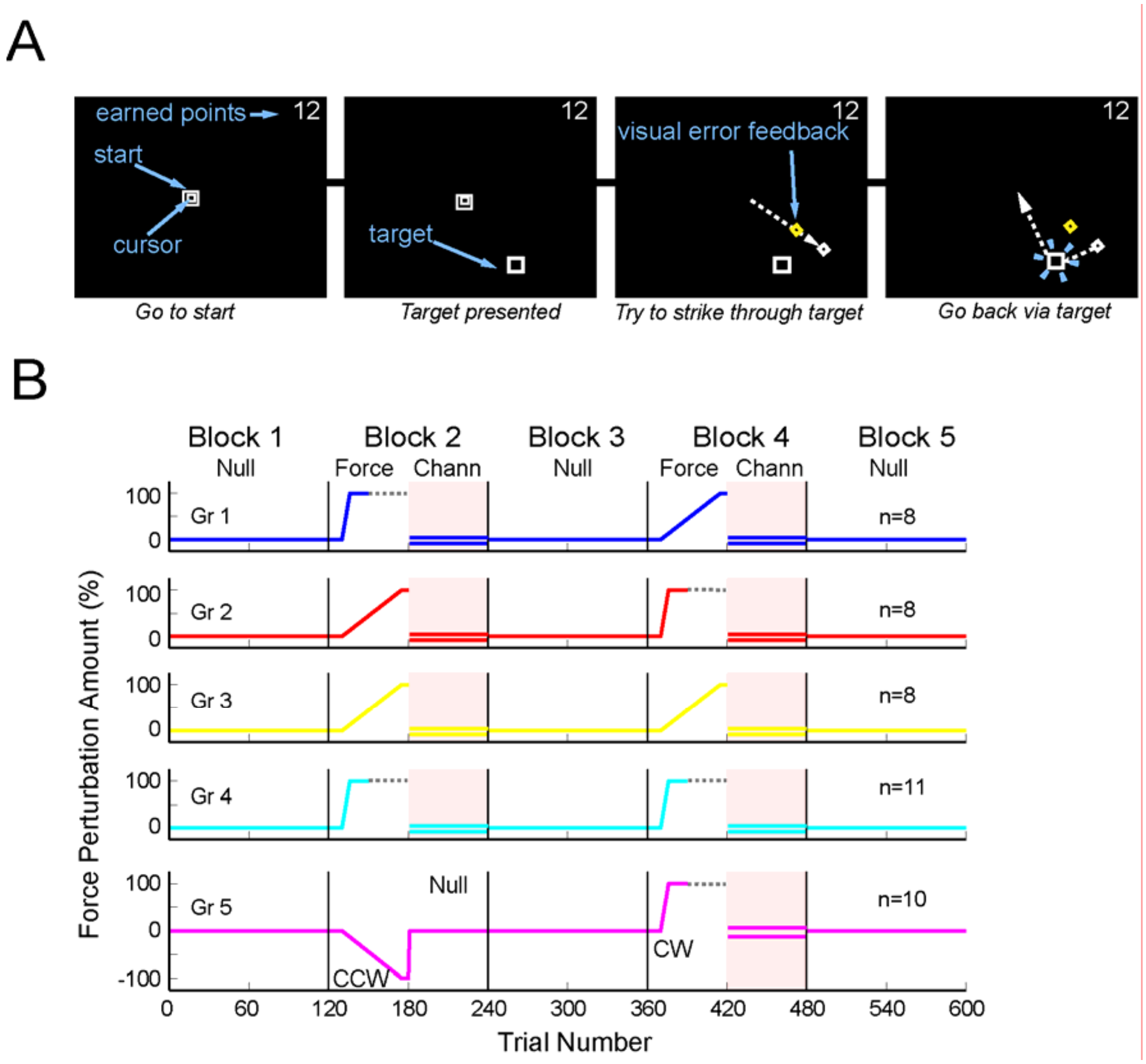

Figure 3.1 Experimental methods **A**: The sequence of prompts and hand movement in a trial. **B**: Experimental training schedules. The curves depict the progression of the perturbation strength through an experimental session. Positive values indicate clockwise, viscous force perturbation while negative values indicate counterclockwise, viscous force perturbation. For clarity, the curves are aligned at the start of each phase in each block. Gray dotted lines indicate the absence of those trials spanned by the dotted lines. Twenty percent of the null and force trials were randomly selected and replaced with channel trials. Blocks of 60 channel trials after adaptation are highlighted with a pink background and a double-line.

3.2.3 Targets

There were two possible targets positioned on an invisible circle of a 10-cm radius at 121.5^o (toward the right shoulder) and 301.5^o (away from the right shoulder). Hand position was displayed at all times as a 5x5 mm white square cursor. At the beginning of each trial, the robot brought the hand to the center mark – a stationary 5x5 mm white crosshair directly ahead of the participants. Targets were then displayed as white squares – this was the "go" signal for the center-out punch. As the strike movement crossed the invisible 10-cm radius circle, a yellow dot appeared at the crossing point to emphasize the distance between the strike and the goal as a measure of reach error. If the movement duration was too long (> 0.23 sec), a blue dot appeared instead. Beyond the invisible circle an elastic force field acted as a "pillow" to absorb the strike. After each strike when the hand hit the pillow, subjects brought their hand back to their intended target. At this point the center mark reappeared and the robot brought the hand back to the center.

3.2.4 Force Perturbation

Each training schedule consisted of three seamless phases in sequential order: a baseline phase, an adaptation phase and, with the exception of the $-G_{LN}$ schedule, a retention phase. In baseline phase trials, participants moved without external forces ("null trials"). To induce errors in adaptation phase trials, we applied a clockwise velocity-dependent force that pushed the hand in a constant direction perpendicular to the hand movement (8 Ns/m) ("force trials"). To measure the force output of participants during the baseline and adaptation phases, we applied a spring/damper force that restricted the hand movement to a straight trajectory toward the target (spring coefficient = 2.5 kN/m; damping coefficient = 25 Ns/m) ("channel trials"). These pseudo-randomly distributed channel trials constituted 20% of the baseline and adaptation phase trials. To measure participants' retained motor outputs, such force channel was used for the entirety of the retention phase. In the -G_{LN} training schedule, the perturbation force during the adaptation phase trials was counter-clockwise, and the retention phase was simply replaced with a block of null trials with channel trials embedded.

3.2.5 Training Schedules

Five groups of participants performed strike movements in five sequential blocks (Figure 3.1B). All groups were exposed to the perturbation in a training schedule in block 2. After an intervening washout block, all participants were re-trained in block 4. The training schedule for each subject group is listed in Table 3.1.

Table 3.1 Training Schedules of subject groups

	Block 1	2	3	4	5
Group 1	N	S_S	N	G_L	N
2	N	G_L	N	S_S	N
3	N	G_L	N	G_L	N
4	N	S_S	N	S_S	N
5	N	-G_{LN}	N	S_S	N

There were three possible types of adaptation training schedules for each block. Schedule "S_S" started with 10 null trials, followed by a 20 force trials that *suddenly* introduced a viscous perturbing force to induce errors, and then ended with 60 channel trials. Schedule "G_L" started with 10 null trials, followed by a 50 force trials that *gradually* introduced the viscous force perturbation, and then ended with 60 channel trials. The numbers of force trials in the S_S and G_L schedules were determined in a preliminary experiment to match the extent of adaptation. In contrast to the force training schedules, the null force schedule ("N") consisted of 120 trials; 80% of which were null trials and 20% channel trials. For Group 5, the first training was performed with the "-

G_{LN}" schedule where force perturbation of the opposite direction (counter-clockwise) was gradually introduced over 50 force trials, followed by 60 null trials.

3.2.6 Performance Measurement

Since full compensation of the force perturbation on a particular trial required a lateral force proportional to the speed profile on the same trial, we computed an adaptation index (AI) on each channel trial found by linear regression of the measured lateral force onto the ideal force required for full compensation of the maximum perturbation(Smith et al., 2006). This index was zero if these forces were uncorrelated and one if they were identical. The index was always relative to the full compensation to the maximum perturbation to allow interpretations based on a single, consistent reference point throughout the entire experimental session. In the modeling results, we simplified the simulation by assuming a scalar perturbation and a scalar output by the model in each trial. Therefore, the simulated AI was simply the ratio of the scalar output normalized by the amount of scalar perturbation.

3.2.7 The time-invariant two-state model

In collaboration with Smith and Kording, one of the authors (Shadmehr) recently proposed a state space model that suggests movement error engages multiple adaptive processes of different time scales (Smith et al., 2006; Kording et al., 2007). We used this model as a springboard to consider the problem of motor retention; however, our behavioral results were inconsistent with the predictions of this model.

A basic version of the time-invariant two-state model has two adaptive units with two distinct internal states; one has a high trial-to-trial sensitivity to error (y) (high learning rate, b_f) and a poor trial-to-trial retention rate a_f of its state (z_f) while the other has a low trial-to-trial sensitivity b_s to error and a high trial-to-trial retention rate a_s of its state (z_s) (Eq.(3.1)). Therefore, the state of the fast process responds quickly to errors while the state of the slow process responds slowly to errors. The overall motor output of the learner for a given trial *n*, is the sum of these two states that responds to error differently (Eq. (3.2)). Importantly, since a and b are both constants, the learner described by this model is time-invariant.

$$\begin{aligned} z_f^{(n+1)} &= a_f z_f^{(n)} + b_f y^{(n)} \quad & b_f > b_s \\ z_s^{(n+1)} &= a_s z_s^{(n)} + b_s y^{(n)} \quad & a_f < a_s \end{aligned} \tag{3.1}$$

$$\begin{aligned} y^{(n)} &= f^{(n)} - \hat{f}^{(n)} \\ \hat{f}^{(n+1)} &= z_s^{(n+1)} + z_f^{(n+1)} \end{aligned} \tag{3.2}$$

It predicted that the gradual training schedule G_L would result in the highest retention during the first channel phase following adaptation, and that after a sufficient number of the 120 wash-out trials (six times the number of S_S phase trials and 2.4 times the number of G_L trials), retention responses after retraining would be similar to those after the first trainings. That is, over-compensation to disturbances would be reduced at a rate that is faster than passive decay.

3.2.8 Statistical Analysis

Unless otherwise indicated, Student’s T-tests and Repeated Measure ANOVA were used in the statistical comparison of the data groups. When Mauchly’s test of sphericity failed, we applied the Greenhouse-Geisser correction in the ANOVA analyses for within-subjects comparisons as needed. All analyses were performed using Matlab or SPSS for Windows (R11.5, SPSS Inc.).

3.3 RESULTS

3.3.1 Fixed task structure within block: First exposure to perturbation by gradual training induces better retention of adapted response

The time-invariant two-state model has a slow learning unit and a fast learning unit; each responds to errors at a time-invariant rate to produce a summed adapted response. Since the slow process has a higher retention rate of its memory (state), a gradual training schedule (G_L) (Figure 3.1B, Group 2, Block 2) would promote the adaptation of the slow state than a sudden training schedule (S_S) (Figure 3.1B, Group 1 / Block 2). To assess the adaptation amount, we computed an Adaptation Index (AI, see Methods) for each 10-trial bins (Figure 3.2). Complete compensation equaled an AI value of one, while zero compensation equaled 0. Figure 3.2A depicts the predictions of the time-invariant two-state model using the parameter values given in the human subject study ($a_f = 0.59$, $a_s = 0.992$, $b_f = 0.21$, $b_s = 0.02$, (Smith et al., 2006)). The relative contributions from the two processes are plotted in Figure 3.2B. To show the passive decay of adaptation, feedback was given to the simulating learner as if it was making perfect adaptations ($y^{(n)} \equiv 0$, i.e. "channel trials"). Since errors were artificially prevented in channel trials, retention of adapted response during the channel phase largely depended on the retention rates of the two processes with the fast state quickly decayed to zero. Therefore, the end retention was contributed mainly by the slow process. Indeed, the gradual training schedule (G_L) produced higher retention after training than the sudden training schedule (S_S).

To probe the difference between the effects of conventional sudden (S_S) and gradual (G_L) training on motor retention, we randomly assigned 16 participants into two corresponding groups (Group 1 and Group 2) in a force adaptation task. The task goal was to make a quick punching movement going through a displayed target. During the adaptation phase, a velocity-dependent, clockwise force acted to redirect the movement laterally. Thus, participants must learn to compensate for the perturbing force in the opposite lateral direction to make their goal. After 50 gradual adaptation trials, Group 2 achieved a compensation amount (AI = 0.53 ± 0.13 mean ± SD) similar to that of Group 1 after 20 sudden adaptation trials (AI = 0.58 ± 0.09). The peak values were not significantly different between the two training groups (two-tail t-test, t(14) = 0.894, p = 0.39), suggesting that similar percentages of the ideal force were produced by both groups (Figure 3.2D).

