State-Space Models of Online Acquisition in Motor Memory

Ali Ghazizadeh Ehsaei

A thesis submitted to Johns Hopkins University in conformity with the requirements for the degree of Master of Science in Engineering

Baltimore, Maryland
July-2005

Contents

Preface:

The aim of this thesis is to try to understand the underlying mechanisms involved in human skill acquisition and motor learning. Skill acquisition and motor learning in general refers to all those learning activities which share a main theme, that is to learn specific movement control strategies to perform a task in order to accomplish a conscious objective. To model human motor learning we are faced with an intricate problem which involves the coordination and control of various limbs in the body and thus different regions of the brain.

In order to overcome the difficulties of the complex problem of human motor learning, in particular the motor learning related to making a movement with the arm, we have adopted a standard approach which has proven useful in all modern sciences for which the scientist tries to design controlled experimental procedures to test his basic, fundamental hypotheses and use them as a proxy to break down and model his seemingly sophisticated world of interest by the way of induction which in our case is about human motor adaptation.

The experiments used to investigate the question of motor learning here, consist of point-to-point reaching movements using a robotic manipulandum. The robotic arm has been used to simulate the effects of the environment producing perturbations and force patterns to be learned by the subjects. The reaching movements in the face of these challenges in

our virtual environment, are supposed to represent our daily struggle to learn motor skills needed for our everyday activities and survival.

In the first chapter we are going to study the state-space model of trial-to-trial motor learning. It will be shown how various versions of this simple model can capture some of the most important features of motor learning. Further it will be discussed how this model can explain the concept of generalization in motor learning.

Chapter Two starts with a list of phenomena unexplained by the current state-space model and will introduce our extension over the existing state-space model arming it to tackle the current shortcomings of learning models. We hypothesized a Multi-Timecourse model consisting of two distinct learning processes. One process adapts slowly but retains information well, while the other adapts quickly but has poor retention. To further distinguish our model from other Single and Multi-State models, we designed a point-to-point reaching experiment using a well studied velocity dependent curl force field. In this paradigm we observed a reemergence of the old, well trained motor behaviors after an apparent washout period (termed *spontaneous recovery*). This phenomenon is predicted by our Multi-Timecourse model while the current models of motor adaptation fail to account for it.

We will also have a short discussion of the various alternatives in formulating the Multi-Timecourse model and the physiological evidence there is to support them. It will be shown that although these different formulations may invoke different interpretations of

how motor memory functions, they can be considered as mathematical equivalents of each other and as such, behavioral experiments like the ones done in this study are hardly sufficient to distinguish among them. The chapter is going to conclude with a summary of results and future questions that needed to be addressed in this context and the possible suggestions for further experiments.

Apart from the differences in their temporal contributions to motor adaptation, the fast and slow learning processes may have different generalization patterns. We are going to investigate this possibility in chapter Three. This difference in generalization patterns can manifest itself in distinct generalization patterns across different movement directions and arm configurations using one hand (*Within arm transfer of learning*) as well as the generalization of the learned skill from one hand to the other (*Between arm transfer of learning*). In chapter Three we will investigate the second possibility. We designed an experiment to study the transfer of the learned force field from one arm to the other. We will use this study to show if there is any evidence for a difference in bimanual transfer properties of the two processes.

Chapter I

Single-State Models of Learning

Understanding how the brain learns to execute movements and modify them in the face of a dynamically changing environment still remains a fundamental challenge in neuroscience. This difficulty arises for several reasons including:

1. The structure of the controller and the role of different controlling modules at the level of the Central Nervous System are highly involved. A large part of the cortex and nucleus including anatomically and functionally diverse regions are contributing to movement generation and control. Neurophysiological and imaging studies as well as psychophysics as helpful as they were to our understanding of the brain and its various regions, have yet a long way to determine some of the most essential questions about the exact role of areas like Cerebellum, Basal ganglia and Motor cortex in control and learning motor skills not to mention to provide us with a general exhaustive model of human motor control.

2. The structures of the plants used in making movements e.g. arms, legs, tongue etc. are not trivial and usually involve large degrees of freedom as well as a complicated anatomy which makes them hard to study in a full blown analysis.

3. In trying to analyze the human motion we are not only faced with a moving and a thinking machine. We have to take into account all the inter-subject variability due to individual differences caused by sensory-motor diversities, attention modulation of behavior, fatigability, emotional states, cognitive abilities etc. all of which makes our already complex case study even more unpredictable and therefore hard to model.

Despite all of these complications stated above there is a surprising level of consistency in the learning abilities and motor behavior of humans. In case of our point-to-point reaching movements it has been reported by several studies that subjects tend to make roughly straight movements with a velocity profile that is smooth and bell shaped (Atkeson and Hollerbach 1985; Flash and Hogan 1985). Moreover the way the subjects move their gaze before and during the movement as well as the way they tend to hold their limbs and their muscle activation all follow stereotypical patterns (Georgopoulos et al., 1981; Soechting and Lacquaniti, 1981; Neggers and Bekkering 2001). It is as though given the constraints of our control systems, our plants and our environment, there are some hidden laws which govern our general perceptions and actions. This assumption along with the intuition that these hidden laws should make our performance tuned to obtain maximum gain possible have led many of the scientists to try to come with optimizing strategies for each task and compare it with human performances to see whether and to what extent they match. Examples of these include the theory of minimum end point variance which tries to model the features of point to point reaching movements by assuming a signal dependent noise and a goal which is to minimize the

end point error or the theory of optimum Bayesian sensory motor integration which tries to model the best strategy in using sensory evidence for executing motor actions.

But can we also think of optimal principles that guide human motor learning?

Let's take the relatively simple task of point to point reaching movements as an example of a motor action. If we assume that moving on a straight path form point A to point B with a smooth velocity is what is optimal and desired in this task, then any internally or externally generated deviations from this trajectory[1] can be thought of as an error from the desired trajectory[2] which needs to be corrected. In this way one can use the errors on previous movements to update his movement strategies. This is the backbone of the online trial to trial learning which we are going to discuss in this chapter.

1. 1 Feed-back vs. Feed-forward learning:

Before moving on to our discussion of online trial-to-trial learning we need to briefly look at the general structure of a feedback controller. This will help us to understand which modules in the controller are targeted by trial to trial learning.

The figure below shows the general block diagram for any closed loop controller. A closed loop controller in general consists of a feed forward block which transforms the

[1] Trajectory in a movement includes position x and velocity $\dot{x}$

[2] Here of course the assumption is that the desired trajectory has not changed due to the new task requirements

inputs of the system into the desired outputs and a feedback block which takes the actual output of the system and uses it to modify the input in order to compensate for the unwanted changes in the output. A system with a well designed negative feedback can give a more robust output in the face of noise and disturbances.

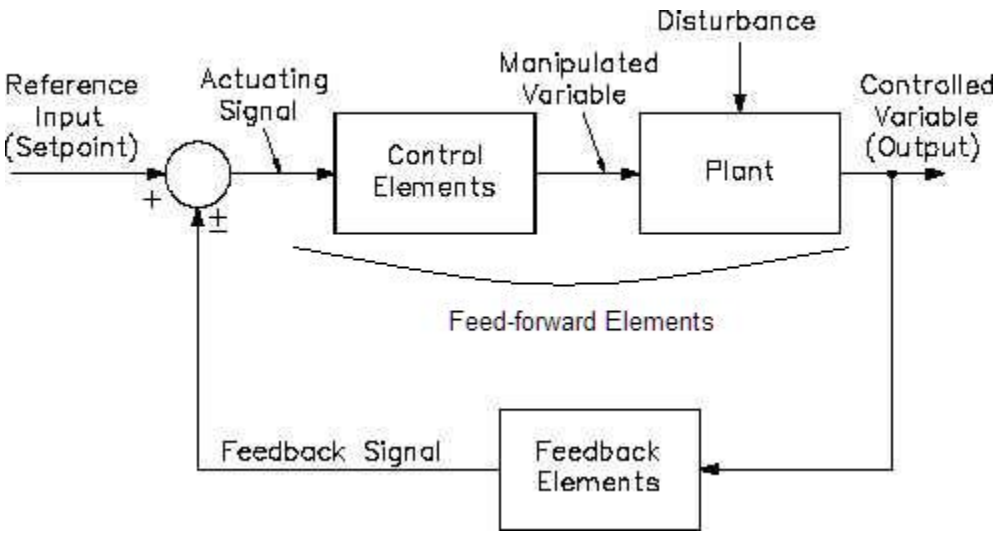

Figure 1.1

In order to perform better in a given task subjects can learn to change the properties of the feed-forward as well as the feed-back module. In the human motor learning literature the feed-forward component of the controller is comprised of two parts. First part includes the dynamics of the arm plus the dynamics imposed by the environment e.g. the robotic arm and its force patterns which are collectively called the “plant”. Second part is the model that the subjects construct from their environment which is called the “Internal model”. It is clear that learning in the feed-forward block can only happen in the latter part i.e. the “internal model” which is plastic enough to change its properties. We are going to investigate how humans can change their internal model of environment based

on their previous experiences. Yet the fact is that subjects can also learn to react more efficiently to the perturbations caused by the environment during the movement by modifying their feedback responses. This phenomenon which is called “Feedback Learning” has been observed in motor learning. For example it is well known that subjects can increase the stiffness of their arms when encountered by a changing environment which is a simple example of “Feedback Learning”.

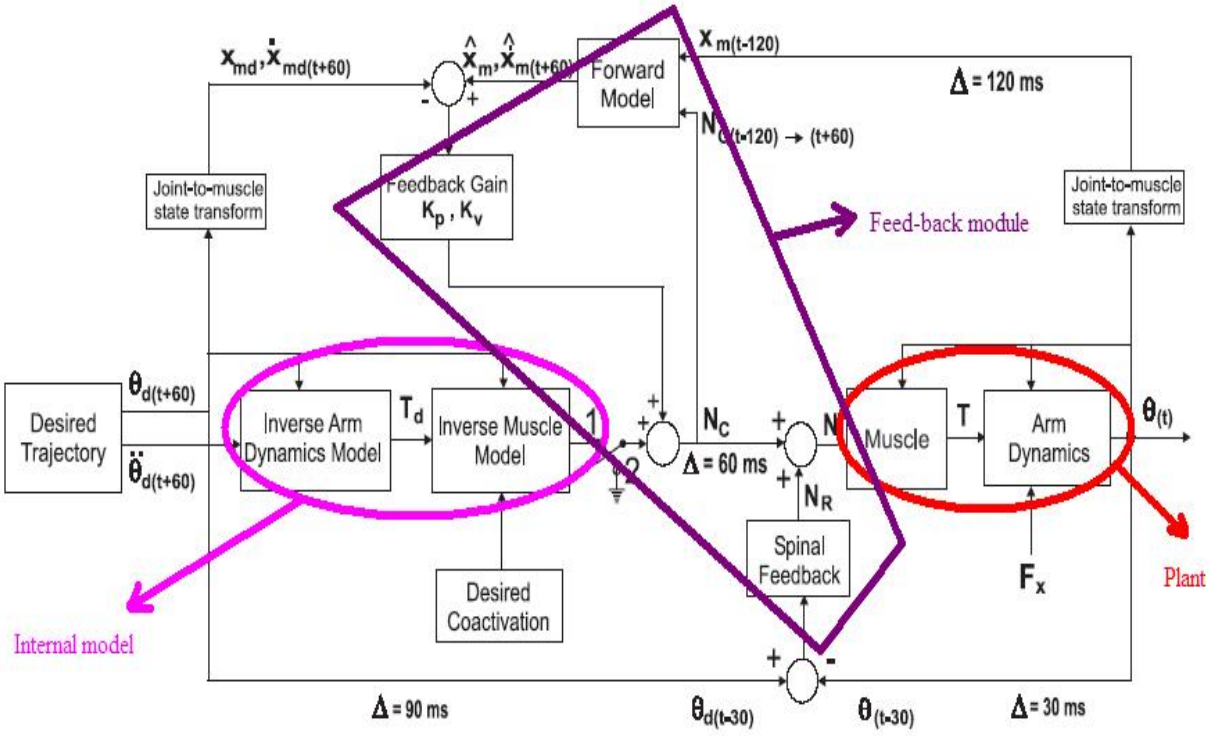

Figure 1.2

As can be seen in figure 1.2 learning can happen in almost all the blocks that are not external to the subject control. Yet we are going to focus mainly on the “Internal model” and its modification through trial to trial learning.

1.2 How to learn: "Quadratic Cost Function"

Another important concept in learning is the way one wants to penalize the deviations from the desired goal. This is a crucial point since the way one defines the costs for various errors in performance can affect how that person will learn from those errors. A recent study by Kording et al. 2004 has looked at different cost functions in an aiming task.

Their results seems to suggest that for a wide range of errors a quadratic cost function can be a good approximation for what people use to penalize errors. A quadratic cost function is defined below:

$$Cost = \frac{1}{N}\sum_{i=1}^{N}(Zd_i - Z_i)^2 \qquad 1.1$$

Zd_i: Desired output on trial i Z_i: Actual output on trial i N: Number of trials

This cost function has been in use for modeling human and machine batch and online learning algorithms by default since it is mathematically tractable and the results of experimental data seem to match the criteria to minimize this cost function well.

For a given set of *Zd_i's* if we take *Z* to be constant then using this cost function we can see that the best constant output to minimize the quadratic cost in equation 1.1 for given *N* trials of $\{Zd_1,...,Zd_N\}$ is the mean of the *Zd_i's*.

$Cost = \frac{1}{N}\sum_{i=1}^{N}(Zd_i - Z_{opt})^2$ is minimum only if $\frac{d(Cost)}{d(Z_{opt})} = 0$ so we have:

$\frac{d(Cost)}{d(Z_{opt})} = \frac{-2}{N}\sum_{i=1}^{N}(Zd_i - Z_{opt}) = 0$ which results in:

$$Z_{opt} = \frac{1}{N}\sum_{i=1}^{N} Zd_i = \overline{Z}_N \qquad \textit{Global mean-track model (batch)} \qquad 1.2$$

In addition Z_{opt} is our best possible estimate using the past *N* trials for the mean over all possible *Zd's,* and as such this Z_{opt} can be used as the best unbiased guess for the next input on trial *N+1*. This is the approach that distinguishes the batch and online learning algorithms. In batch learning of the mean we have to wait for all possible data points to be able to calculate Z_{opt} while in the online learning we update our measure of the mean using each single data point. Next section will show different possible formulations of online learning and their interpretations as well as their possible advantages over one another.

1.3 Learning an Internal model on trial-to-trial bases

Finding the optimum output using equation 1.2 is an example of general learning strategies that are called batch learning algorithms in which one has to keep a long history of the past events in order to be able to predict the future events. Another way that one can think about learning is what is called the online method of learning. In contrast to

batch learning here one only needs to keep the last estimate of the process in order to come up with the next estimate using the last input seen.

This way of thinking about learning although may be mathematically equivalent to batch learning, is more biologically sound and computationally tractable. To derive the online equivalent of batch estimate for $\overline{Z}$ we start from equation 1.2, we have:

$$\overline{Z}_N = \frac{1}{N}\sum_{i=1}^{N} Zd_i = \frac{\sum_{i=1}^{N-1} Zd_i + Zd_N}{N} = \frac{(N-1)}{N} \times \overline{Z}_{N-1} + \frac{Zd_N}{N} = \overline{Z}_{N-1} + \frac{1}{N} \times (Zd_N - \overline{Z}_{N-1})$$

$$\overline{Z}_N = \overline{Z}_{N-1} + \frac{1}{N} \times (Zd_N - \overline{Z}_{N-1}) \qquad \textit{Global mean-track model (online)} \qquad 1.3$$

Therefore one can construct an optimal output $\overline{Z}_N$ on each trial based on the previous estimate $\overline{Z}_{N-1}$ and the difference between the last measurement Zd_N and $\overline{Z}_{N-1}$. Using the decreasing coefficient $\frac{1}{N}$ will guarantee that our estimate at each step using equation 1.3 will be equal to what would have been resulted if the batch equation 1.2 would be used for the same number of inputs. If we substitute $\frac{1}{N}$ with a constant 1-α it means that we weigh recent experiences more as we go through more and more trials. In this case we define our $\overline{Z}_N$ as a weighted average of the past experiences, we have:

$$\bar{Z}_N = \frac{\sum_{i=1}^{N} \alpha^{N-i} \times Zd_i}{\sum_{i=1}^{N} \alpha^{N-i}} \qquad \textit{Local-mean track model (batch)} \qquad 1.4$$

We can derive the online version of the Local mean-track model also we have:

$$\bar{Z}_N = \frac{\sum_{i=1}^{N} \alpha^{N-i} \times Zd_i}{\sum_{i=1}^{N} \alpha^{N-i}} = \frac{\alpha \times \sum_{i=1}^{N-1} \alpha^{N-1-i} \times Zd_i + Zd_N}{\sum_{i=1}^{N} \alpha^{N-i}} = \frac{\alpha \times \sum_{i=1}^{N-1} \alpha^{N-1-i} \times \bar{Z}_{N-1} + Zd_N}{\sum_{i=1}^{N} \alpha^{N-i}}$$

$$= \frac{\alpha \times \sum_{i=1}^{N-1} \alpha^{N-1-i} \times \bar{Z}_{N-1} + \bar{Z}_{N-1} - \bar{Z}_{N-1} + Zd_N}{\sum_{i=1}^{N} \alpha^{N-i}} = \frac{\sum_{i=1}^{N} \alpha^{N-i} \times \bar{Z}_{N-1} + Zd_N - \bar{Z}_{N-1}}{\sum_{i=1}^{N} \alpha^{N-i}}$$

$$\bar{Z}_N = \bar{Z}_{N-1} + \frac{Zd_N - \bar{Z}_{N-1}}{\sum_{i=1}^{N} \alpha^{N-i}}$$

If N>>1 and $\alpha < 1$ we get:

$$\bar{Z}_N = \bar{Z}_{N-1} + (1-\alpha) \times (Zd_N - \bar{Z}_{N-1}) \qquad \textit{Local-mean track model (online)} \qquad 1.5$$

The term $Zd_N - \bar{Z}_N$ which appears both in 1.3 and 1.5 is the difference between our previous estimate of a process and the last evidence of that process and is called the "Prediction error". We will update our estimate by adding a fraction of the prediction error to our current estimate to come up with the next. The learning rate of the model in

equation 1.3 is $\frac{1}{N}$ which keeps decreasing as new trials are observed. This decrease in learning rate means that we assume all the previous data observed regardless of their recency are equally informative about what is going to come next. This is not a good assumption in most cases where recent trials are much better predictors of the future than the older trials.

For example this strategy performs worse than the online learning with a constant learning rate in equation 1.5 in the case of an environment with stepwise changes where recent evidence provide a much better estimate for the optimal prediction for the next trial. As shown below equation 1.5 gives a lower value for the quadratic cost function 1.1. The simulation illustrates the performance of both models in a static versus a dynamic environment with a step change.

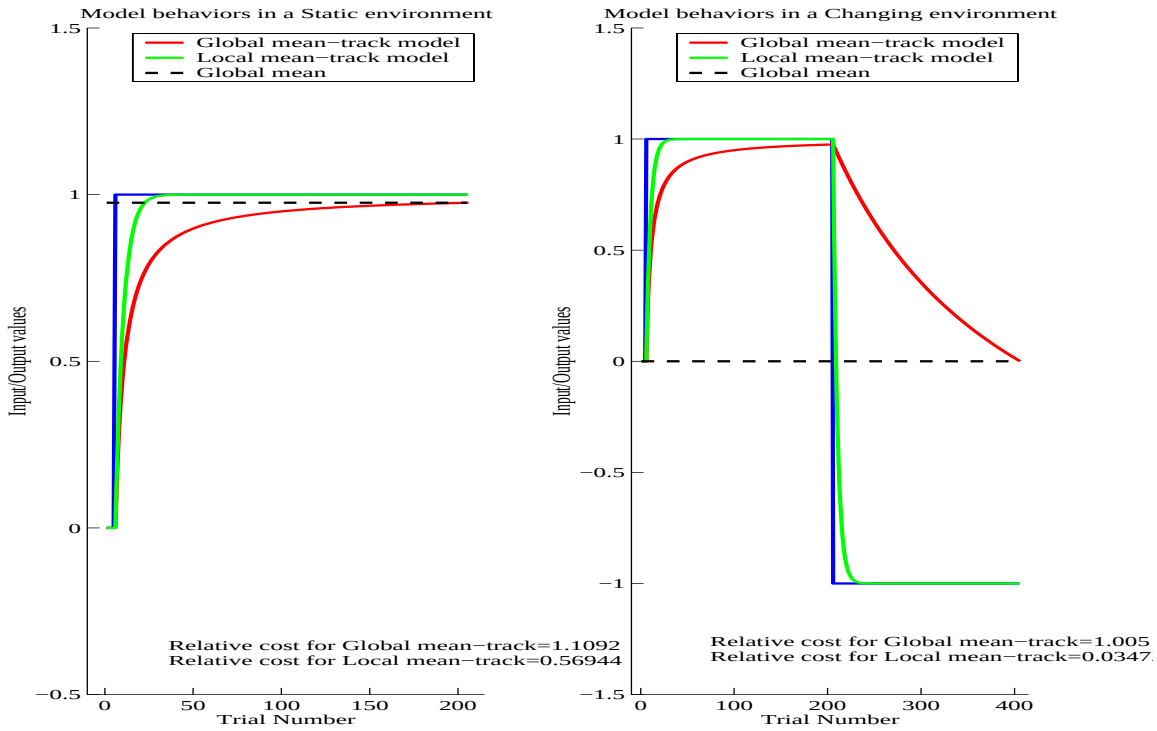

Figure 1.3

As stated before the best constant output to minimize a quadratic cost function is the global mean of the input across all the trials. Figure 1.3 shows the performance of the global mean-track model derived in 1.3. One can see how the value of the Global mean-track system converges to the global mean of the input after all trials are visited. As a measure of relative performance of this model to the performance of the global mean, we found the ratio of the quadratic cost function for model 1.3 to the cost function for the global mean. Notice that the cost function for the global mean is the same as the variance of the input multiplied by the number of trials. Therefore this relative measure is the ratio of the sum of squared errors to the input variance divided by the total number of trials. Surprisingly this number is bigger than one for both static and dynamic input in this case

and even more so for the static input (Figure 1.3 Left). In general the relative cost is a function of the input dynamics as well as the number of trials.