After both groups had attained similar behavioral adaptation, they underwent 60 channel trials in which their movements were restricted to a straight path toward the target by a stiff spring-damper force "channel." In these trials, angular errors were kept very small (absolute error = 0.0087 ± 0.0075 rad, mean ± SD) when compared to those in the absence of the force channel. These trials allowed us to examine the timeline of decay of the participants' motor output without the obfuscation of error-driven de-adaptation. Despite having the same amount of adaptation at the end of training, gradual training group showed markedly better retention ratio than the sudden training group in both their AI (Figure 3.2C) and raw motor outputs (Figure 3.2D). Both gradual and sudden training groups' motor outputs declined initially. However, participants in the gradual training group stabilized in the second half the channel trial block (AI = 0.30) while the motor outputs of the sudden training group participants continued to decay (AI = 0.14) (Figure 3.2C). Repeated-measure ANOVA analysis showed a significant main effect from the training schedule (F(1,14) = 8.634, $p < 0.05$). We fitted a single exponential function to estimate the time constant for the decays. As predicted by the time-invariant two-state model, the sudden training group ($1/\tau = 0.44 \pm 0.32$ trial^{-1}) declined significantly faster than the gradual training group ($1/\tau = 0.08 \pm 0.10$ trial^{-1}) (two-tailed t-test, t(14) = 2.95, $p < 0.05$). In 60 channel trials, participants in the sudden training group lost 76% of their adaptation, while participants in the gradual training group lost merely 42% by comparison. The data suggested that even though behaviorally participants in both groups showed identical adaptation after training, the specific training schedules that were followed had a profound effect in the mechanism through which the force compensations were produced.

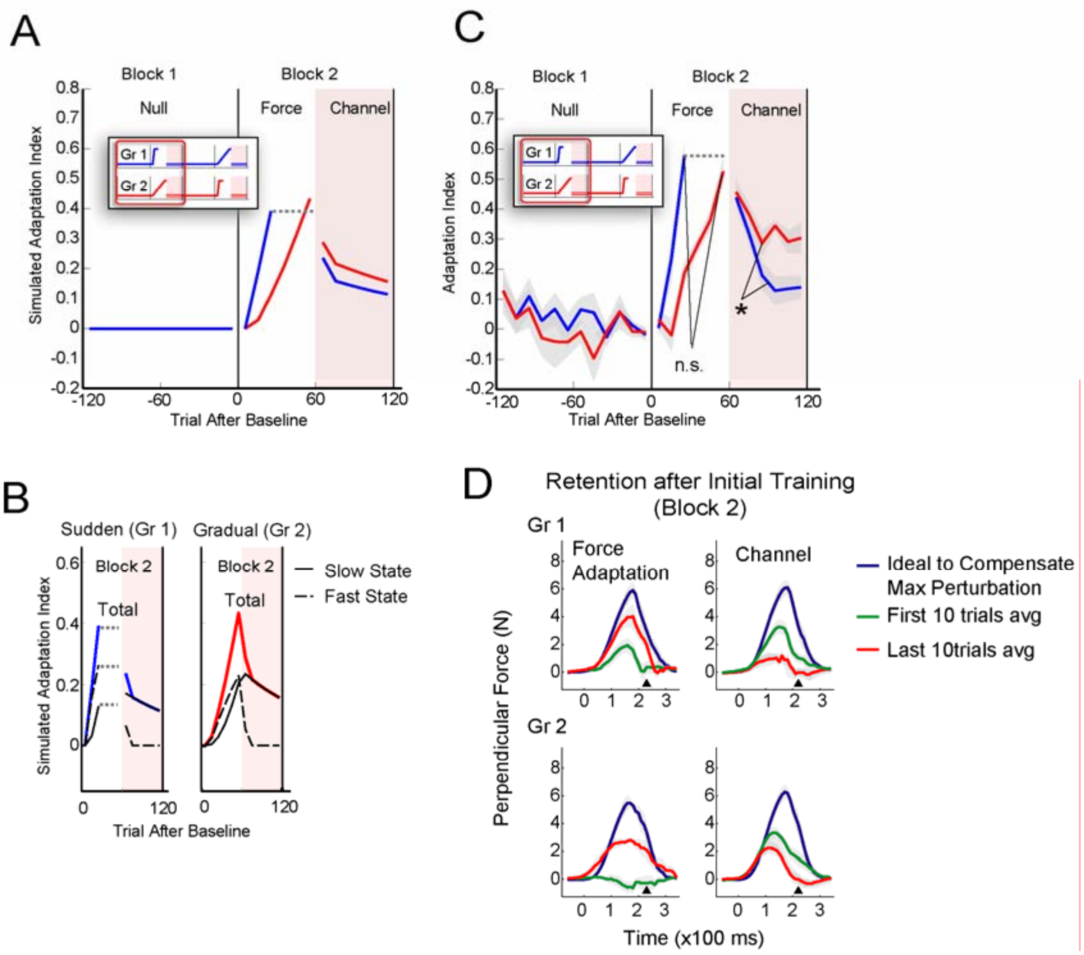

Figure 3.2 Adaptation and retention profiles of the gradually and suddenly trained groups in naïve learners. **A:** Predictions of the time-invariant two-state model in the gradual training schedule G_L (red) or the sudden training schedule S_S (blue). **B:** Relative contributions from the slow (solid black) and the fast processes (dashed black) of the time-invariant two-state model in 10-trial bins. Sum of the contributions are plotted in blue (S_S) and red (G_L). **C:** Adaptation and retention response in bins of channel trials averaged across subjects after gradual or sudden trainings. Adaptation Index is the linear correlation regression coefficient between subjects' own force and the ideal force to compensate for max perpendicular perturbations in channel trials. Grey shades indicate SEM. **D:** Mean perpendicular ideal force to compensate for the full perturbation (blue), and the actual force produced by the participants in the first 10 (green) and the last 10 trials (red) in the adaptation phase and the channel (retention) phase of block 2. Triangles indicate the time when the hand cursor passes the target. Force profiles were averaged across subjects. Grey shades indicate SEM.

3.3.2 Motor adaptation to task structure was not completely washed out after 120 trials

One interesting observation in the field of motor learning is the phenomenon of savings. In a setting where a learner learns a task twice, his learning rate is higher the second time around(Medina et al., 2001; Kojima et al., 2004; Lebron et al., 2004). It is not clear, however, how prior training experience affect the temporal profile to retained motor response especially after an intervening washout. The data thus far have illustrated the effect of gradual training on immediate retention of adapted motor response to an external perturbation in block 2. If gradual training and sudden training engage different learning mechanisms preferentially, how does having prior training affect retention after another retraining?

The time-invariant model failed to produce savings after an extended washout block in simulation. During the extended wash-out block (block 3), both processes de-adapted from over-compensation errors. The fast process, in fact, produced negative compensation while the slow process slowly de-adapted. By the end of the washout block, both fast and slow states were very near their naïve state at 0. Therefore, the learner should respond to the retraining as if the first training never took place. We simulated the model's behavior after the first training and after retraining with the same training schedule (S_S). Despite having adapted to the perturbation in a previous block, the model did not retain the re-adaptation more (Figure 3.3A). These qualitative observations did not change when we used best-fit parameter values obtained from fitting the model to our data ($a_f = 0.85$, $a_s=0.988$, $b_f = 0.08$, $b_s = 0.04$, $r^2 = 0.89$, AIC = -5.02, F(4, 2356)=4737, $p < 0.001$).

Contradicting the predictions of the time-invariant model, the behavioral data showed that prior training experience had an influence on subsequent adaptation and future retentions, even after an extended washout block. As in the simulation, all participants performed in an intervening wash-out block where the perturbation force was removed for 120 trials (Block 3). Averaged AI showed that there was no significant compensatory force produced across all subjects (one-sample t-test, $t(7)=1.69$, $p = 0.14$) at the end of the washout (Figure 3.3B). In block 4, Group 1 and Group 2 swapped their training schedules; Group 1 underwent the gradual training schedule while Group 2 the sudden training schedule.

Since behaviorally the adapted response was washed out for every group, we expected their retention profile after retraining to be identical to that after the first training. Specifically, if participants' adapted response had truly been reset to their naïve states (i.e. identical to block 1), retention profiles of subjects who re-trained with the sudden schedule (S_S) in block 4 should be similar to the retention profiles after the sudden training in the absence of any previous training (block 2). However, Group 2's retained motor response after a sudden training in block 4 was significantly higher than Group 1's retained motor response after an identical training in block 2 (Figure 3.3B, C, two-way ANOVA, main effect on group number, $F(1, 14) = 5.97$, $p < 0.05$).

Taken together, these results suggested that despite the fact that behavioral motor outputs had been washed out successfully, the 120 wash-out trials with no force did not return the adaptive mechanisms and states back to their naïve status. Therefore, additional prior training with the gradual training schedule increased motor retention beyond temporal immediacy.

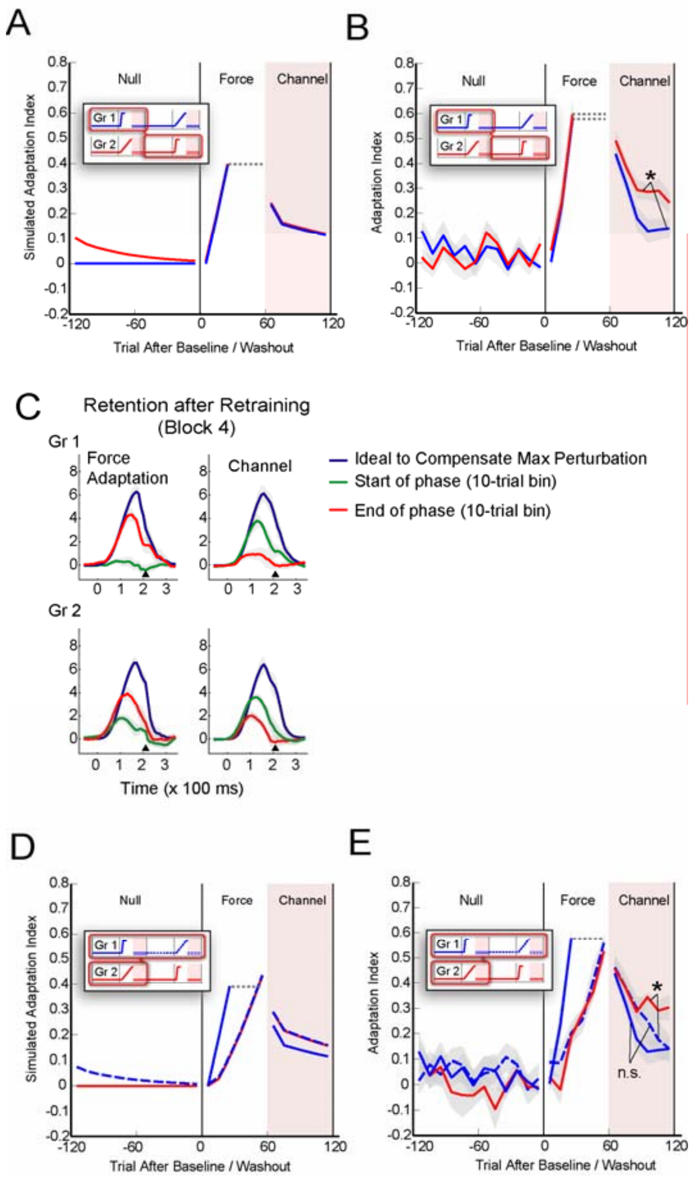

Figure 3.3 Comparisons between naïve and experienced re-adaptation and retention profiles after washout **A:** Predictions of the time-invariant two-state model on the effect of a prior gradual training experience, after washout, on future motor retention. **B:** The behavioral effect of a prior gradual training experience on future retention (red) as compared to that of naïve learners (blue). Grey shades indicate SEM across subjects. **C:** Mean perpendicular ideal force to compensate for the full perturbation (blue), and the actual force produced in the first 10 (green) and the last 10 trials (red) in the adaptation phase and the channel (retention) phase of block 4. Triangles indicate the time when the hand cursor passes the target. Force profiles were averaged across subjects. Grey shades indicate SEM. **D:** Predictions of the time-invariant two-state model on the effect of a prior short, sudden training experience, after washout, on future motor retention. **E:** The behavioral effect of a prior short, sudden training experience on future retention of gradual retraining (blue-dash) as compared to that of naïve learners of gradual training (red). Grey shades indicate SEM across subjects.

3.3.3 Prior training influenced the retention of adapted motor response following subsequent trainings – evidence of structural learning

If adaptation was strictly a cumulative process, we would expect Group 1 to retain much of its adapted motor response in the channel phase of block 4 after gradual training (G_S). If the wash-out process during block 3 was not complete, this could only mean more retained response after a retraining with the gradual schedule (Figure 3.3D).

Neither behavior, however, was observed. Surprisingly, while Group 1 participants did show a marginally higher retention in block 4 when compared to the itself in block 2, the increase was not significant (two-way ANOVA, no block order main effect, F(1,14)=0.972, p=0.34). After retraining with the gradual schedule, the adapted response plummeted to the same low level as it did after a naïve sudden training two blocks prior (Figure 3.3E). Similarly, despite the intervening washout block, Group 2's retained motor response profile after the sudden training was not significantly different from its retention profile after the gradual training two blocks prior. Two-way ANOVA showed that there were no main within-subjects effects found in block order (F(1,7)=0.376, p = 0.559) nor a block × movement bin interaction effect with Greenhouse-Geisser correction (F(2.7, 18.7)=1.196, p=0.335). Thus, it appeared that the gradualness or suddenness of the first exposure to the perturbation set the retention behaviors when the participants were re-exposed to the perturbation. In other words, prior training influenced how participants retain a motor skill after retraining and this influence was dependent on the type of prior experience.

We tested this hypothesis by looking at retention profiles of participants who had identical retraining schedules but different first training schedules before the washout. Since we had two types of retraining schedules (G_S and S_S), we performed two additional experiments to make the comparisons. In the first comparison, we recruited 8 participants (Group 3) who first trained with a gradual training schedule (G_S) in block 2 (different from Group 1) and then again trained with a gradual training schedule (G_S) in block 4 (same as Group 1). We found that participants who were exposed to the perturbation gradually (Group 3) retained significantly more adapted response than those who were exposed to the perturbation suddenly (Group 1) (Figure 3.4B). Two-way repeated-measure ANOVA with Greenhouse-Geisser correction showed that there was a significant interaction effect between the prior training schedule and movement bin ($F(2.7, 38.2) = 3.98$, $p < 0.05$) and a nearly significant main effect of prior training schedule ($F(1, 14) = 4.56$, $p = 0.051$). We fitted a single exponential to the adaptation decay. Group 1's motor response decayed significantly faster ($1/\tau = 0.028 \pm 0.007$ trial^{-1}) than Group 3 ($1/\tau = 0.007 \pm 0.003$ trial^{-1}) (two-sample t-test, $t(14) = 2.68$, $p < 0.05$). If Group 3's higher retention was a result of savings, then we should see increased learning rate during the training phases in block 2 and 4. Indeed for Group 3, adaptation was faster in block 4 when compared to block 2 (two-way repeated measure ANOVA, block number main effect, $F(1,13) = 7.67$, $p < 0.05$). This observation supported that the wash-out block did not return the underlying adapted response back to its naïve state.

In the second comparison, we recruited a fourth group and focused on the effect of prior training schedules on the retention profiles after a second, sudden exposure (S_S) to the same perturbation. Group 4 (n=11) first trained with S_S while Group 2 first trained with G_L in block 2; however, both retrained with S_S in block 4. Again, we found that Group 2's retention in block 4 was significantly better than that of Group 4 (repeated measure 2-way ANOVA, $F(1,17)=5.56$, $p<0.05$), which suggested that retention after retraining was affected by the trainees' first exposure to the perturbation and the benefits of this first gradual training (G_S) lingered beyond the washout block (Figure 3.4C). Since both Group 1 and Group 4 first trained with the S_S schedule, their retention file after the first training were identical (two-way repeated measure ANOVA, $F(1, 17)=0.131$, p=0.72) (Figure 3.4E). Since we observed that the gradual retraining did not increase retention in Group 1, we expected the same from Group 4 who repeated S_S in retraining. Indeed, Group 4's retention profile in block 4 were not significantly different from that of Group 1 after the first training (block 2) (two-way repeated measure ANOVA, F(1,17)=0.04, p=0.84) nor were them significantly different from that of Group 1 after retraining (F(1,17)=0.08, p=0.36) (Figure 3.4D). Also as expected, participants who repeated G_S in retraining (Group 3) were significantly better in retaining adaptation than Group 4. Two-way ANOVA showed that there was a main between-subjects effect (F(1,17)=19.20, p<0.001). Therefore, in addition to the experience of a previous training, the type of prior training also influences motor retention as well.

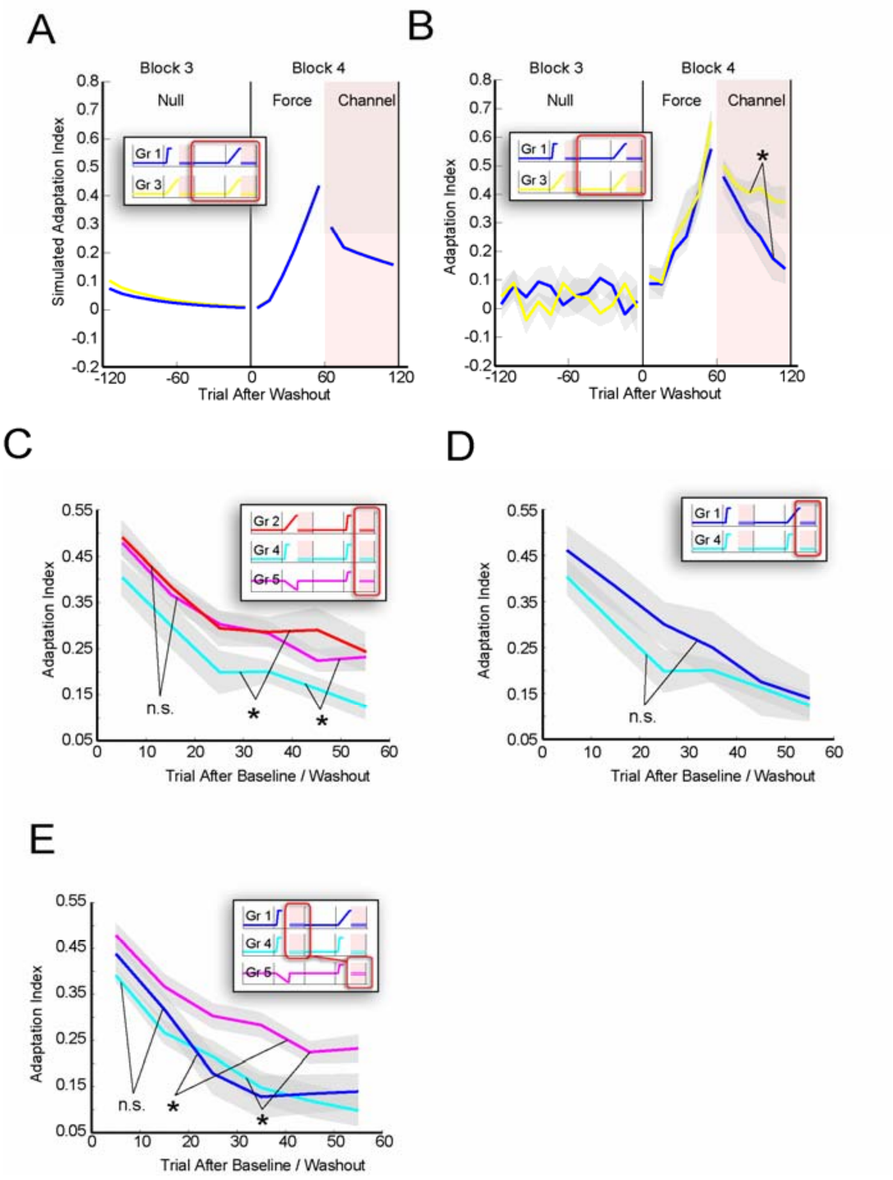

Figure 3.4 The effects of the type of prior training experience on re-adaptations and retentions after washout. Grey shades indicate SEM across subjects. **A:** Predictions of the time-invariant two-state model on the effect of a prior short, sudden training (blue) or a gradual training (yellow), after washout, on future motor retention of gradual retraining. **B:** Behavioral effect of a prior short, sudden training (blue) or a gradual training (yellow), after washout, on future motor retention of gradual retraining. **C:** Behavioral effect of a prior gradual (red), short-sudden training (cyan) with the clockwise perturbation, or a gradual training with the counter-clockwise force perturbation (magenta), after washout, on future motor retention of short, sudden retraining. **D:** Behavioral effect of gradual retraining on retention when subjects were previously exposed to the perturbation with a short-sudden training. **E:** The behavioral effect of a prior gradual training experience in the counter-clockwise force field on future retention of an opposite force field (magenta) when compared to naïve retention profiles for the clockwise force perturbation (blue and cyan). Grey shades indicate SEM across subjects.

3.3.4 The learner is not time-invariant: prior training experience alters the structure of the learner

It is possible that the assumption made by the constant-rate model that the learner is time-invariant was incorrect. For example, if the learner attributed the gradualness or suddenness of the first perturbation introduction as a fundamental characteristic of the novel environment he or she was encountering, the learner may change the way he or she respond to future exposures to the perturbation in the same environment. Namely, the learner not only adapted to the magnitude of the perturbation (i.e. estimation of the state of the perturbation by changing the learner's internal state), but also adapted to the temporal structure of the perturbation by changing his or her own learning structure (i.e. estimation of the trial-to-trial coefficients, b and a) to match the gradualness or suddenness of the exposure. If the learner altered his or her structure as an adaptive response to the temporal structure of the first training structure, we should expect that a gradual training schedule, irrespective to the direction of its content (e.g. clockwise or counter-clockwise perturbation) would induce higher subsequent retention of retraining content.

It was also possible that making movement in a block of 60 channel trials constituted an experimental context which gave rise to the observation that the retention performances after retraining seemed to correlate with prior retention profiles. If such was the case, replacing the first channel phase (block 2) with null trials should remove the context. The additional null trials would only serve to extend the washout period by 50% which again, can only help to abolish the effect of any prior training history experienced in block 2, and thus lower retention.

To verify these predictions, we recruited 10 participants (Group 5) who underwent the same training schedule as Group 2 with two alterations: the first training was now in the counter-clockwise force field (block 2) and the 60 channels at the end of block 2 were replaced with null trials (schedule $-G_{SN}$). The null hypothesis here was, if the learning structure was time-invariant or if participants used the first channel block as a context, final retention should be worse than, or at best, equal to the lower retention amount as seen (i.e. Group 1 block 2). We found that this was not true as retention in Group 5 after retraining was significantly better than naïve retention (Figure 3.4E). Repeated measure ANOVA showed a significant effect between Group 1 and Group 5 ($F(1,16)=6.317$, $p<0.05$). Group 5 retention after retraining was not significantly different from Group 2 at all after ($F(1,16)=0.222$, $p=0.644$) (Figure 3.4C). Similarly, there was a significant difference between Group 4 and Group 5 retention after retraining ($F(1,19)=5.34$, $p<0.05$). These results directly contradicted with the idea that the learner structure was time-invariant. In fact, the structure must have changed to adapt to perturbations of different time scales. Despite a short, sudden retraining, the first gradual exposure to the perturbation was a key factor that led to the increase in retention observed in Group 2 when compared to naïve participants.

To summarize, retention behaviors after retraining were complex. Table 3.2 shows the results of the repeated-measure ANOVA analysis. The type of prior training schedule, and in turn, the temporal structure of the training biased the rate of retention decay after retraining. Consequently, the type of retraining schedule seemed to have little effect in dictating the rate of retention decay.

Table 3.2 Statistical comparison of the retained motor outputs after training / retraining

		Block 2 Retention				Block 4 Retention					
	Gr	1	2	3	4	1	2	3	4	5	
Block 2 Retentio	1		*	*	n.s.	~n.s.	*	**	n.s.	*	S_S
	2			n.s.	**	n.s.	n.s.	~n.s.	*	n.s.	G_L
	3				**	n.s.	n.s.	n.s.	*	n.s.	G_L
	4					n.s.	**	**	n.s.	**	S_S
Block 4 Retention	1						n.s.	~n.s.[1]	n.s.	n.s.	G_L
	2							~n.s.	*	n.s.	S_S
	3								**	*	G_L
	4									*	S_S
	5										S_S
		S_S	G_L	G_L	S_S	G_L	S_S	G_L	S_S	S_S	**Sch**

☐	Repeated-measure ANOVA Between-Subjects Main Effect
▨	Repeated-measure ANOVA Within-Subjects Main Effect
**	p < 0.01
*	p<0.05
~n.s.	0.05 < p < 0.10
n.s.	p>0.10
[1]	Significant mvt-bin x training schedule interaction effect

3.4 DISCUSSION

In the current study, we used a novel paradigm that fixed the structure of a force adaptation task within a movement training block, but varied the structure across two training blocks separated by a washout block. We looked at the passive retention of the adaptation after training. In the first training block, the task structure could be considered fixed because naïve participants had not experienced any task-related perturbation. We observed the benefit of gradual training on the retention of adapted motor response in the first training block. After 20 to 50 training trials, participants were able to compensate for up to 65% of the perturbation force. We probed participants' retention profile of motor outputs without allowing them to de-adapt by artificially eliminating their movement errors. Participants who first underwent 50 trials of gradual training sustained their adapted response more than participants who trained with 20 trials of sudden training. We also verified that when the task structure was fixed, the time-invariant two-state model was able to account for the retention benefit of the gradual training observed in naïve learners by assuming simultaneous adaptation of a slow and a fast learning state.

After an extended washout block, more complex retention behaviors were observed. Prior adaptation experience influenced the amount of retention after re-adaptation regardless of the exposure rate of the retraining. Despite that behavioral adaptation had been cleared by the end of the washout block, participants showed increased retention response after retraining if the first training was gradual; participants showed lower retention after retraining if the first training was sudden and short. Clearly, the washout block did not return the participants to their naïve states prior to adaptation, despite the fact that their motor output was similar to their baseline output prior to any adaptation. Lastly, we demonstrated that gradualness of the first training alone was sufficient to promote retention of future adaptations, even if the retraining was in a dynamical environment different from the first. This was the first evidence that structural learning of the task took place, and estimation of task parameter could be washed out independent of estimation of the task structure.