On the other hand the Local mean track system can follow the input much faster and converges to the local mean of the input since it weighs recent evidence more. The relative cost index calculated for this model shows that the model outperforms the constant compensation offered by the global mean of the input. This advantage is even bigger when the model tries to learn to perform in a changing environment as shown in the right part of figure 1.3 due to the increased variance of the input.

Notice that the global mean of the input over certain number of trials is the best "constant" to minimize the quadratic cost function for those trials but this does not mean that we cannot do better than a cost equal to the variance of input times the input length. If we allow for changes to happen in the output we can achieves costs much smaller than what would be resulted with just the mean, and this is what happens in case of model 1.5.

Our Local mean-track model can converge to the current level of the input given enough trials of constant input presented to the system. Yet it is not clear if subjects can reach this level of performance in real world. To upgrade our Local mean-track model into a more general form, one can modify equation 1.5 to have a partial learning of the local mean assuming that one can only learn a fraction β of the weighted mean as follows:

$$\overline{Z}_N = \beta \times \frac{\sum_{i=1}^{N} \alpha^{N-i} \times Zd_i}{\sum_{i=1}^{N} \alpha^{N-i}} = \frac{\alpha \times \beta \times \sum_{i=1}^{N-1} \alpha^{N-1-i} \times Zd_i + \beta \times Zd_N}{\sum_{i=1}^{N} \alpha^{N-i}} = \frac{\alpha \times \sum_{i=1}^{N-1} \alpha^{N-1-i} \times \overline{Z}_{N-1} + \beta \times Zd_N}{\sum_{i=1}^{N} \alpha^{N-i}}$$

$$= \frac{\alpha \times \sum_{i=1}^{N-1} \alpha^{N-1-i} \times \overline{Z}_{N-1} + \overline{Z}_{N-1} - \overline{Z}_{N-1} + \beta \times Zd_N}{\sum_{i=1}^{N} \alpha^{N-i}} = \frac{\sum_{i=1}^{N} \alpha^{N-i} \times \overline{Z}_{N-1} + \beta \times Zd_N - \overline{Z}_{N-1}}{\sum_{i=1}^{N} \alpha^{N-i}}$$

$$\overline{Z}_N = \overline{Z}_{N-1} + \frac{\beta \times Zd_N - \overline{Z}_{N-1}}{\sum_{i=1}^{N} \alpha^{N-i}} = \overline{Z}_{N-1} + \frac{\beta \times Zd_N - \beta \times \overline{Z}_{N-1} + \beta \times \overline{Z}_{N-1} - \overline{Z}_{N-1}}{\sum_{i=1}^{N} \alpha^{N-i}}$$

$$= (1 - \frac{1-\beta}{\sum_{i=1}^{N} \alpha^{N-i}}) \times \overline{Z}_{N-1} + \frac{\beta}{\sum_{i=1}^{N} \alpha^{N-i}} \times (Zd_N - \overline{Z}_{N-1})$$

$$\alpha, \beta < 1 \qquad 1.6$$

Again assuming that $N >> 1$ we can rewrite 1.6 as follows:

$$\overline{Z}_N = (1 - (1-\beta) \times (1-\alpha)) \times \overline{Z}_{N-1} + \beta \times (1-\alpha) \times (Zd_N - \overline{Z}_{N-1}) \qquad 1.7$$

We can define a learning rate and a retention capacity for our online model as follows:

$$B = \beta \times (1-\alpha)$$
$$A = 1 - (1-\beta) \times (1-\alpha)$$

B: Learning rate A: Retention capacity 1-A: Forgetting rate

Put together the prediction error and the online update gives us the single state-space equation below:

$$E_N = Zd_N - \overline{Z}_N$$

$$\overline{Z}_N = A \times \overline{Z}_{N-1} + B \times (Zd_N - \overline{Z}_{N-1}) \qquad \textit{Local mean-track with forgetting} \qquad 1.8$$

The two equations in 1.8 are the basis of the state-space model used in this thesis to model the human motor control learning. Figure 1.4 shows the behavior of this model in comparison with models derived in 1.3 and 1.5

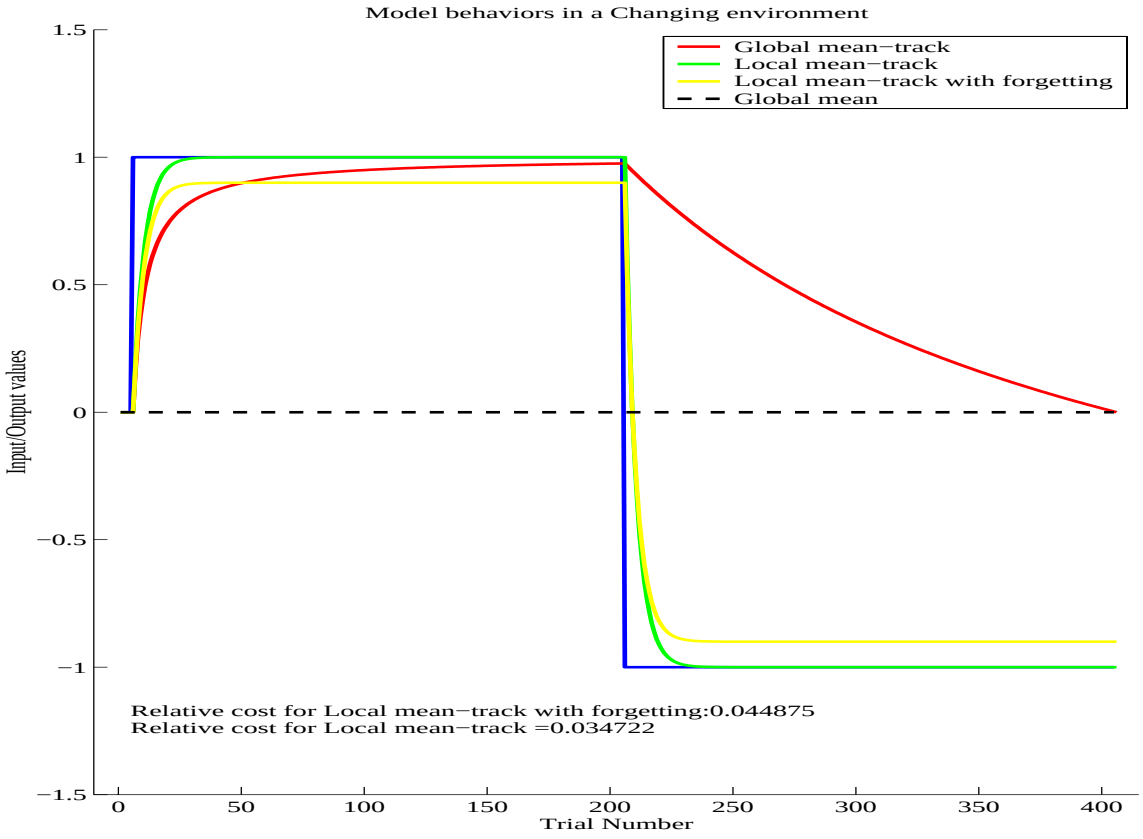

Figure 1.4

In Figure 1.4 we have used $\beta = 0.9$ and $\alpha = 0.8$ for the local mean-track with forgetting. As expected we see that this model eventually compensates for 90% of the constant level

of input. This will of course increase the cost relative to the local mean-track with no forgetting which has the same value for α while its β is equal to "one" as it is shown in figure 1.4.

1.4 Steady state error and learning timecourse of the system

Looking at figure 1.4 one can see that the system simulated using equation 1.8 which has a forgetting factor different from "one" never reaches the same level of performance as the system of equation 1.5. This error that is sustained in response to a constant input after theoretically infinite time is called "Steady state error" and is defined as:

$$\underset{N \to \infty}{E_N = Zd_N - \overline{Z}_N} \quad \text{Steady state error} \qquad 1.9$$

If $A=1$ in the state-space equation as in equation 1.5, to calculate the steady state error we have:

$$\overline{Z}_\infty = \overline{Z}_\infty + B \times (Zd_\infty - \overline{Z}_\infty) \Rightarrow \overline{Z}_\infty = Zd_\infty \Rightarrow E_\infty = 0 \qquad 1.10$$

As expected the steady state error for a system that does not forget and have $A=1$ is zero independent of the learning rate B. On the other hand if we let $A \neq 1$ as in equation 1.8 we have:

$$\bar{Z}_\infty = A \times \bar{Z}_\infty + B \times (Zd_\infty - \bar{Z}_\infty) \Rightarrow \bar{Z}_\infty = \frac{B}{1 - A + B} Zd_\infty \Rightarrow E_\infty = \frac{1 - A}{1 - A + B} Zd_\infty \qquad 1.11$$

It can be seen that when we let the forgetting to take place in the learning system, we are just going to learn to compensate for a fraction of the input which is equal to $\frac{B}{1 - A + B}$. Substituting for α and β we have:

$$\bar{Z}_\infty = \frac{\beta(1-\alpha)}{1-(1-(1-\beta)(1-\alpha))+\beta(1-\alpha)} Zd_\infty \Rightarrow \bar{Z}_\infty = \beta Zd_\infty \quad \textit{Steady state} \qquad 1.12$$

As expected the final value of $\bar{Z}_\infty$ does not depend on the value for α. It will be a fraction of the constant input dictated by β only.