The result that the mere gradualness of the first exposure to the perturbation was sufficient to promote future adaptation retention was in contrast to the conventional view of the A-B paradigm which had shown anterograde interference when training in the opposite force field, "A," immediately prior (e.g. counter-clockwise) hinders subsequent adaptation to the normal force field, "B" (e.g. clockwise)(Brashers-Krug et al., 1996). There was a protocol difference. In the present paradigm, we separated the two training blocks with a null block, and instead of the rate of the adaptation process, we were interested in the retention response after adaptation. The conventional view of the A-B paradigm which assumes time-invariance in the learner structure, however, would predict that the final retention should be worse than, or at best, equal to naïve retention amount because of the anterograde interference. The presence of the washout block, therefore, might be a key signifying the termination of a force adaptation experience (Krakauer et al., 2005).

The time-invariant two-state model was unable to account for these observations because the behavioral result would mean that the state of the slow process must be inhibited after and only after a short sudden training had previously taken place. The fact that gradualness of the training alone was sufficient to promote retention suggested that there is a third, top-down process that takes precedence of the two-state model. This top-down process inhibits the slow process after a sudden training. For example, an early and abrupt exposure to large perturbations may draw the participants' attention to the fact that their movement errors were more likely to be caused by an external source rather than an intrinsic variation in motor outputs. Attention to the perturbation source, and therefore, context may influence how well motor memories are retained (Kording et al., 2007; Michel et al., 2007).

It is interesting to compare our results to several other studies that also gradually introduced perturbations. Indeed, recent studies of visuomotor adaptation tasks have shown that gradual exposure to kinematics errors led to greater after-effects when compared to sudden exposure of the perturbation (Kagerer et al., 1997; Michel et al., 2007). In these studies, however, the error signal is the misalignment between proprioception and vision of the moving limb. Therefore, it was not clear adaptation occurred at the sensory or the motor plasticity level. Here we eliminated the discrepancy between proprioception and vision. Our results indicated that differential response errors of different time-scale occur at least at the motor plasticity level.

Intuitively, if the environment (perturbation) consistently results in big movement errors, the variance of the environment may be too high for the learner to adapt. The learning rate of the slow process, however, minutely increased with the error size. This agrees with the study of Smith et al. where the learning rate was shown to be inversely proportional with the trial-to-trial correlation of the perturbation(Smith, 2004). If the environment is not consistent, the best strategy is not to adapt to its minute-to-minute changes using the fast process; instead the learner slowly changes his adapted response using the slower process.

In rehabilitation, tool use is common. For example, a stroke patient may practice making movements with a load to reclaim use of his arm. One of the problems facing rehabilitation therapy using assistive tools is that patients may show adaptation only with the rehabilitation aid (Krakauer, 2006). That is, patients made movement adjustment in association with the tool rather than their body. One interpretation of the slow and the fast process in the two-state model is that the slow process corresponds to changes in the body while the fast process responds to fast-varying changes in the environment or context. If this interpretation is true, then we should be able design rehabilitation programs that “fool” the patients to assign the error source to their own and therefore retain their regained motor skills better.

While the time-invariant two-state model allowed us to consider many short-term motor learning phenomena, the present study forces us to consider the global effect of prior experience and context when learning is concerned. Specifically, the initial delivery of training to naïve learners has a significant effect on how the content of the training is remembered in immediate and long-term future. The results of this study, therefore, exposed an imminent need to update our current thinking of motor learning to include models of structural learning such as algorithms employed by the Bayesian network.

CHAPTER 4

SELF-DIRECTED STRATEGY IN MOTOR LEARNING: ACTIVE LEARNING

4.1 INTRODUCTION

A game like tennis involves a number of specific skills: the serve, the forehand, the backhand, etc. A successful player needs to develop a high proficiency in each of these skills. While we may have a coach that closely supervises our training and tells us on which stroke to practice, most of our training time is spent alone: we must autonomously choose the skill component to practice. What factors influence our training strategy?

In the field of machine learning, this problem is termed "active learning". Available to the learner is a list of training examples; each example focuses on a different component of the task. There is no preset curriculum, nor is there an instructor; therefore the learner picks out its own curriculum, one practice example at a time. Examples can be repeated if necessary, but the ultimate goal is to improve in the task overall with a minimum number of examples. The learner may already know some elements of the task. Which example, then, is the best to pick next?

Mathematically, we can summarize the active learning problem of a motor task as follows. A learner has to learn a task represented by a desired, but unknown function $y(x)$. The task is consisted of P different skill components or behaviors ($x_1 ... x_P$). For each component, the learner has a corresponding parameter ($w_1 ... w_P$) representing what the learner has learned about the task. In other words, the parameters constitute the learner's model of the task. On trial n, the learner selects a component $x^{(n)}$ and observes the result of its output $\hat{y}^{(n)}$ and the correct answer $y^{(n)}$. Using the learning gain $K^{(n)}$, the learner updates the corresponding parameter by the difference between actual and desired output:

$$w_x^{(n+1)} = w_x^{(n)} + K^{(n)}\left(y^{(n)} - \hat{y}^{(n)}\right) \tag{4.1}$$

$$w_x^{(n+1)} = w_x^{(n)} + K^{(n)}\left(y^{(n)} - \hat{y}^{(n)}\right) \tag{1}$$

In coached (passive) learning and most motor learning experiments, the environment (e.g. the teacher) instructs the learner which component $x^{(n)}$ to train next. In active learning, the learner itself needs to pick $x^{(n)}$ with the goal to maximizing overall performance by minimizing the expected squared error of the output. If we assume that the learner is on average unbiased, then the expected squared error across all components is equal to the sum of the motor variability σ^2 and the average uncertainty about the desired output.

$$\underset{x}{E}\left[(\hat{y}(x)^{(n)} - y(x)^{(n)})^2\right] = \underset{x}{E}\left[\operatorname{var}(\hat{y}(x)^{(n)}) + \sigma^2\right] = \underset{x}{E}\left[\Sigma_x^{(n)} + \sigma^2\right] \tag{4.2}$$

$$\underset{x}{E}\left[(\hat{y}(x)^{(n)} - y(x)^{(n)})^2\right] = \underset{x}{E}\left[\operatorname{var}(\hat{y}(x)^{(n)}) + \sigma^2\right] = \underset{x}{E}\left[\Sigma_x^{(n)} + \sigma^2\right] \tag{2}$$

$\Sigma_x^{(n)}$ represents the uncertainty of the parameters $w_x^{(n)}$ prior to movement *n*. A policy that reduces the overall uncertainty of the output on trial n+1 optimally is to choose an action $x^{(n+1)*}$ where one is most uncertain from the list of all possible components (see section 4.1.1 for derivation and assumptions, (Cohn et al., 1996):

$$\begin{aligned} x^{(n+1)*} &= \arg\min_x\left(E\left[\left(y^{(n+1)} - \hat{y}^{(n+1)}\right)^2 \mid x^{(n)}\right]\right) \\ &= \arg\max_x\left(\Sigma_x^{(n)}\right) \end{aligned} \tag{4.3}$$

$$\begin{aligned} x^{(n+1)*} &= \arg\min_x\left(E\left[\left(y^{(n+1)} - \hat{y}^{(n+1)}\right)^2 \mid x^{(n)}\right]\right) \\ &= \arg\max_x\left(\Sigma_x^{(n)}\right) \end{aligned} \tag{3}$$

In this derivation we have assumed that the motor noise σ^2 is constant across all components, and that the amount of noise cannot be changed by learning. By practicing in a part of the task space where the estimated output has the highest uncertainty, one can minimize the expected squared error.

We can ask the question whether people's choice can be described using an uncertainty-based model; and if so, how the learner would estimate its own model uncertainty for the task. For example, an algorithm that estimates the model uncertainty during learning is the Kalman filter. The Kalman filter updates the model using Eq. (4.1) with a learning rate $K^{(n)}$ that optimally reduces the model uncertainty. Under this learning rule, the uncertainty of the model parameters will always be reduced for the component of the task that was practiced. Therefore, the estimated output variance (Eq. (4.2)) would be lowered after this particular $x^{(n)}$ was practiced. Together with Eq. (4.3), this leads to the prediction that, for optimal performance, the learner should avoid practicing the same components two trials in a row (i.e. $x^{(n+1)*} \neq x^{(n)}$).

The standard Kalman filter reduces the model uncertainty with every learning instance independent of the error that was encountered. This would lead to the counterintuitive prediction that errors during active learning will not affect action selection. However, there are versions of the Kalman filter that would increase the parameter uncertainty when a big error is observed. One example is a system where the output noise has a distribution that is not Gaussian but drawn from a mixture of Gaussians of identically zero means and different variances. In such case, an observation of a big error would lead to a reduced rate of learning from the large error (Kording and Wolpert, 2004), and an increase in model uncertainty. However, any versions of such model still would predict that a learner should decrease their model uncertainty when a zero error is observed; the general prediction that the learner should not repeat an action also holds true in this case.

To summarize, the uncertainty-based selection rule (Eq. (4.3)) predicts that an optimal learner should choose the component with the highest uncertainty to practice and therefore should not immediately repeat the same action. Here we have taken the first steps in quantifying the factors that affect human choices in active learning of a motor skill. Can their choices be understood in the framework of uncertainty estimation? Do motor errors affect our choice of training?

4.1.1 The best next choice

Let us assume a task consists of *P* different skill components or behaviors ($[x_1 ... x_P]$), each of which is based on a set of corresponding parameters $\mathbf{w} = [w_1 ... w_P]$. After an action on trial *n*, of a particular component, $x^{(n)}$, the system learns from the difference between desired $y^{(n)}$ and produced output $\hat{y}^{(n)}$. Thus, on each trial:

$$\hat{y}^{(n)} = w_x^{(n)} + \varepsilon^{(n)} \tag{4.4}$$

$$w_x^{(n+1)} = w_x^{(n)} + K^{(n)} \left(y^{(n)} - \hat{y}^{(n)} \right) \tag{4.5}$$

$$\hat{y}^{(n)} = w_x^{(n)} + \varepsilon^{(n)} \tag{4}$$

$$w_x^{(n+1)} = w_x^{(n)} + K^{(n)} \left(y^{(n)} - \hat{y}^{(n)} \right) \tag{5}$$

The best next skill component $x^{(n+1)*}$ to train on trial *n+1* is the component that, after learning, will reduce the expected squared error (over all components x) of the system on trial *n+1*.

$$\underset{x}{E}\left[\left(y_x^{(n+1)} - \hat{y}_x^{(n+1)} \right)^2 \right] = \underset{x}{E}\left[\Sigma_x^{(n+1)} + \sigma^2 \right] \tag{4.6}$$

$$x^{(n+1)*} = \arg\min_x \left(E\left[\left(y^{(n+1)} - \hat{y}^{(n+1)} \right)^2 \mid x^{(n+1)} \right] \right) \tag{4.7}$$

$$\underset{x}{E}\left[\left(y_x^{(n+1)} - \hat{y}_x^{(n+1)}\right)^2\right] = \underset{x}{E}\left[\Sigma_x^{(n+1)} + \sigma^2\right] \tag{6}$$

$$x^{(n+1)*} = \arg\min_x\left(E\left[\left(y^{(n+1)} - \hat{y}^{(n+1)}\right)^2 \mid x^{(n+1)}\right]\right) \tag{7}$$

If we keep track of the variance of $w_x^{(n)}$ in form of $\Sigma_x^{(n)} = \operatorname{var}\left(w_x^{(n)}\right)$ and assume independence of the variance of producing or perceiving the output of the system $\sigma^{(n)2} = \operatorname{var}(\hat{y}^{(n)} \mid w^{(n)})$, our new uncertainty of the model after perceiving behavior $x^{(n)}$ is

$$\Sigma_x^{(n+1)} = \left(1 - K^{(n)}\right)^2 \Sigma_x^{(n)} + K^{(n)2}\sigma^{(n)2} \tag{4.8}$$

$$\Sigma_x^{(n+1)} = \left(1 - K^{(n)}\right)^2 \Sigma_x^{(n)} + K^{(n)2}\sigma^{(n)2} \tag{8}$$

It should be noted that the uncertainty of the model parameter associated with the skill component *x* is independent of errors, but only the number of times the component *x* has been trained. The expected squared error over all components for the next trial is

$$E\left(\operatorname{var}\left(\hat{y}^{(n+1)}\right)\right) = \sum_{x=1}^{P} \frac{\Sigma_x^{(n)}}{P} + \sigma^{(n)2} \tag{4.9}$$

$$E\left(\operatorname{var}\left(\hat{y}^{(n+1)}\right)\right) = \sum_{x=1}^{P} \frac{\Sigma_x^{(n)}}{P} + \sigma^{(n)2} \tag{9}$$

For a constant learning rate $0 < K < 1$, it follows directly from Eq. (4.8) that the expected squared error will be reduced most, if we pick a behavior $x^{(n)*}$, for which the corresponding skill component *x* and the associated parameter uncertainty $\Sigma_x^{(n)}$ is highest.