The magnitude of the steady state error will depend on values of *B, A* and Zd_∞. For example for learning rate *B* equal to *0.2* and the forgetting factor *A* equal to *0.9* we get:

$$\bar{Z}_\infty = \frac{0.2}{1 - 0.8 + 0.2} Zd_\infty = 0.5 \times Zd_\infty \Rightarrow E_\infty = \frac{1 - 0.8}{1 - 0.8 + 0.2} Zd_\infty = 0.5 \times Zd_\infty$$

This system only compensates for half of the input. Increasing the learning rate as well as the retention capacity will increase the asymptotic performance of the system.

But how the values of *A* and *B* define the timecourse of learning before reaching this asymptote?

To derive the timecourse of learning we have to rearrange our single state-space model in 1.8 into a standard discrete differential equation, we have:

$$\bar{Z}_N = A \times \bar{Z}_{N-1} + B \times (Zd - \bar{Z}_{N-1}) \Rightarrow \bar{Z}_N - \bar{Z}_{N-1} = (A - 1 - B) \times \bar{Z}_{N-1} + B \times Zd \qquad 1.13$$

This is a first order discrete differential equation. Here we assume that the input to the system is constant that is: $Zd_N = Zd = const.$ The solution to the standard first order differential equation is as follows:

$$\bar{Z}_N - \bar{Z}_{N-1} = \gamma \times \bar{Z}_{N-1} + \varphi \times Zd \qquad \textit{where} \quad \begin{matrix} \gamma = A - 1 - B \\ \varphi = B \end{matrix} \qquad 1.14$$

$$\bar{Z}_N = C \times (1-\gamma)^N - \frac{\varphi}{\gamma} \times Zd$$

The solution consists of a homogeneous part which is $C \times (1-\gamma)^n$ where *C* is a constant defined by the initial condition and the non homogeneous part which is $-\frac{\varphi}{\gamma} \times Zd$.

If $\bar{Z}_0 = 0$ then $C = \frac{\varphi}{\gamma} \times Zd$, substituting for values of C, γ and φ we get:

$$\bar{Z}_N = \frac{B}{1 - A + B} \times Zd \times (1 - (A - B)^N) \qquad 1.15$$

Equation 1.15 is an important equation since it shows the temporal (trial to trial) solution for the single state space model developed so far. We can substitute for α and β we get:

$$\bar{Z}_N = \beta \times Zd \times (1 - \alpha^N) \qquad 1.16$$

Since $0 < \alpha < 1$ as $N \to \infty$ we get $\bar{Z}_\infty = \beta \times Zd$ which is the same steady state value derived previously. If we represent the trial constant[1] of our discrete system by η, we have:

$$1 - \alpha^\eta = 1 - e^{-1} \Rightarrow \eta = \frac{-1}{\ln(\alpha)} = \frac{-1}{\ln(A - B)} \qquad \textit{Trial constant} \qquad 1.17$$
$$0 < A - B = \alpha < 1$$

Therefore the time course of adaptation is controlled by the difference between the retention capacity *A* and learning rate *B*.

Notice that our initial assumption about α being between zero and one, naturally leads to state-space equations with $A > B$, that is learning rate will always be smaller than the retention capacity. What will happen if we let the learning rate become equal or larger than the retention capacity?

[1] "Trial constant" is the counterpart of the "Time constant" for discrete systems and it is equal to the trial number in which the system reaches $1\text{-}e^{-1}$ ~ %63 of the final value.

Remember that as long as $|\alpha| < 1$ the weighted of sum of the past experiences in our local mean-track model for a bounded input remains bounded and thus our single state space model is stable for $-1 < \alpha < 1$. Figure 1.5 shows the performance of the system with three different values of α {0.8; 0; -0.8}. The value for β is set to 0.5 to yield the same asymptotic performance for all three models.

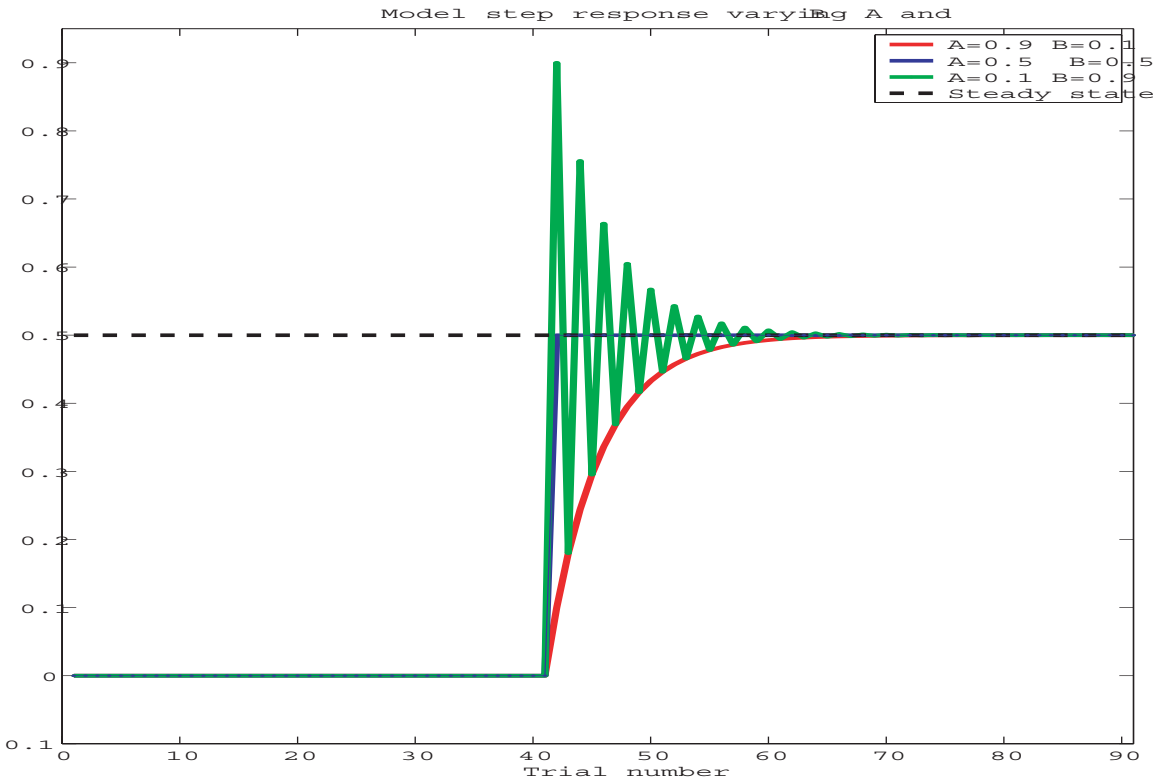

Figure1.5

These values for α and β correspond to the following values for *A* and *B* and η:

Model I: A=0.9 B=0.1 $\eta = \dfrac{-1}{\ln(0.8)} = 4.48$ *trls*

Model II: A=0.5 B=0.5 $\eta = \frac{-1}{\ln(0.0)} = 0$ *trls*

Notice that although the learning rate is different from one ($B = 0.5$) , yet the model learns to reach its asymptote on the very first trial (see figure 1.5). This is because the true time course of learning is a function of *A*-*B*.

Model III: A=0.4 B=0.6

As expected or negative values of α the model shows oscillatory behavior. Yet the response envelope of model III seems to be the same as the response of model I. A simple modification to 1.17 can be used to predict this phenomenon if we replace α with its magnitude $|\alpha|$ we have:

$$\eta = \frac{-1}{\ln(|\alpha|)} = \frac{-1}{\ln(|A-B|)} \qquad \textit{Generalized trial constant} \qquad 1.18$$

If $|\alpha| \to 1$ equation 1.18 can be approximated as:

$$\eta = \frac{-1}{\ln(|\alpha|)} = \frac{-1}{\ln(1-(1-|\alpha|))} \Rightarrow \eta \to \frac{1}{1-|\alpha|} = \frac{1}{1-|A-B|} \qquad 1.19$$
$$|\alpha| \to 1$$

It is worth mentioning that although the steady state value of the system and its time course seem to be governed by an interaction of *A* and *B*, from the point of view of α

and β they are separately specified. As shown previously the steady state of the system depends only on β while timecourse of the system is defined independently by α. Smaller values of α causes the system to be faster while values of α close to "1" make the system arbitrarily slow. Remember that smaller values of α means that the model weighs recent evidence more than the older evidence and this will make the system to have a faster learning rate.

We can use the example of an electrical circuit to illustrate our derivation of the steady state and the time course of learning for our single state-space model. The circuit shown has its input-output relation governed by the same state space equation as in 1.8.

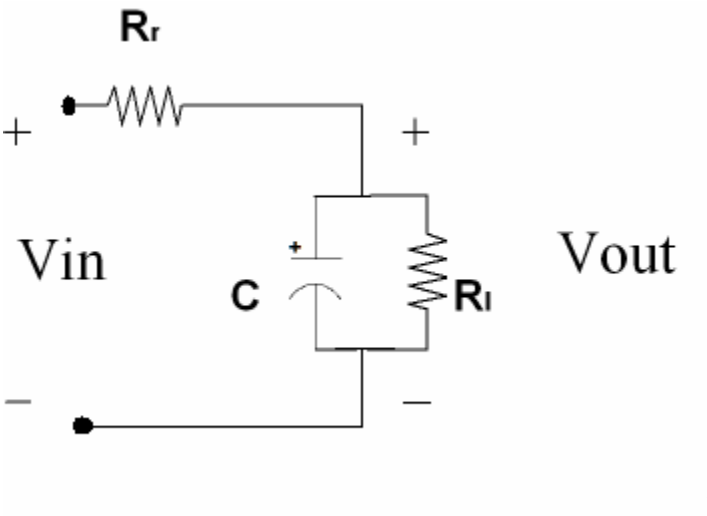

Figure 1.6

To derive the input output relation we can use Kirchhoff's voltage and current equations as follows:

$$\frac{dV_c}{dt} = \frac{I_c}{C} = \frac{1}{C}\left(\frac{V_{in} - V_c(t)}{R_r} - \frac{V_c(t)}{R_l}\right) \Rightarrow V_c(t+\Delta t) = V_c(t) + \frac{1}{C}\left(\frac{V_{in} - V_c(t)}{R_r} - \frac{V_c(t)}{R_l}\right)\Delta t$$

Rearranging the above equation results in:

$$V_c(t+\Delta t) = \left(1 - \frac{\Delta t}{CR_l}\right) \times V_c(t) + \frac{\Delta t}{CR_r}(V_{in} - V_c(t)) \qquad 1.20$$
$$\Delta t \to 0$$

Therefore the output voltage across the capacitor V_c at time $t + \Delta t$ depends on V_c at time t plus the difference between the input and V_c at time t. This is similar to equation 1.8 where the output $\overline{Z}_{N+1}$ depends on $\overline{Z}_N$ and the prediction error which is $Zd_N - \overline{Z}_N$.