We also considered the result for non-constant learning rate. For a Kalman filter, the learning rate depends on the parameter uncertainty and measurement uncertainty itself.

$$K^{(n)} = \frac{\Sigma_p^{(n)}}{\Sigma_p^{(n)} + \sigma^{(n)2}} \tag{4.10}$$

$$K^{(n)} = \frac{\Sigma_p^{(n)}}{\Sigma_p^{(n)} + \sigma^{(n)2}} \tag{10}$$

With such flexible learning rate the updated parameter uncertainty becomes

$$\Sigma_p^{(n+1)} = \left(\Sigma_p^{(n)} - \frac{\Sigma_p^{(n)}}{\Sigma_p^{(n)} + \sigma^{(n)2}} \Sigma_p^{(n)} \right) \tag{4.11}$$

$$\Sigma_p^{(n+1)} = \left(\Sigma_p^{(n)} - \frac{\Sigma_p^{(n)}}{\Sigma_p^{(n)} + \sigma^{(n)2}} \Sigma_p^{(n)} \right) \tag{11}$$

Again, it can be easily seen that the variance is minimized when the learner picks the behavior for which the parameter uncertainty is maximal. It should be noted that we assume that there is negligible generalization between behaviors, and a homogenous variance for the output for each behavior.

4.2 METHODS

4.2.1 The hitting game

Participants held the handle of a robotic manipulandum whose movement was displayed as a small cursor (Hwang et al.). Participants were instructed to make rapid center-out movements to punch through one of four targets displayed. Subjects' proficiency was rewarded with points in randomly distributed test trials. The closer the hand cursor came to passing through the target, the greater the number of points.

There were four possible targets, arranged on an invisible circle of a 10-cm radius. The targets were positioned at -15^{o}, 75^{o}, 165^{o}, and 255^{o}. Hand position was displayed at all times as a 0.5x0.5 cm white square cursor. At the beginning of each trial, the robot brought the hand to the center mark – a stationary 0.5x0.5 cm white square. Targets were then displayed as red squares. After a short, variable delay, the targets turned white and the center mark vanished – this was the “go” signal for the center-out strike. As the strike movement crossed the invisible 10-cm radius circle, a yellow dot appeared at the crossing point to emphasize the distance between the strike and the goal as a measure of reach error. If the movement duration was too long (> 0.23 s), a blue dot appeared instead. Beyond the invisible circle an elastic force field acted as a “pillow” to absorb the strike. On some trials the robot perturbed the movement with a velocity-dependent force field (see below), which deflected the movement in clockwise or counter clockwise direction.

Because there were multiple targets present during active training trials, we needed to ascertain toward which target the subject was aiming. Therefore, after each strike when the hand hit the pillow, subjects brought their hand back to the center of their intended target. At this point the center mark reappeared and the robot brought the hand back to the center.

4.2.2 Active, passive, and test training

A trial was either a test trial or a training trial (Figure 4.1A). At the beginning of the experiment participants were instructed to pick their sequence of directions to train so that they would maximize their performance in the test trials. To further motivate participants to maximize their performance on test trials, we paid participants a monetary reward dependent on the amounts of points earned in the test trials. It was made very clear that the score only reflected their performance in the test trials, but not their performance on the training trials.

Test trials were randomly interspersed between the training trials (1 out of 4) and clearly announced with the word “test” on the screen immediately before the trial started. The computer pseudo-randomly picked the target among the 4 available to test the participants’ performance. On these trials (and only these trials), points were awarded based on performance. Performance was measured as the angular distance to the target at the point where the cursor crossed an invisible circle of 10 cm. Four critical accuracy levels were established: 5.16°, 4.49°, 3.61°, and 2.48°. For each additional accuracy level achieved, the movement was award one additional point in that trial for up to a maximum of 4 points.

Three out of four trials were training trials. The Experiments were divided into blocks of 60 trials (experiment 1) or 160 trials (experiment 2); each block was either active training or passive training. In **active training trials**, all four targets were available and the subject chose their target direction to aim. In **passive training trials**, the computer pseudo-randomly picked the target among the four available. The schedule of active and passive training trials is shown in Figure 4.1A (experiment 1) and Figure 4.1B (experiment2).

4.2.3 Experiment 1

Our objective was to ask whether errors that subjects experienced during active learning affected the subsequent choices that they made. To that aim, we considered two kinds of errors: errors that were due to a consistent perturbation, and errors that were due to an inconsistent perturbation. To produce errors in the shooting task, the robot produced a velocity dependent curl force field (Thoroughman and Shadmehr, 2000) (viscosity of 10 Ns/m) that perturbed the hand clockwise with respect to direction of the target, counterclockwise, or not at all.

The perturbations of each block of trials followed one of two patterns (Figure 4.1C). In the Constant-in-Null pattern, movements toward 3 of the 4 targets were unperturbed (Null, or “N”), while movements toward one target were perturbed with a consistent curl force field (Constant, or “C”). The C target was assigned pseudo-randomly for each block. The C target had a clockwise field for the first 30 trials of the block, and then the field switched to a counter clockwise field. A good active learning policy would have been to find the C target and continue to train mostly on that target.

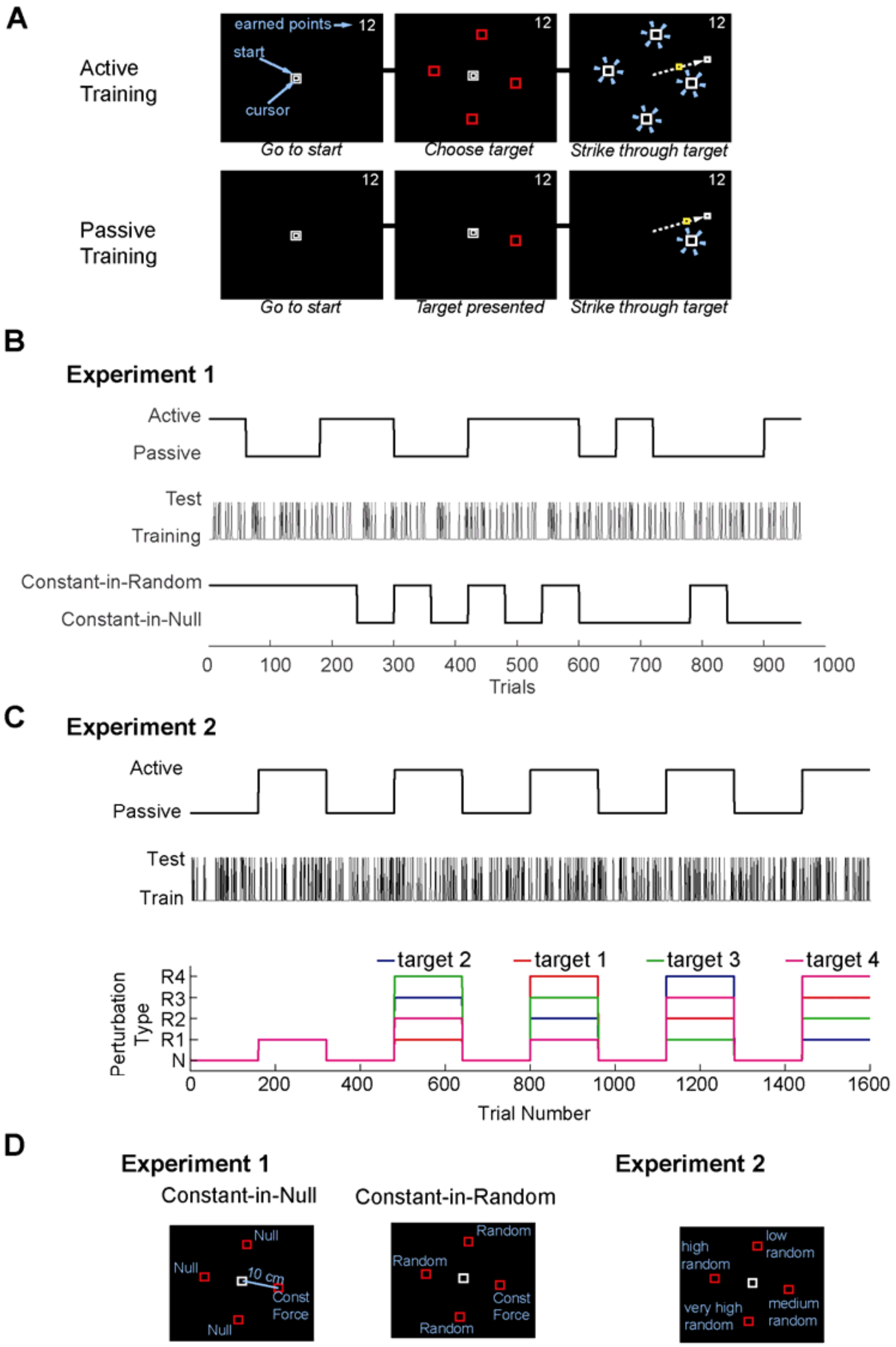

Figure 4.1 **A.** Experimental design. During active training sets, subjects chose a target to aim. During passive training sets, the computer chose a target for subjects. Participants were tested for their performance at a random direction chosen by the computer at random trials. **B.** Example experiment 1 protocol for one subject. Each subject performed 16 sets of movements with pseudo-randomly ordered conditions. **C.** Example experiment 2 protocol for one subject. **D.** Force-target association patterns. Unknown to subjects, each of the four targets was associated with a type of perturbation during a block.

In the Constant-in-Random pattern, again one target was picked to have a constant perturbation (C) that switched after 30 trials. In contrast to a Constant-in-Null block, however, a curl field was also presented during movements to the 3 remaining targets. These curl fields switched randomly between clockwise or counterclockwise fields (Random, or "R"). In these movements, the field had a random viscosity with a uniform distribution from -10 to +10 Ns/m.

4.2.4 Experiment 2 -- Perturbations of variable variances

To further explore the idea that the variance of the perturbations – i.e., the reliability of errors – influenced choices during active learning, we conducted a second experiment where the mean of the perturbations associated to the various targets were identical, but their variance differed (Figure 4.1B). Once again, four targets were available. At the start of each block (now 160 trials long), each target was assigned a curl field with a viscosity that had a mean of 10 Ns/m, but a variance that was either low (R1 target, viscosity uniformly drawn from 6 to 14 Ns/m), medium (R2 target, viscosity uniformly drawn from 2 to 18 Ns/m), high (R3 target, viscosity uniformly drawn from -2 to 22 Ns/m), and very high (R4 target, viscosity uniformly drawn from -6 to 26 Ns/m). Similar to Experiment 1, participants earned points only during the sparsely distributed test trials in which the computer randomly tested participants' performance in one of the four targets. During these test trials, the viscosity was always 10 Ns/m (C targets).

4.2.5 Softmax Regression Procedures

We modeled the probability of choosing a target using a generalized linear model. We used a multinomial extension of logistic regression –softmax regression. The probability, *p*, of selecting a target of direction *j* ($x^{(n)}=j$, $j=1\ldots P$) depends on the vector

$\mathbf{v}^{(n)}$, which in turn was a linear function of three factors: the bias ($\mathbf{\Theta}_{\text{bias}}$, 3 free parameters) modeled the preference of participants for particular targets, Θ_{repeat} modeled the preference of each subject to repeat the last movement direction $\mathbf{x}^{(n-1)}$, and finally Θ_{error} modeled the increase in probability to repeat the last movement direction, as the size of the last error increased.

$$\mathbf{v}^{(n)} = \boldsymbol{\theta}_{bias} + \theta_{repeat}\mathbf{x}^{(n-1)} + \theta_{error}\mathbf{x}^{(n-1)}\left|y^{(n-1)}\right| \tag{4.12}$$

$$\mathbf{p}^{(n)} = \begin{bmatrix} p\left(x^{(n)} = 1 \mid \mathbf{v}^{(n)}\right) \\ \dots \\ p\left(x^{(n)} = P \mid \mathbf{v}^{(n)}\right) \end{bmatrix} == \text{softmax}\left(\mathbf{v}^{(n)}\right) = \frac{\begin{bmatrix} \exp\left(v_{(1)}^{(n)}\right) \\ \dots \\ \exp\left(v_{(P)}^{(n)}\right) \end{bmatrix}}{\sum_{j=1}^{P} \exp\left(v_{(j)}^{(n)}\right)} \tag{4.13}$$

$$\mathbf{v}^{(n)} = \boldsymbol{\theta}_{bias} + \theta_{repeat}\mathbf{x}^{(n-1)} + \theta_{error}\mathbf{x}^{(n-1)}\left|y^{(n-1)}\right| \tag{12}$$

$$\mathbf{p}^{(n)} = \begin{bmatrix} p\left(x^{(n)} = 1 \mid \mathbf{v}^{(n)}\right) \\ \dots \\ p\left(x^{(n)} = P \mid \mathbf{v}^{(n)}\right) \end{bmatrix} == \text{softmax}\left(\mathbf{v}^{(n)}\right) = \frac{\begin{bmatrix} \exp\left(v_{(1)}^{(n)}\right) \\ \dots \\ \exp\left(v_{(P)}^{(n)}\right) \end{bmatrix}}{\sum_{j=1}^{P} \exp\left(v_{(j)}^{(n)}\right)} \tag{13}$$

We fitted the parameters $\mathbf{\Theta}_{\text{bias}}$, Θ_{repeat}, and Θ_{error} by maximizing the log-likelihood of the data given our model using numerical methods (Matlab fminsearch).