A comparison between 1.8 and 1.20 gives:

$$A = 1 - \frac{\Delta t}{CR_l}$$
$$B = \frac{\Delta t}{CR_r} \qquad 1.21$$

The steady state value of this system is equal to $\frac{R_l}{R_r + R_l}$. Using equation 1.21 we have:

$$V_c(\infty) = \frac{R_l}{R_r + R_l} = \frac{B}{1 - A + B}$$

The time constant of this circuit is known to be equal to $C \times \frac{R_r R_l}{R_r + R_l}$. Using equation 1.21 we have:

$$\tau = C \times \frac{R_r R_l}{R_r + R_l} = \frac{\Delta t}{1 - A + B} \xrightarrow{\Delta t = 1} \tau = \frac{1}{1 - A + B} = \frac{1}{1 - \alpha} \qquad 1.22$$

The time constant derived here is accurate when $\Delta t \to 0$ (see 1.20) which makes $\alpha \to 1$. Notice that this is the same condition we used to derive the time constant in 1.19 as an especial case.

1.5 Experimental results

To test the prediction of our state space model 1.8, and to see how the subjects learn to respond to the changes in the environment we designed a point-to-point reaching experiment using a robotic manipulandum shown below.

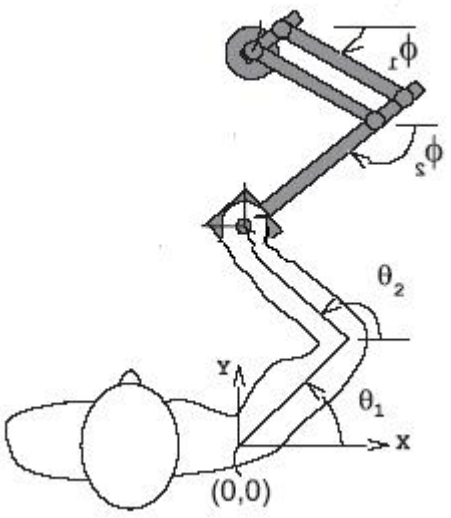

Figure 1.7

Methods:

We asked 14 healthy right handed the subjects to sit in front of a monitor screen while holding a robotic arm. Subjects were supposed to move from an origin to a target 10cm away (Outward movement) and after finishing the movement come back to the origin (Back movement). Both origin and target were represented by circles 1cm in diameter.

Subjects were required to finish each out or back movement in *500±50 ms*. Subjects received feedback based on their total movement time at the end of each movement by target color change. A blue color at the end of the movement meant a slow movement while a red color meant fast movement. Subjects who managed to make a *500±50 ms* movement would receive a target "explosion" as their reward

The experimental sequence is shown below:

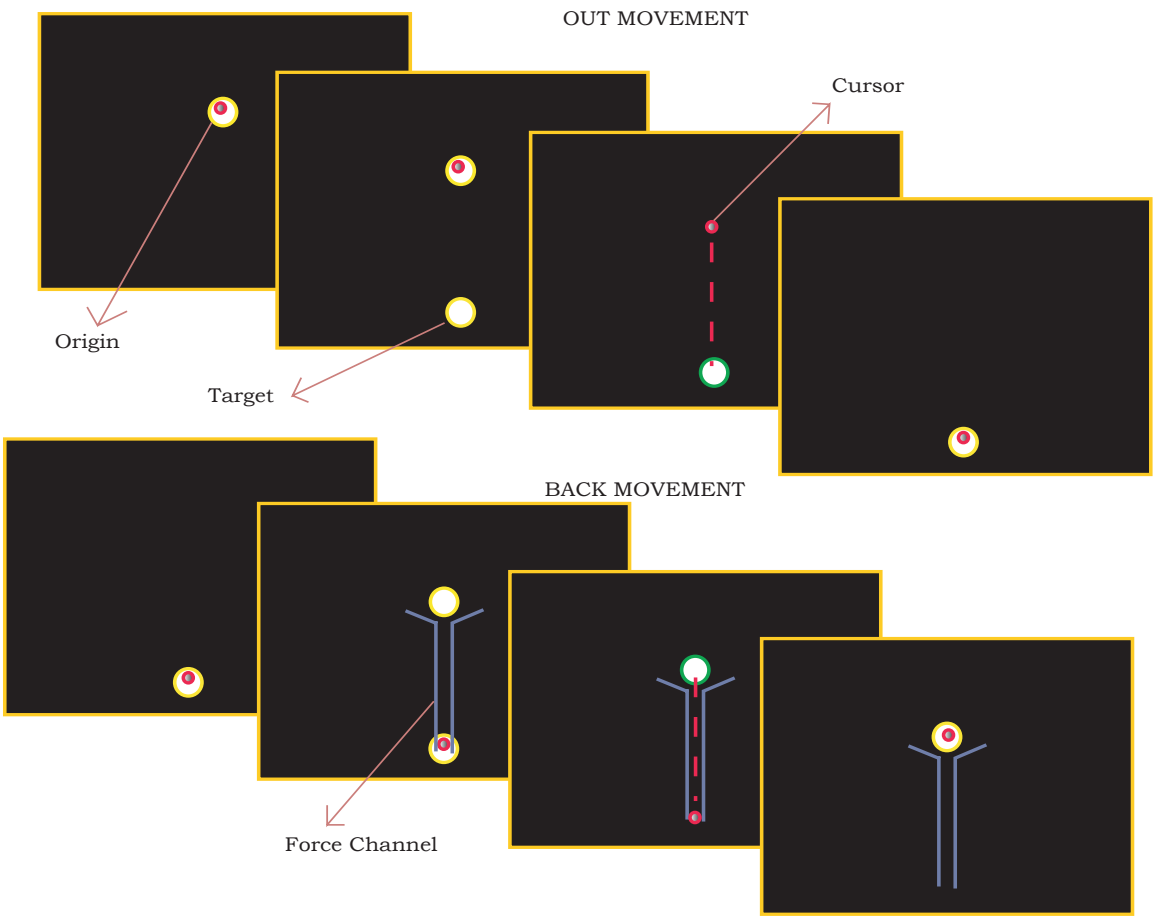

Figure 1.8

Back movements were all carried out in the “Force channel” environment. During force channel trials robot restricted the movement of the arm to the straight line connecting the origin to the target. Since the deviation from straight line is considered to be used by the subjects as an error signal to modify their subsequent motor output, using force channel on back trials ensures that only out movement would be used to change motor behavior.

Subjects were first trained to reach with the robotic arm for two blocks of trials (Baseline training). Each block of consists of 60 out and 60 back movements. Then the robot

started to exert a velocity dependent curl force on subjects' arms during the out movements for two blocks of trials. This type of force field is a well studied environment with which people have studied motor skill learning. The structure of the field in its relation to hand velocity is shown below:

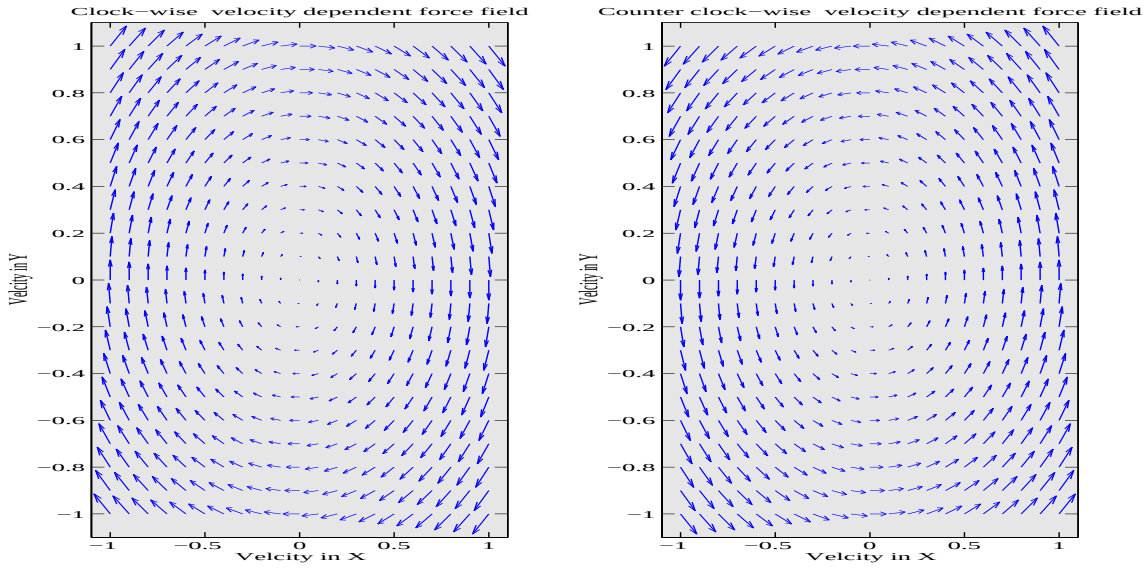

Figure 1.9

This force always acts perpendicular to the direction of motion and as such makes the hand trajectory deviate from the straight line. As subjects move in this field they learn to compensate for the forces and gradually restore the original smooth straight paths to the target. The perpendicular displacement (PD) from the straight line at peak velocity has been used as a measurement of learning.

After the baseline training we divided our subjects into two groups. One group experienced a clockwise curl field while the other one learned a counter clockwise field (see Figure 1.9). As shown below both groups of subjects continuously reduced their PD to eventually have PDs that were not significantly different from zero. The results of a *t-test* on the initial 10 trials in force field compared to the last 10 trials performed in null for both groups shows a significant difference (*Clockwise tstat=5.2606, P<0.0001 ;Counter clockwise tstat=-2.8606, P=0.01<0.05 df=17*) while the last 10 movements in the field were not different from the baseline (*Clockwise tstat=0.48979, P=0.6>0.05 ;Counter clockwise tstat= 0.043821, P=0.9>0.05 df=17*).

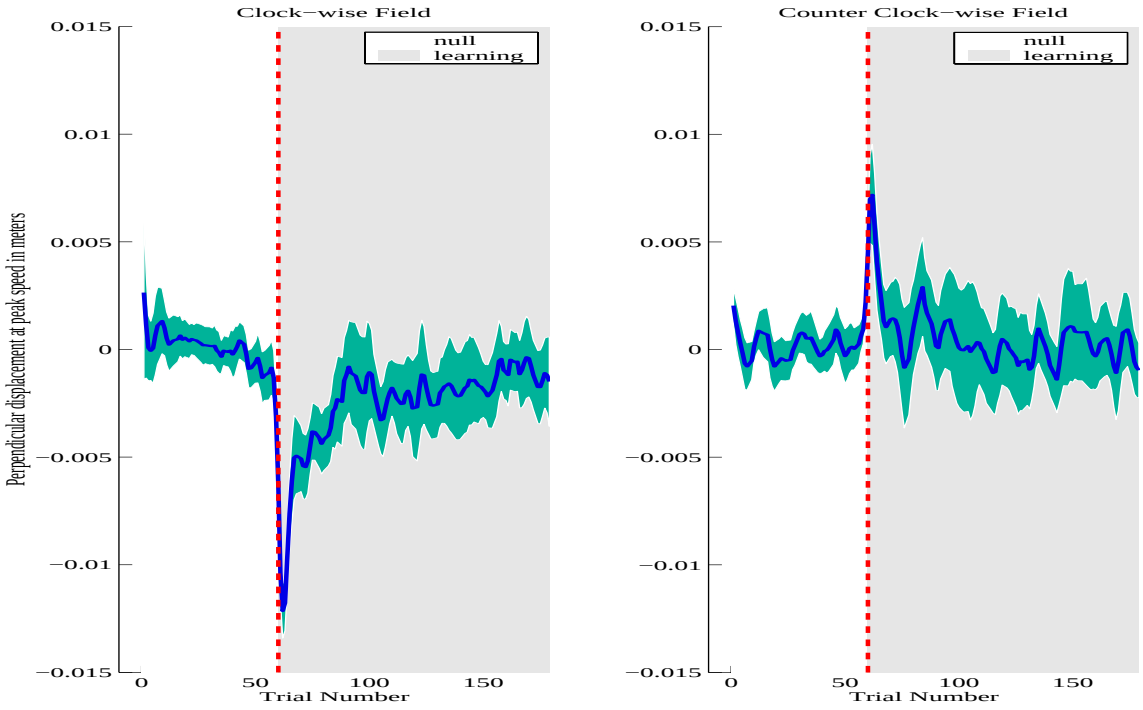

Figure 1.10

To directly monitor the force output change as subjects learned to compensate for the field, we included occasional force channel trials in outward movements as well. These trials were considered to have negligible effect on learning since perpendicular errors are clamped at zero during these trials and any change in the motor output is caused by a hypothetically small forgetting rate and thus may be ignored. The greatest advantage of looking at the force readings in the force channel trials is that it can be considered to be mainly affected by the feed forward output known as the internal model since the feedback module gets inactivated when the errors are clamped at zero.

The total motor output is effected both by the formation of the internal model and the modification of feedback responses. Here our main interests lie with studying how subjects learn the model of their environment through feed forward responses and as such force channel trials provides us with an ideal situation to study the changes of the internal model.

We can fit our state space model to the data by using perpendicular displacements or by directly measuring the lateral forces subjects produced in force channel trials recorded by the transducer. It is important to note that the learning processes inferred from PDs contain the learning effects from both feedback and feed forward block while the fits to the force channel trials should only include the feed forward contribution to learning. Figure 1.11 shows both leaning indexes used and their corresponding model fits. The data shown is the combined data for both groups of subjects.

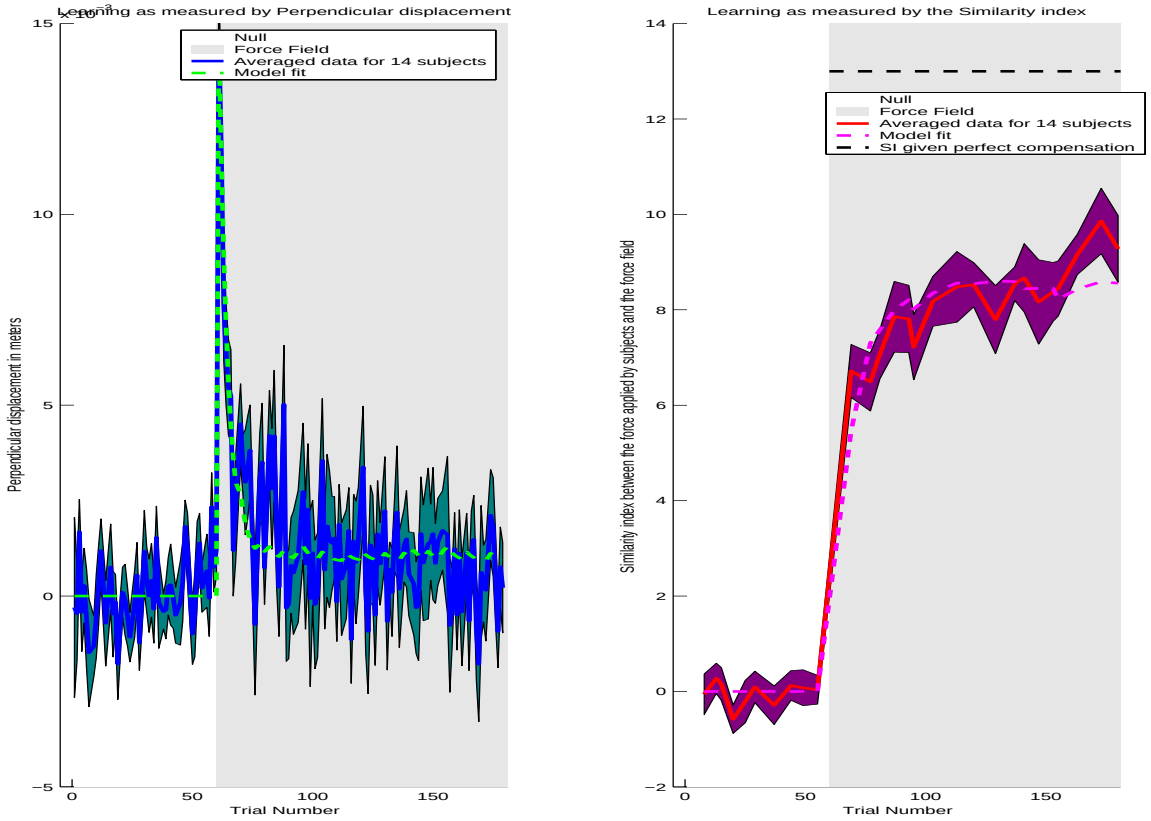

Figure 1.11

To quantify learning using force transducer readings in force channel trials we looked at the Similarity index (*SI*) between the force timecourse during the movement to the ideal velocity dependent force given full compensation using the following formulation:

$$SI = \frac{\vec{F}_T . \vec{F}_I}{\left|\vec{F}_I\right|^2}$$

$\vec{F}_T$: *Lateral force during a trial as measured by force transducer*

$\vec{F}_I$: *Ideal lateral force needed for complete force compensation*

1.23

Similarity index has been calculated using the dot product of these two force profiles in time divided by the dot product of the ideal force profile by itself. *SI* values that are close to one indicate close to complete learning of the force pattern while smaller values mean under-compensation of the force field.

The fitted parameters for the single state-space model using PDs are:

$A = 0.985, B = 0.222$ with $R^2 = 0.71$ therefore $\overline{Z}_\infty = 0.94$ and $\tau = 4.2\, trials$

The fitted parameters for the single state-space model using SIs are:

$A = 0.963, B = 0.078$ with $R^2 = 0.94$ therefore $\overline{Z}_\infty = 0.67$ and $\tau = 8.7\, trials$

Both measures provide reasonable fits for the single state-space model. Yet the fact that the fit to the SIs has a higher *R-squared* may imply that our single state-space model is more suited to describe the formation of the internal model compared to the total motor output.

Notice that while the data for PD predicts a close to perfect compensation for $\overline{Z}_\infty$, SI data finds the final learning to be less than *70%* of the total force needed for complete compensation. Also since the learning rate calculated from PDs can roughly be considered as a summation between the learning rate of the feed-forward and feed-back

modules, the learning rate derived using PDs overestimates the true learning rate of the internal model as expected.

1.6 Extending the State-space learning models to account for generalization

So far we have only discussed how we can learn a particular motor skill for a specific movement direction which in our case involved only outward movements. One of the most important issues in motor learning is the question of generalization and that is how learning a motor behavior in a specific set of states or situations may manifest itself in other similar or non similar sets of states or situations.

To model the generalization of a learned motor output to all sets of possible motor states, a weighted combination of local basis functions of input space should be used, we have:

$$Z(x) = \sum_{j=1}^{m} w_j g_j(x) \qquad \textit{x: Input space } \textit{w}_j\textit{: Combination weights} \qquad 1.24$$

$$\textit{g}_j\textit{(x): Gaussian basis functions}$$

$$\textit{Z(x): Motor output as a function of the state}$$

Notice that here we have defined *Z(x)* as a learned motor output for each state *x*. This is a more general notation than what was presented as $\overline{Z}$ in 1.8 which was formulated to represent a single output index for one movement direction only. Position, velocity and acceleration are some of the important states to be considered in the case of our point to

point reaching tasks as represented by *x*. During the course of a single movement subjects visit various states of different positions, velocities and accelerations and he is supposed to learn the force associated with each state visited to be able to perform ideally in the task. The learning indexes like PD and SI used to fit the state-space model would only give us an overall measure of learning for the whole movement and in that sense equation 1.16 is a more exact way of formulating the detailed association that takes place between motor output and each input state.

But how does our new model update its output at each trial?