4.2.6 Participants

Sixteen subjects participated in experiment 1 and another sixteen in experiment 2. For experiment 1, the experiment was counter-balanced across subjects for the order of the perturbations (Constant-in-Null and Constant-in-Random) and training conditions (active and passive). Subjects were healthy, right-handed, and naïve to the purpose of the experiment. Procedures and protocols were approved by the Johns Hopkins Medicine Institutional Review Board and participants gave their written consent prior to the experiments.

4.3 RESULTS

We used a hitting-game to examine how the statistical properties of the training experience influenced the subjects' strategy in active learning (Figure 4.1A-D). The goal of the game was to become as proficient as possible at hitting a small target in one of four directions with a rapid center-out punch using a robotic manipulandum (Hwang et al., 2003)(Figure 4.1A). Participants were tested on randomly interspersed test trial, in which one direction was chosen at random for them. Their score and the financial incentive depended solely on the performance on these test trials. In between test trials, participants could train for directions of their choice (active learning) or a direction (passive learning) chosen randomly for them (Figure 4.1B). To induce errors, we applied a velocity-dependent force that pushed the hand in a *constant* direction perpendicular to the hand movement (Thoroughman and Shadmehr). Movements toward the other targets could either be unperturbed (Constant-in-Null perturbation blocks, Figure 4.1D) or be perturbed by a *random* force perpendicular to the movement (Constant-in-Random blocks).

4.3.1 Errors and last movement influence action selection during active learning

If a human active learner estimates the uncertainty about the desired output and then choose to train on a component on which they are most unsure (Eq.(4.3)), they should have a tendency *not* to practice in the same direction as the last trial. Contrary to this prediction of uncertainty-based models, our participants repeated the movement direction on the last trial with a probability of 35% (±12%) of the trials (Figure 4.2A), significantly higher than just choosing a direction at random (two-tail T-test, $t(15)=3.46$, $p<0.01$). When the participants decided to switch, they picked each of the other directions with equal probability.

The probability of repeating a direction was also strongly modulated by the amount of error that the subject experienced in the previous trial. In figure 2B, we plotted the probability of repeating the last direction as a function of the absolute size of the error on the last trial. We found that probability of repeating a direction was an increasing function of the error size (one-factor ANOVA, $F(9, 159)=5.51$, $p<<0.001$). Therefore, the larger the error in trial n, the more likely that the same direction would be repeated in trial $n+1$. This was the case for both blocks in which movement to the remaining targets were unperturbed (Constant-in-Null pattern) or perturbed by a random force field (Constant-in-Random pattern).

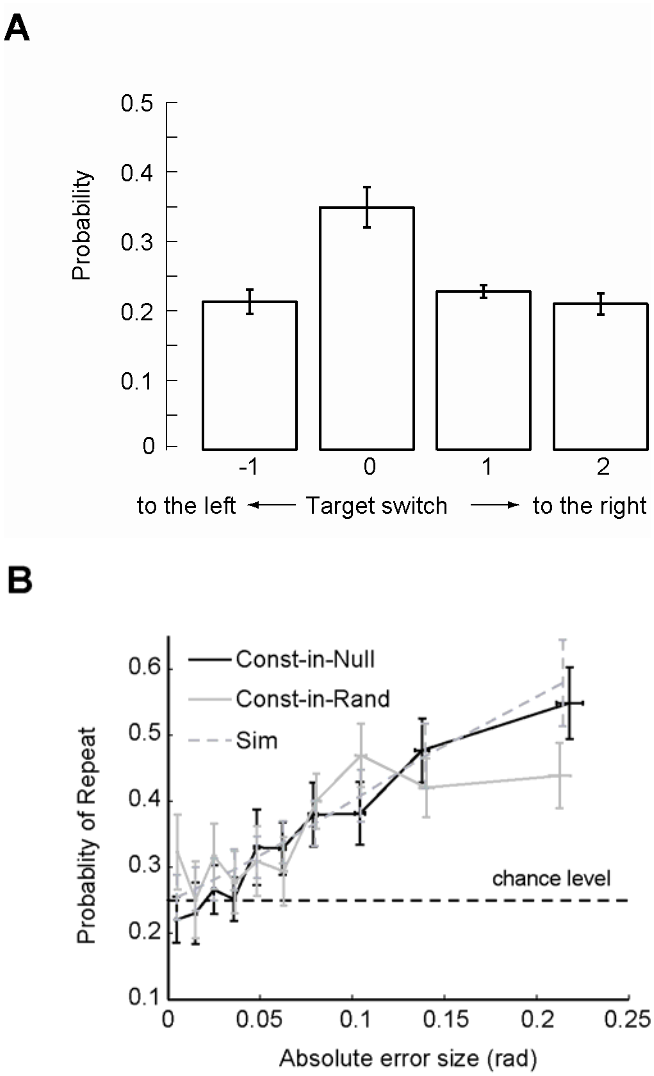

Figure 4.2 **A.** Probability of switching training directions from trial *n* to trial *n+1*. The abscissa indicates the direction and the magnitude of the switch; 0 target switch means that the same target in trial *n* was repeated in trial *n+1*. White bars indicate the average probabilities of each switch options. Error bars indicate S.E.M. across participants. **B.** Probability of repeating a target as a function of the size of the recently experienced error, separated by the type of force-target patterns. Dashed line is the average fit using the softmax regression model. Error bars indicate S.E.M. across participants.

When there were little or no errors in a trial, the probability of repeating the target approached 25%, the rate of random selection. Under uncertainty-based models, the observation of a zero error should have decreased the uncertainty about the corresponding task component. This would have then lowered the learner's probability of selecting this direction again below the probabilities of the other directions. While our data showed that error was a robust factor in encouraging repetition of a previously selected direction, a trial with a small error did not reduce the probability of the re-selection of the same movement direction below chance. Participants, therefore, clearly violated a fundamental prediction of active learning models based on model uncertainty.

To quantify these observations we estimated the contribution of error and the tendency to repeat a direction even in the absence of an error using softmax regression (a multinomial equivalent of logistic regression, see methods). The regression included a term to capture biases towards specific targets, a constant term for repeating the direction (Θ_{repeat}) independent of error size, and a term that dependent on the absolute size of the last error (Θ_{error}). From the estimated parameters, we were able to reproduce the sequence and trends of participants' choices (dotted line, Figure 4.2B). While we did not find any significant bias toward any of the four targets (one-factor ANOVA, F(3, 60) = 1.04, p=0.38), nearly all participants showed a positive Θ_{error}, indicating that they were more likely to repeat an action when a large error was encountered (two-tailed t-test, t(1,15)=3.6, p<0.01). Once we accounted for the size of the error, Θ_{repeat} was not significantly different from zero, (t(1,15)=1.0, p=0.32).

4.3.2 Relationship between learning strategy and performance

How did the participants' active learning strategies affect their performance? We looked at the correlation between strategy and performance on test trials after active learning and passive training. Since performance was determined largely by the overall proficiency of the participants, we subtracted the average errors during test trials of active learning blocks from the average errors of passive learning blocks for each participant.

The difference was then correlated with individual parameter estimates (Θ_{repeat} and Θ_{error}) from the softmax regression. A positive correlation meant that adapting these strategies hurt the participants' performance while a negative correlation means they benefited their proficiency.

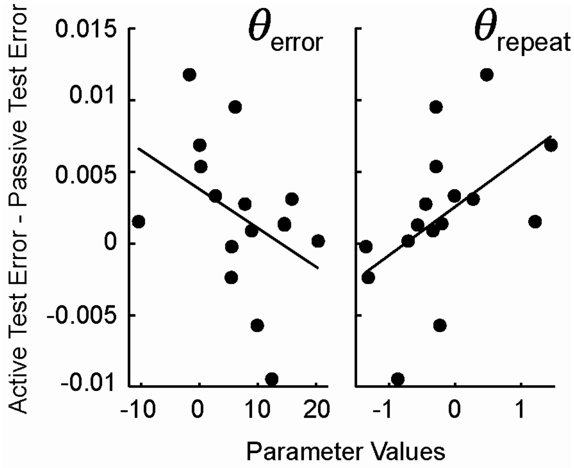

Figure 4.3 Scatter plots of the values of model parameters (abscissa) against the difference of the sizes of test trial errors between active training blocks and passive training blocks (ordinate). The straight lines are the linear regression fits. Negative slopes indicate a positive relationship between participants' sensitivities to a particular parameter in his/her active training strategy and test performance when compared to passive training.

There was a strong positive correlation between Θ_{repeat} and subsequent test performance (Figure 4.3). The more likely participants repeated a direction (after the influence of the error size has been accounted for), the worse was their test trial performance (one-tailed Spearman's correlation, r=0.60, p<0.01). This means that the violation of the optimal active learning strategy indeed hurt the performance of the participants in the active compared to the passive learning condition. We also saw that error sensitivity had a positive effect on active test performance (one-tail Spearman's correlation, r=0.48, p<0.05). Participants who sought to train in directions where their errors were big performed better in subsequent testing (Figure 4.3). Importantly, two participants that displayed error avoidance (negative values for Θ_{error}) showed poorer performance relative to their own performance after passive learning.

4.3.3 Variance of the error signal

One reason for participants to re-select the last direction even when their movement was perfect may be that they were trying to estimate the consistency of the observation. For example, if a participant found that for one target the perturbation changed randomly from trial-to-trial, a good active learning strategy would be to ignore this direction because training here could not lead to further improvement. Did the variance of the perturbations affect the choices?

To test this idea, we introduced a condition – Constant-in-Random perturbation pattern –where one direction was perturbed with a constant force field ("C" target), while the other three were perturbed with a random force field ("R" targets). To assess the influence of consistency, we attempted to match the absolute sizes of the errors of the movements toward all the directions. Since participants would adapt to the constant force field, we introduced a stronger force field to the constant target than to the random ones and flipped the perturbation direction after 30 trials (Figure 4.4A). As a result the errors for the constant target were large immediately after the onset of the block and after the switch, but lower than for the random targets in the end of each phase.

To account for these remaining differences in error, we selected trials that had similar performance in C and R targets (blue dotted trials in Figure 4.4A; paired t-test for each trial and each participants, p>0.15). For these trials, we found that the probability of choosing a "C" target was not different from choosing an "R" target (Figure 4.4B, two-tailed t-test, t(30)=-0.43, p=0.67). Therefore, variance of perturbations did not appear to influence choice.

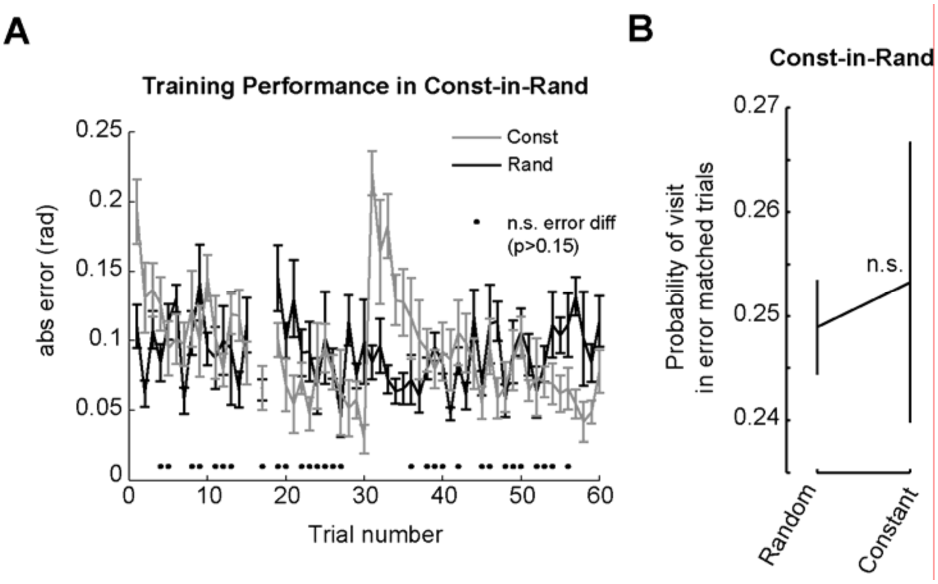

Figure 4.4 **A.** Average error for the constant and random targets as a function of trial number in active training. Graph shows initial high errors for the constant target with subsequent learning. Dots indicate the trials that were picked in which the average absolute error size were matched for constant and random targets. The vertical bars indicate the S.E.M. across participants. **B.** Overall probability of visiting a constant-force target versus a random-force target for trials of matched absolute error sizes (dotted trials in Figure 4.4A).

While we attempted to match the absolute error size in experiment 1, the average size of the force field was different. Furthermore, the number of trials in a block (60) might have been too small to allow participants to estimate variance for the different skill components. To address these concerns and test for the influence of error variance more explicitly, we designed a second experiment in which the averages of the force perturbations were matched, and participants made substantially longer sequences of movements (160 per block). The perturbations associated with each target were drawn from a distribution that had identical mean (10 Ns/m) – and we expected the means of the errors to be equal. Each target in a block had a perturbation variance that was low, medium, high, or very high (Figure 4.5B, abscissa). We labeled the corresponding targets

R1, R2, R3, and R4 (Figure 4.1C). Participants adapted to the mean force field for equally for all four variance levels (Figure 4.5A).