It can be proved easily that to minimize the same quadratic cost function in 1.1 representing the error between the current and ideal motor output, one should update the combination weights in 1.24 using a gradient descent rule. We have:

$$E(x_N) = Zd(x_N) - Z_N(x_N)$$
$$w_j^{(N+1)} = w_j^{(N)} + \eta g_j(x_N) E_N(x_N) \qquad \eta : \textit{Learning rate for gradient descent} \qquad 1.25$$
$$x_N : \textit{State visited on trial “ N”}$$

Depending on how broad the basis functions $g_j(x)$ are the error in one state can affect the performance in the neighboring states as well as its own. If we rewrite 1.25 in a state-space format we get:

$$Z_{N+1}(x) = \sum_{j=1}^{m} w_j^{(N+1)} g_j(x) = \sum_{j=1}^{m} (w_j^{(N)} + \eta g_j(x_N) E_N(x_N)) g_j(x) = Z_N(x) + \eta \sum_{i=1}^{m} g_j(x) g_j(x_N) E(x_N)$$

$$Z_{N+1}(x) = Z_N(x) + \eta G^T(x) G(x_N) E(x_N) = Z_N(x) + B(x, x_N) E(x_N)$$

Where: $$G(x) = \begin{bmatrix} g_1(x) \\ g_2(x) \\ . \\ . \\ g_m(x) \end{bmatrix} \quad \text{and} \quad B(x, x_N) = \eta G^T(x) G(x_N) \qquad 1.26$$

Here *B* is a learning rate which broadcasts the error observed in the last visited state x_N onto all the other states *x*. The broadcast of the error is maximum for x_N itself since $\eta G^T(x) G(x_N)$ has its peak at $x = x_N$ and this value will decline depending on the width of $g_i(x)$ for all other *x*. To make 1.26 a full generalization of 1.8 we can rewrite 1.25 by adding a state dependent retention capacity to the gradient descent update rule. It follows:

$$E(x_N) = Zd(x_N) - Z_N(x_N)$$
$$w_j^{(N+1)} = A(x, x_N) w_j^{(N)} + \eta g_j(x_N) E_N(x_N) \qquad 0 < A(x, x_N) < 1$$

$$\therefore Z_{N+1}(x) = A(x, x_N) Z_N(x) + B(x, x_N) E(x_N) \qquad 1.27$$

The retention capacity *A* governs the forgetting rate of the motor output for each state *x* as a function of the last state visited x_N. It is conceivable that the amount of forgetting in the motor output for a particular state to be dependent on the last state in which motor output has been executed and error has been observed.

This shows one of the complications of our experimental design for fitting the single state-space model 1.8 to the out movements. That is our estimates of learning rate and forgetting factor for the out direction can be affected by the generalized forgetting factor 1- $A(x, x_N)$ caused by the back movement ($180°$ away in movement direction) in the force channel. Since our current experimental design leaves us with no intelligent way to extract the forgetting effects caused by moving back in the channel, we need to just assume such effects are negligible to be able to keep our current interpretations of the fitted values for *A* and *B* intact. Later in the chapter we will see that such an assumption about minimal effects of a movement $180°$ away might indeed not be so far from the reality.

Equation 1.27 is the general form of the single-timecourse state-space equation 1.8 which formulates the evolution of motor output as a function of each state *x*. Recent studies have tried to explore the properties of $B(x, x_N)$ and $A(x, x_N)$ using force field paradigms in motor learning by applying a velocity dependent curl field (Donchin et al 2003).

Figure 1.12 shows the fitted values for the generalization properties of *B* as a function of movement direction (Thoroughman and Shadmehr 2000). It can be seen that *B* is a

narrow function of movement direction. Its value falls to 50% of its peak value at $B(x_N, x_N)$ for $45°$ degrees away and it almost goes to zero $90°$ away.

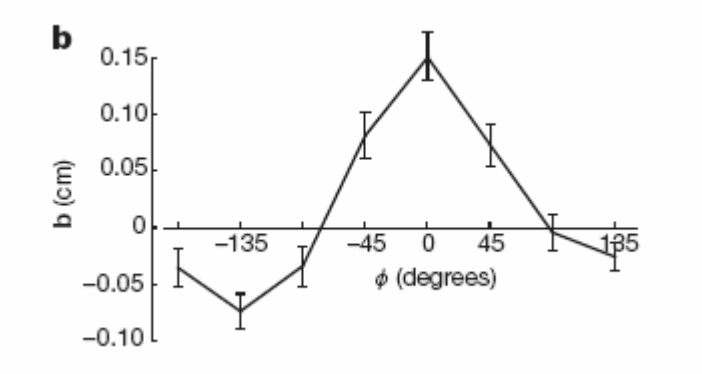

Figure 1.12

An interesting phenomenon here is that the fitted values for *B* using PD suggests a small negative generalization to $180°$ away for the curl force field. We wondered if our experimental results can be used to explore this idea.

1.7 Bimodality in generalization: Reality or a Fitting artifact?

In our point to point reaching experiment, subjects always moved back to the origin in a force channel. This would give us the opportunity to measure the changes in the behavior of the internal model for back movements as learning takes place in outward direction. It

is worth mentioning that the use of force channel would clamp the error in back movements providing us with two advantages:

1. Since subjects observed little deviation from straight line in back movements *<1mm*, learning would be limited to out movements and as such all motor output changes in back movements can be assumed to arise form generalization to 180° away.

2. Clamping errors on back movements minimizes the effects of feedback output contaminating the output of the internal model. Therefore we can be confident of the force readings to reflect mostly the changes in internal model.

Figure 1.13 shows trial by trial evolution of the force output perpendicular to movement direction in back movements as learning takes place in out movements. The data shown is the combined data from our both groups of subjects.

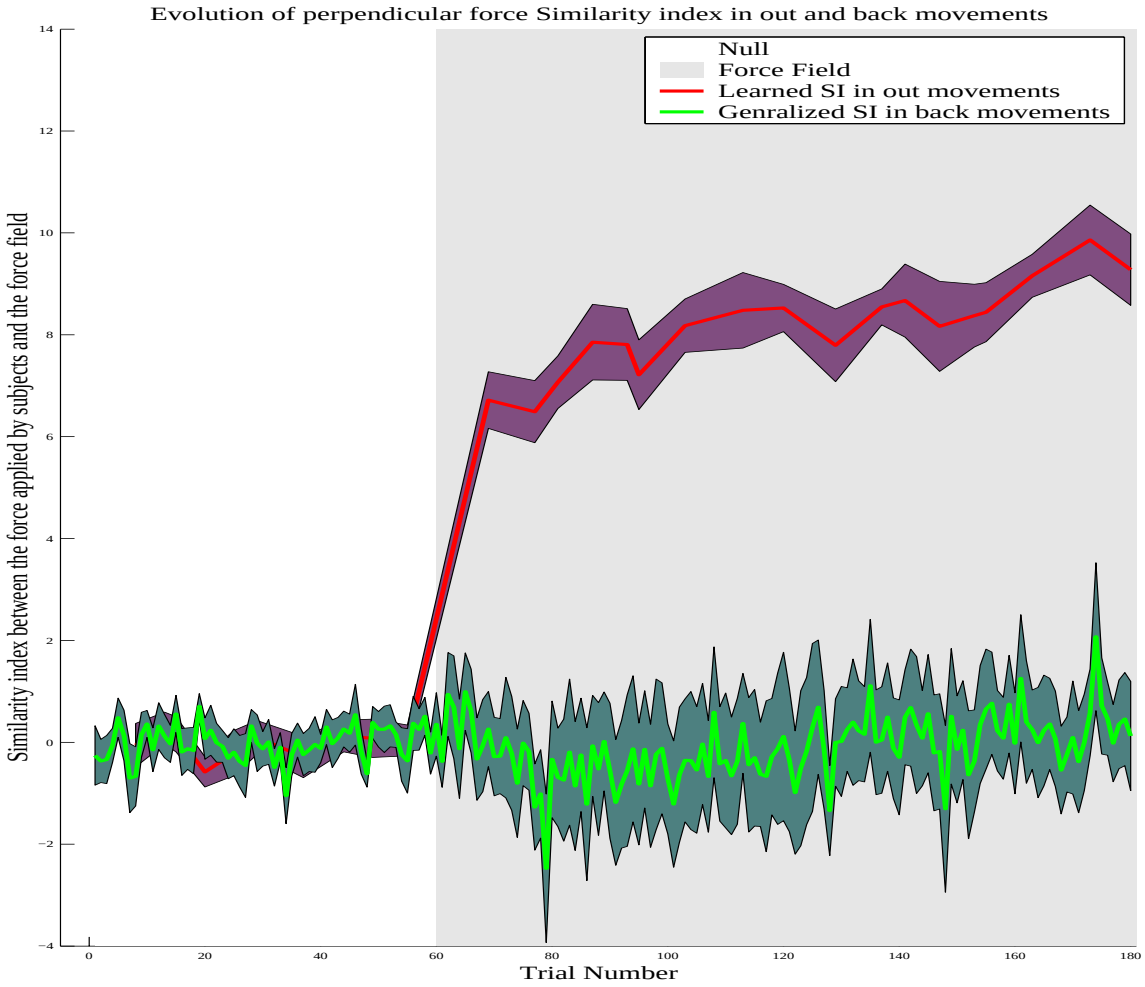

Figure 1.13

As learning progresses in out direction, there is hardly any evidence for generalization of learning to 180° away as shown in figure 1.13. Nevertheless one should remember that the result in figure 1.12 is calculated from measures of PD and we already have shown that this measure can also reflect the learning processes in feed-back module although the

fact that the generalization predicted for 180° away is negative make this explanation to be an unlikely candidate.

A statistical analysis on the mean SI level before and after introduction of the field for the back movements as subjects learned to compensate for the force fielding on the out movements does not show a meaningful difference (*p=0.28,Two tailed t-test, df=178*). The same t-test this time on the out movement where learning is taking place, shows a significant difference for SI levels in force field vs. null (*p<0.05,CI=[7.5921 8.7732], df=26*). Yet it seems that the variability of the subjects' performance as measured by SI increases significantly for force field vs. null for both out movements as well as back movements (*p<0.05 df=26, for out movements; p<0.5 df=26, for back movements*). This means that learning the force field in one direction can increase the variability of motor control in the same direction as well as in the180° away but it has almost no significant effect on the mean motor output for the 180° away.

1.8 Summary

In this chapter we introduced the idea of online trial to trial motor learning using a single state-space model. We showed that this model can explain a great portion of the general learning trend in the data (high *R-squared*) . In the next chapter we are going to discuss some of the shortcomings of the current state-space model in explaining some of the other well known phenomena in motor learning. We then will introduce an extension to the current model to overcome these difficulties. We also report for the first time a

transient recovery of the older memories in motor adaptation and show how this new peculiar phenomenon can be explained in the light of the new extension to our current single state-space model.

Chapter II

Multi-State Models of Learning

In the previous chapter we introduced an online learning mechanism which was able to explain the main learning trends in motor skill learning as studied by using a velocity dependent curl force field. Here we are going to present some of the motor skill learning behaviors that are not easily explained by our current version of the state-space model.

2.1 Savings

"Savings" is a fundamental behavioral phenomenon where prior adaptation to a novel environment increases the speed of subsequent adaptation to the same environment. In some cases, savings is observed even after performance has been brought back to baseline during a washout period. This means that one can still observe a faster relearning of a certain motor skill after all the measurable performance indexes has been brought back to base line. This phenomenon was observed initially in classical paradigms for eye blink conditioning where re-exposure to paired stimuli after extinction induces reacquisition that is much faster than the original acquisition (Frey and Ross, 1968; Napier et al., 1992).