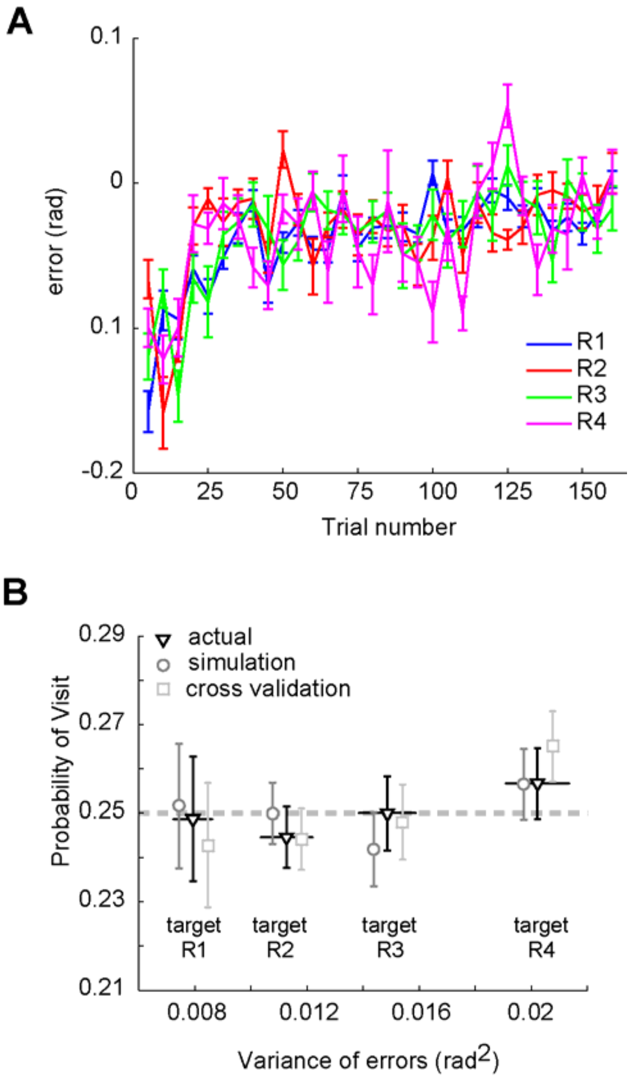

Figure 4.5 **A.** Average movement error separated by the variance of the perturbation (target R1=smallest, R4=highest variance). The vertical bars indicate the S.E.M. across participants. **B.** Overall probability of selecting a target direction with small (R1), medium (R2), high (R3), and very high (R4) perturbation variance. Plotted are the actual probabilities, simulated probabilities using parameters obtained from logistic regression of experiment 2, and simulated probabilities using parameters from experiment 1.

To determine whether the variance of the error influenced the choices during active learning, we first needed to account for the influence of the mean absolute error on choice since it increases with perturbation variance. We therefore used the same softmax regression approach to capture the strategies as for experiment 1. As in experiment 1, we found that the error size of the last trial posed a significant influence on the target selection of next trial (two-tailed t-test, $t(15)=2.54$, $p<0.05$). Participants also showed a slight tendency to repeat the last target even in the absence of error (two-tail t-test, $t(15)=1.77$, $p=0.09$), again violating the uncertainty-based models. From these parameter estimates we then predicted the probability to practice on targets of each variance level (Figure 4.5B). The observed probabilities were not significantly different from these values (two-factor ANOVA, $F(3, 93) = 0.443$, $p=0.72$). For cross-validation purposes, we also used the parameters fitted using experiment 1 data, and again, the probabilities were not significantly different ($F(3,93)=1.123$, $p=0.334$).

In summary, the results of experiment 2 provided clear evidence that participant's choices were influenced by the error of the last movement, but not by the variance of these errors over a number of trials.

4.4 DISCUSSION

The presented studies constitute to our knowledge the first systematic investigation into active learning strategies in humans. We used a task in which participants autonomously picked, trial-by-trial, a sequence of different skill components (movement directions) to practice to become proficient in the overall task. In this situation, we found two main factors dominating the choices made by the learners.

First, nearly all participants repeated the last movement when a large movement error was made in the last trial. It is intuitively clear that this strategy should lead to better learning: big errors can indicate a mismatch between the current estimate of the force perturbation and the desired output and therefore indicate the need to learn more. Participants who sought out movement directions with large errors were more successful in subsequent test trials; participants who avoided errors were comparatively less successful.

The second and a surprise finding is that even after perfect movements (i.e. no errors), participants has a tendency to repeat the last movement direction. Current active learning algorithms in machine learning have the opposite behavior: making an observation near *x* reduces the model uncertainty in the neighborhood of *x* and therefore reduces the probability of re-selecting this component in the next training trial. Humans did not follow this optimal strategy during active learning. When no error was observed in a movement direction – a situation in which the estimated uncertainty of the output should have been reduced, they did not avoid this movement direction. Thus, the uncertainty-based active learning models do not provide a good description of natural human behavior in the context of active learning.

This strategy of re-selecting the target was sub-optimal, as it correlated negatively with subsequent performance. Indeed, as most participants showed this behavior, the performance after active learning was not better, or even slightly worse, than after passive learning, in which a trial was picked at random.

The observation that alternating the practiced skill during training increases subsequent performance agrees with a host of behavioral studies that used passive, intermixing training schedules. In a task where participants were asked to learn to perform three different punch styles, Shea and Morgan (Shea and Morgan, 1979) showed that people who trained with a random schedule – practice trials on all three styles were conducted intermittently – retained their performance better after 10 minutes and after 10 days, when compared to people who trained one style at a time. Similar results emphasizing the benefits of concurrent and intermixed training of several skills were found in pistol shooting (Keller et al., 2006), surgery training (Brydges et al., 2007), and

3D spatial orienting (Shebilske et al., 2006). Interestingly, in a task where participants were to complete an arm movement that consisted of three segments of distinct speeds, participants learned to time their movement segments more accurately when feedback of their accuracy was given for one random segment of the three, and less accurately when feedback was given for only one segment several trials in a row (Lee and Carnahan, 1990). An uncertainty-based active learning strategy reduces the uncertainty of the desired output more when error feedback is obtained across the whole task space by quickly alternating from task component to component, than continuing the components one type at a time in blocks.

Finally we found that, while participant's choices were dependent on the absolute size of the last error, they were insensitive to the cross-trial variance of these errors. Under the assumption that this variance cannot be reduced through training, a better learning strategy would have been avoiding task components that have a very high variance because very little can be learned here. The results can imply one of two things. First, participants may not estimate the variance of the error signal over multiple trials. This is congruent with recent results that showed that variance of reward values does not influence decisions (Daw et al., 2006). Alternatively, it may imply that participants were trying to reduce the variance of the errors for movement directions with high trial-to-trial perturbation variance. However, they were not successful in doing so as the behavioral results of experiment 2 showed that for the trial-to-trial variance in performance remained large for these movements (Figure 4.5A).

Our results highlight that humans do not always choose the optimal learning strategy when given the chance to select their own training sequence. We showed that we can separate aspects of the strategy that improved and aspects that did not improve performance. Our findings open up the possibility to design adaptive algorithms that would lead to better learning than random training. Specifically, the present results would predict faster learning and better performance when movement directions are repeated only when the last error was in that direction was large. Such adaptive training algorithms may play an important role in robot-based rehabilitation training after stroke or developmental disabilities.

CHAPTER 5

FROM MODELS TO LEARNING IMPROVEMENT – A DISCUSSION

Computational models of learning have become important tools for a critical examination of the underlying laws that govern motor learning, since, by definition, mathematical models cannot be self-conflicting and ambiguous. The goal of this thesis is to use theoretical models as a foundation to extend our current understanding of the computational procedures required to achieve each of these components. The thesis focuses on the computational steps that modulate with the temporal structure of training and the statistical characteristics of the learning history and seeks behavioral data that unambiguously verify the validity of these theoretical components of learning. Specifically, the thesis aims to apply novel understandings of motor learning to facilitate the enhancement of motor learning. In this thesis, I made attempts to identify components of motor adaptation acquisition, motor adaptation retention, and motor decision that allow for learning enhancement.

In the hierarchy of motor system, it is thought that motor learning mechanisms supersede motor control mechanisms. However, it was not until the acceptance of the concept of a dynamic internal model that a link between control and learning was made. A fundamental principle of motor learning is that the degree of success relies on the amount of practice. In analogy to practice, therefore, it is the iterative updates of the internal model that completes the computational steps of motor learning. Motor learning involves the learner to first identify the structures of the task and second, estimate the optimal parameters necessary to describe the task. The former can be thought as a system identification problem. However, when the task structure is fixed, then in computational terms, motor adaptation is about finding the optimal parameter of the model of the task.

Previous models of learning adapt to change in the environment by a gradient descent algorithm in an attempt to minimize the error between the observation and desired motor output. In Chapter 2, I showed that this adaptation depends not just on error, but also on passage of time between the two movements. It was observed that subjects learned a reaching task faster, i.e., with fewer trials, when the inter-trial time intervals (ITIs) were lengthened. I hypothesized two computational mechanisms that could have accounted for this. First, learning could have been driven by a Bayesian process where the learner assumed that errors are due to perturbations that have multiple timescales. In theory, longer ITIs can produce faster learning because passage of time might increase uncertainty, which in turn increases sensitivity to error. Second, error in a trial may result in a trace that decays with time. If the learner continued to sample from the trace during the ITI, then adaptation would increase with increased ITIs. The two models made separate predictions: The Bayesian model predicted that when movements are separated by random ITIs, the learner would learn most from a trial that followed a long time interval. In contrast, the trace model predicted that the learner would learn most from a trial that preceded a long time interval. We performed two experiments to test for these predictions and in both experiments found evidence for the trace model. We suggest that motor error produces an error memory trace that decays with a time constant of about 4 seconds, continuously promoting adaptation until the next movement.

In Chapter 3, I showed that motor learning is more than estimating the proper scaling of motor outputs. Rather, motor learning also encompasses generating rules to describe the task itself. Here we used a similar force adaptation task with hand-punching movements to examine the retention of human adaptation to different task structure. The task structure was varied by gradually or suddenly introducing the force perturbation. We fixed the task structure within a training block, but varied the task structure across two training blocks separated by a washout block. In the first training when the participants were initially exposed to the perturbation with the task structure fixed, the gradualness of the first training significantly improved the retention of adaptation. The current time-invariant two-state model of motor adaptation predicted that errors caused by overcompensation in the extended null-force washout block would actively engage the learner in de-adaptation and completely wash out adaptation. However, while this adaptation appeared to have washed out in an extended null-force block, we found that participants' first experience with the perturbation continued to influence the rate of decay in their future retention when they were re-exposed to the perturbation in a subsequent retraining block. Despite the intervening washout block, the gradualness of the first training alone was sufficient to promote future retention of adaptation regardless of the rate of perturbation exposure in retraining even when the first training was in the opposite force field. Therefore, subjects generated separate sets of rules to describe the two task structure. Separate rules that govern the retention of adaptation, therefore, followed. As a result, the first exposure to the task created an acquisition and retention bias for subsequent learning scenarios. Together, these observations suggested that the structure of the learning process is not time-invariant as previously assumed and the human motor retention process possibly receives a top-down control based on the history of training. In practice, because the gradualness of first exposure to training promotes future retention of adaptation, we can exploit this structural learning bias to enhance learning.

When we learn a new skill without a coach, we become "active learners". We choose the specific component of the task to train on (e.g. iron, driver, putter, etc.). In Chapter 4, I asked what guides our selection of the training sequence. How do choices that people make compare to choices made by machine learning algorithms that attempt to optimize performance? I asked subjects to learn the novel dynamics of a robotic tool while moving it in four directions. They were instructed to choose their practice directions to maximize their performance in subsequent tests. I found that their choices were strongly influenced by motor errors: subjects tended to immediately repeat an action if that action had produced a large error. This strategy was correlated with better performance on test trials. However, even when participants performed perfectly on a movement, they did not avoid repeating that movement; they tended to repeat that choice with the same or higher probability as other movements. This behavior led to sub-optimal performance. It also violated a strong prediction of current machine learning algorithms, which solve the active learning problem by choosing a training sequence that will maximally reduce the learner's uncertainty about the task. While I showed that these algorithms do not provide an adequate description of human behavior, these results suggested that it should be possible to improve human motor learning by helping people choose their training sequence.

The individual features of motor learning identified in Chapter 2 and 3 can be used alone or in combination in motor learning applications – for example, rehabilitation in post-stroke recovery. Importantly, the principles identified in this thesis augment our understanding in how practice improves the degree the motor performance. Since most if not all motor learning paradigms rely on repeated movements, manipulation in the temporal distribution and task structure bias on the trial-to-trial practice schedule should be applicable across many training protocols.

The introduction of the thesis mentioned that animals evolved to develop a sensory-motor system to sense, manipulate and interact with their surroundings. The results of the active learning study re-emphasize the interactive component of motor learning that is often ignored in the computational motor control literature. In Chapter 3, we observed that people can abstract task structures implicitly without given explicit description or instruction to do so. That is, people possess the ability to generate new rules to describe their environment even in a passive learning environment. It is natural to think, then, with the freedom to actively direct their learning strategies, such as in the case described in Chapter 4, animals have the tools to aggressively collect data about their environment. In Chapter 4, however, subjects needed to learn task structure while optimally estimate motor outputs. While unbiased human learners do not appear to be optimal in their active learning strategy in this dual-task, whether people become more effective extracting structural information alone with active learning paradigms remains to be elucidated. When the sensory system is damaged (e.g. as a result of parietal stroke), it should be expected that the efficacy of the active learning component would be impaired.

Another common theme among the studies presented in the current thesis is that a robotic manipulandum was used. In handling the manipulandum, the subject must learn not just the dynamical disturbances (usually in the form of a viscous curl force field), but also a change in the inertial properties of his or her arm because of the additional physical contact to the robot itself. In all studies, we have allowed time for the subjects to familiarize themselves with the handling of the manipulandum through introductory practice prior to manipulating experimental variables. The presence of after-effects after adapting to the viscous force field is strong evidence that the nervous system constructs a new mapping between limb state and muscle outputs, and it has been shown that learning of one movement type (e.g. direction) can generalize to another. If a common internal model serves to represent limb dynamics across movement tasks, adaptation in one task should be generalized to another. However, it is unclear whether a common internal model of limb dynamics is used not only across movement types but also across movement tasks. In the case of these studies, the internal model would also encompass a component that accounts for the dynamics of the robot arm that may not necessarily be appropriate for other tasks other than the original training. This identifiable differentiator would be important in rehabilitation therapy because it undermines the generalization of skills learned in one training task in clinics to new tasks at home. The results in Chapter

3, however, raise the possibility to promote generalization by providing structural commonality across tasks. This possibility is especially interesting in rehabilitation researches that investigate motor learning in patients with hemiparesis with robotic assistive devices.

REFERENCES

Aboukhalil A, Shelhamer M, Clendaniel R (2004) Acquisition of context-specific adaptation is enhanced with rest intervals between changes in context state, suggesting a new form of motor consolidation. Neurosci Lett 369:162-167.

Bhushan N, Shadmehr R (1999) Computational nature of human adaptive control during learning of reaching movements in force fields. Biol Cybern 81:39-60.

Bizzi E (1993) Intermediate representations in the formation of arm trajectories. Curr Opin Neurobiol 3:925-931.

Bock O, Thomas M, Grigorova V (2005) The effect of rest breaks on human sensorimotor adaptation. Exp Brain Res 163:258-260.

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

Brydges R, Carnahan H, Backstein D, Dubrowski A (2007) Application of motor learning principles to complex surgical tasks: searching for the optimal practice schedule. J Mot Behav 39:40-48.

Cohn DA, Ghahramani Z, Jordan MI (1996) Active Learning with Statistical Models. J Artificial Intelligence Research 4:129-145.

Coltz JD, Johnson MT, Ebner TJ (1999) Cerebellar Purkinje cell simple spike discharge encodes movement velocity in primates during visuomotor arm tracking. J Neurosci 19:1782-1803.

Coltz JD, Johnson MT, Ebner TJ (2000) Population code for tracking velocity based on cerebellar Purkinje cell simple spike firing in monkeys. Neurosci Lett 296:1-4.

Comas D, Petit F, Preat T (2004) Drosophila long-term memory formation involves regulation of cathepsin activity. Nature 430:460-463.

Commins S, Cunningham L, Harvey D, Walsh D (2003) Massed but not spaced training impairs spatial memory. Behav Brain Res 139:215-223.

Criscimagna-Hemminger SE, Donchin O, Gazzaniga MS, Shadmehr R (2003) Learned dynamics of reaching movements generalize from dominant to nondominant arm. J Neurophysiol 89:168-176.

Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. Nature 441:876-879.

Dayan P, Yu AJ (2006) Phasic norepinephrine: A neural interrupt signal for unexpected events. Network 17:335-350.

de Haan A, Jones DA, Sargeant AJ (1989) Changes in velocity of shortening, power output and relaxation rate during fatigue of rat medial gastrocnemius muscle. Pflugers Arch 413:422-428.

Diedrichsen J, Hashambhoy Y, Rane T, Shadmehr R (2005) Neural correlates of reach errors. J Neurosci 25:9919-9931.

Donchin O, Francis JT, Shadmehr R (2003) Quantifying generalization from trial-by-trial behavior of adaptive systems that learn with basis functions: theory and experiments in human motor control. J Neurosci 23:9032-9045.

Ebbinghaus H (1964) Memory: A contribution to experimental psychology. New York: Dover.

Efron B, Tibshirani RJ (1993) An introduction to the bootstrap monographs on statistics and applied probability. New York: Chapman and Hall.

Feldman AG (1986) Once more on the equilibrium-point hypothesis (lambda model) for motor control. J Mot Behav 18:17-54.

Fiala JC, Grossberg S, Bullock D (1996) Metabotropic glutamate receptor activation in cerebellar Purkinje cells as substrate for adaptive timing of the classically conditioned eye-blink response. J Neurosci 16:3760-3774.

Francis JT (2005) Influence of the inter-reach-interval on motor learning. Exp Brain Res:1-4.

Fusi S, Drew PJ, Abbott LF (2005) Cascade models of synaptically stored memories. Neuron 45:599-611.

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

Giszter SF, Mussa-Ivaldi FA, Bizzi E (1993) Convergent force fields organized in the frog's spinal cord. J Neurosci 13:467-491.

Gormezano I (1966) Classical conditioning. In: Experimental methods and instrumentation in psychology (Sidowski, J. B. Ed.). New York: McGraw-Hill.

Han JS, Gallagher M, Holland P (1998) Hippocampal lesions enhance configural learning by reducing proactive interference. Hippocampus 8:138-146.

Huang VS, Shadmehr R (2007) Evolution of motor memory during the seconds after observation of motor error. J Neurophysiol 97:3976-3985.

Hwang EJ, Shadmehr R (2005) Internal models of limb dynamics and the encoding of limb state. J Neural Eng 2:S266-278.

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

Hwang EJ, Smith MA, Shadmehr R (2006b) Adaptation and generalization in acceleration-dependent force fields. Exp Brain Res 169:496-506.

Hwang EJ, Donchin O, Smith MA, Shadmehr R (2003) A gain-field encoding of limb position and velocity in the internal model of arm dynamics. PLoS Biol 1:E25.

Isabel G, Pascual A, Preat T (2004) Exclusive consolidated memory phases in Drosophila. Science 304:1024-1027.

Jaric S, Radovanovic S, Milanovic S, Ljubisavljevic M, Anastasijevic R (1997) A comparison of the effects of agonist and antagonist muscle fatigue on performance of rapid movements. Eur J Appl Physiol Occup Physiol 76:41-47.

Josselyn SA, Shi C, Carlezon WA, Jr., Neve RL, Nestler EJ, Davis M (2001) Long-term memory is facilitated by cAMP response element-binding protein overexpression in the amygdala. J Neurosci 21:2404-2412.

Kagerer FA, Contreras-Vidal JL, Stelmach GE (1997) Adaptation to gradual as compared with sudden visuo-motor distortions. Exp Brain Res 115:557-561.

Keller GJ, Li Y, Weiss LW, Relyea GE (2006) Contextual interference effect on acquisition and retention of pistol-shooting skills. Percept Mot Skills 103:241-252.

Kettner RE, Mahamud S, Leung HC, Sitkoff N, Houk JC, Peterson BW, Barto AG (1997) Prediction of complex two-dimensional trajectories by a cerebellar model of smooth pursuit eye movement. J Neurophysiol 77:2115-2130.

Kojima Y, Iwamoto Y, Yoshida K (2004) Memory of learning facilitates saccadic adaptation in the monkey. J Neurosci 24:7531-7539.

Kording KP, Wolpert DM (2004) Bayesian integration in sensorimotor learning. Nature 427:244-247.

Kording KP, Tenenbaum JB, Shadmehr R (2007) The dynamics of memory as a consequence of optimal adaptation to a changing body. Nat Neurosci 10:779-786.

Kording KP, Tenanbaum, T., Shadmehr, R. (2007) The dynamics of memory as a consequence of optimal adaptation to a changing body. Nature Neuroscience in press.

Krakauer JW (2006) Motor learning: its relevance to stroke recovery and neurorehabilitation. Curr Opin Neurol 19:84-90.

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

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

Lackner JR, Dizio P (1998) Gravitoinertial force background level affects adaptation to coriolis force perturbations of reaching movements. J Neurophysiol 80:546-553.

Lackner JR, DiZio PA (2003) Adaptation to rotating artificial gravity environments. J Vestib Res 13:321-330.

Lebron K, Milad MR, Quirk GJ (2004) Delayed recall of fear extinction in rats with lesions of ventral medial prefrontal cortex. Learn Mem 11:544-548.

Lee TD, Carnahan H (1990) When to provide knowledge of results during motor learning: scheduling effects. Human Performance 3:87-105.

Locatelli F, Maldonado H, Romano A (2002) Two critical periods for cAMP-dependent protein kinase activity during long-term memory consolidation in the crab Chasmagnathus. Neurobiol Learn Mem 77:234-249.

Maldonado H, Romano A, Tomsic D (1997) Long-term habituation (LTH) in the crab Chasmagnathus: a model for behavioral and mechanistic studies of memory. Braz J Med Biol Res 30:813-826.

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

Miall R (1995) Motor control, biological and theoretical.

Michel C, Pisella L, Prablanc C, Rode G, Rossetti Y (2007) Enhancing visuomotor adaptation by reducing error signals: single-step (aware) versus multiple-step (unaware) exposure to wedge prisms. J Cogn Neurosci 19:341-350.

Mussa-Ivaldi FA, Hogan N, Bizzi E (1985) Neural, mechanical, and geometric factors subserving arm posture in humans. J Neurosci 5:2732-2743.

Nezafat R, Shadmehr R, Holcomb HH (2001) Long-term adaptation to dynamics of reaching movements: a PET study. Exp Brain Res 140:66-76.

Platt ML, Glimcher PW (1997) Responses of intraparietal neurons to saccadic targets and visual distractors. J Neurophysiol 78:1574-1589.

Raymond JL, Lisberger SG (1998) Neural learning rules for the vestibulo-ocular reflex. J Neurosci 18:9112-9129.

Savion-Lemieux T, Penhune VB (2005) The effects of practice and delay on motor skill learning and retention. Exp Brain Res 161:423-431.

Schall JD, Thompson KG (1999) Neural selection and control of visually guided eye movements. Annu Rev Neurosci 22:241-259.

Scharf MT, Woo NH, Lattal KM, Young JZ, Nguyen PV, Abel T (2002) Protein synthesis is required for the enhancement of long-term potentiation and long-term memory by spaced training. J Neurophysiol 87:2770-2777.

Scheidt RA, Dingwell JB, Mussa-Ivaldi FA (2001) Learning to move amid uncertainty. J Neurophysiol 86:971-985.

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.

Shadmehr R, Arbib MA (1992) A mathematical analysis of the force-stiffness characteristics of muscles in control of a single joint system. Biol Cybern 66:463-477.

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

Shea JB, Morgan RL (1979) Contextual interference effects on the acquisition, retention and transfer of a motor skill. J Exp Psychol (Hum Learn) 3:179-187.

Shebilske WL, Tubre T, Tubre AH, Oman CM, Richards JT (2006) Three-dimensional spatial skill training in a simulated space station: random vs. blocked designs. Aviat Space Environ Med 77:404-409.

Smith MA, Shadmehr R (2005) Intact ability to learn internal models of arm dynamics in Huntington's disease but not cerebellar degeneration. J Neurophysiol 93:2809-2821.

Smith MA, Ghazizadeh A, Shadmehr R (2006) Interacting Adaptive Processes with Different Timescales Underlie Short-Term Motor Learning. PLoS Biol 4:e179.

Smith MA, Shadmehr R. (2004) Modulation of the rate of error-dependent learning by the statistical properties of the task. In: Advances in Computational Motor Control. San Diego.

Staddon JE, Higa JJ (1996) Multiple time scales in simple habituation. Psychol Rev 103:720-733.

Staddon JE, Chelaru IM, Higa JJ (2002) Habituation, memory and the brain: the dynamics of interval timing. Behav Processes 57:71-88.

Sutton RS, Barto AG (1981) Toward a modern theory of adaptive networks: expectation and prediction. Psychol Rev 88:135-170.

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

Thoroughman KA, Shadmehr R (2000) Learning of action through adaptive combination of motor primitives. Nature 407:742-747.

Todorov E (2004) Optimality principles in sensorimotor control. Nat Neurosci 7:907-915.

Todorov E, Shadmehr R, Bizzi E (1997) Augmented Feedback Presented in a Virtual Environment Accelerates Learning of a Difficult Motor Task. J Mot Behav 29:147-158.

Tully T, Preat T, Boynton SC, Del Vecchio M (1994) Genetic dissection of consolidated memory in Drosophila. Cell 79:35-47.

Wainscott SK, Donchin O, Shadmehr R (2005) Internal models and contextual cues: encoding serial order and direction of movement. J Neurophysiol 93:786-800.

Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. Neuron 46:681-692.