A recent study by Kojima et al 2004 reported the conditions under which savings may or may not occur in saccade adaptation. It is known that humans as well as the animals can

change the gain of their saccades to ensure accuracy of saccades throughout life despite growth, aging, and some pathologies of the oculomotor plant or nervous system. This means that if saccades consistently overshoot the target, their amplitude gradually decreases, and if they consistently fall short, their size gradually increases. There is good evidence that saccadic adaptation is a form of motor learning driven by visual error immediately after the end of saccades (Wallman and Fuchs, 1998; Shafer et al., 2000; Noto and Robinson, 2001). This is consistent with our online trial-to-trial motor learning discussed in previous chapter. Yet our current model seems to be unable to provide a satisfactory explanation for the effects of savings seen in saccade adaptation. Below we are going to discuss the experimental results by Kojima et al and show single state-space model predictions for each of their experimental paradigms.

2.2 Experimental results (Kojima et al 2004)

Two rhesus monkeys were prepared for eye movement recording by the magnetic search coil method with the use of surgery. During recording sessions, the monkey sat in a primate chair in a darkened booth with its head restrained. The animal was required to make saccades toward a target spot presented on a monitor screen

Double reversal paradigm

In their experiment Kojima et al. induced a series of three alternating gain changes by reversing the polarity of visual error twice during ongoing adaptation. After collecting 100-400 pre-adaptation saccades to horizontal steps of the target, they started the first block of adaptation session by subjecting the animal to 35% forward or backward target

jump during the saccade (learn block). When the gain was altered by ~0.1-0.2 after 400-800 saccades, the direction of target jump was reversed to bring the gain back to the pre-adaptation value, i.e., ~1.0 (unlearn block). They then reversed the target jump direction again and induced a gain change using a jump of the same size as in the learn block.

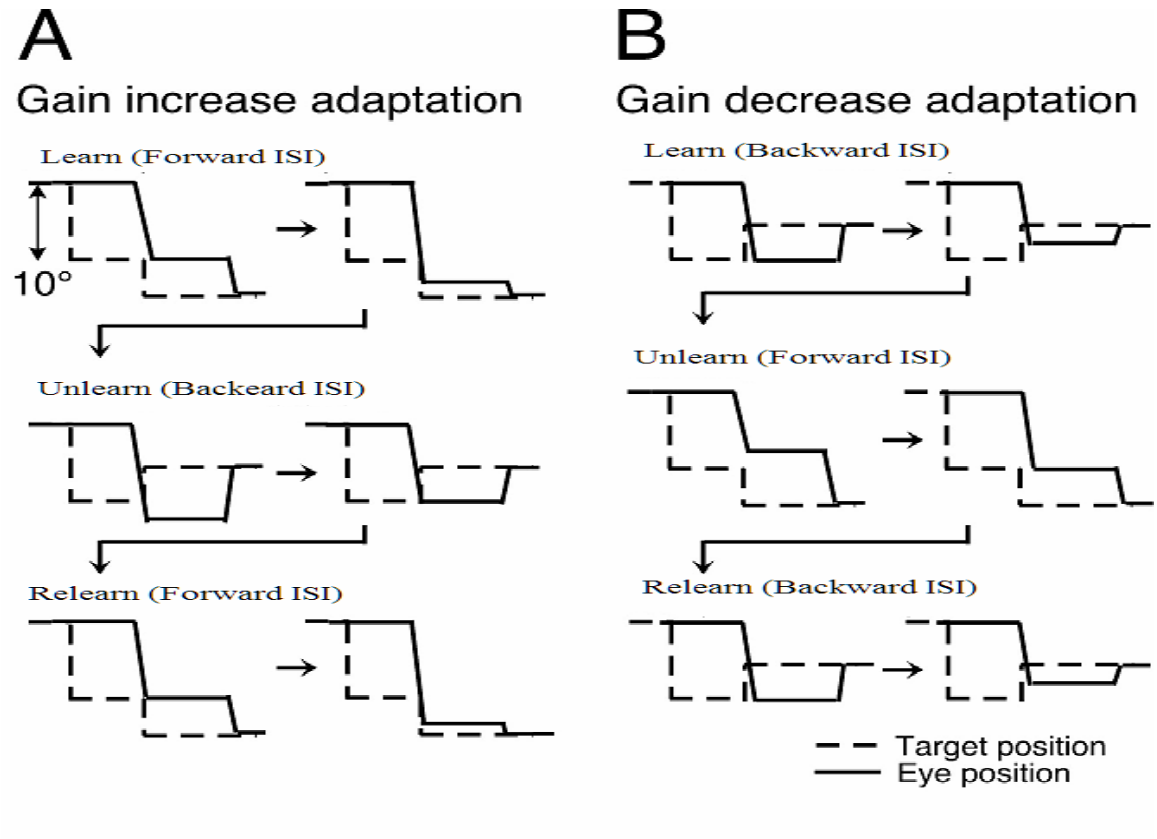

Figure 2.1

Figure 2.1 schematically illustrates these procedures for gain increase (A) and gain decrease (B) adaptation. These paradigms will be called the standard gain-increase or gain-decrease paradigms. The experimental data obtained indicates that indeed the

relearning block of adaptation evolves faster than the initial learning block. Figure 2.2 shows their results from two monkeys.

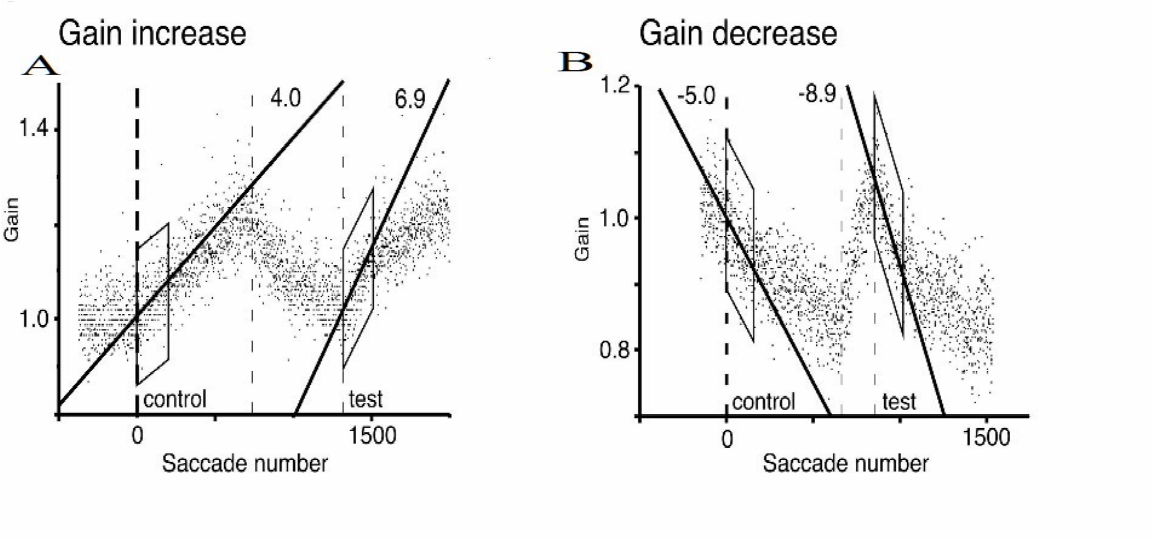

Figure 2.2

Linear regression lines shown are fitted for the first 150 saccades of learn and relearn blocks. The slope of relearn adaptation was larger than that of learning block as shown in both Figure2.2 A for gain-increase and B for gain-decrease. The current state space model with a single learning process cannot explain this result. In other words our single state space model will always show a single learning rate which is independent of the paradigm and the learning history. To demonstrate savings, the system somehow needs to have the ability to remember its previous adaptation state while all apparent manifestations of such adaptation have been wiped out.

To explain this data the authors suggested that savings occurs because the nervous system maintains separate neural mechanisms for gain-up vs. gain-down adaptation (termed the gain-specific model). In other words in this model we have a system whose output is a summation of two independent learning processes. One of the processes can only produce a positive output (gain-up) while the other one can only produce a negative output (gain-down) during adaptation. Figure 2.3 shows the schematic that Kojima et al suggested for their gain-specific model for the adaptation to the standard gain-increase paradigm.

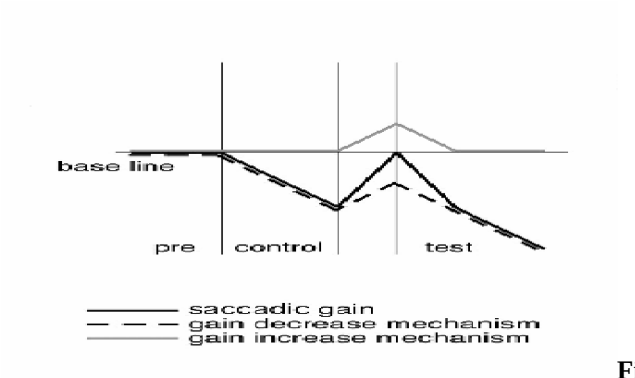

Figure 2.3

It is important to notice that both processes can be active simultaneously in this model during gain-increase or gain-decrease depending on the history of adaptation. Actually this co-activation of gain-up and gain-down processes together is the reason behind the increase in learning rate to almost the double of initial learning block during which only one of the two processes are active. The idea for the gain-specific model can be used to modify our current formulation of online learning as follows:

$$E_N = Zd_N - \overline{Z}_N$$
$$\overline{Z}_N = Z_N^{up} + Z_N^{dn}$$

Gain-specific model 2.1

$$Z_{N+1}^{up} = \max(A_{up} Z_N^{up} + B_{up} E_N ; 0)$$
$$Z_{N+1}^{dn} = \min(A_{dn} Z_N^{dn} + B_{dn} E_N ; 0)$$

Notice that in 2.1 we did not make any assumptions about the learning rate and retention capacity of Z_N^{up} vs. Z_N^{dn}. Based on the relationship between the learning rates and retention capacities of the two processes we can further divide this model into two general categories:

- The model is symmetric gain-specific if $A_{up} = A_{dn}$ and $B_{up} = B_{dn}$

- The model is asymmetric gain –specific if $A_{up} \neq A_{dn}$ or $B_{up} \neq B_{dn}$

Unless explicitly mentioned, from this point on, by gain-specific we imply a symmetric gain-specific model. Figure 2.4 shows how facilitation in learning can be predicted by the gain-specific model while a single state model fails to do so.

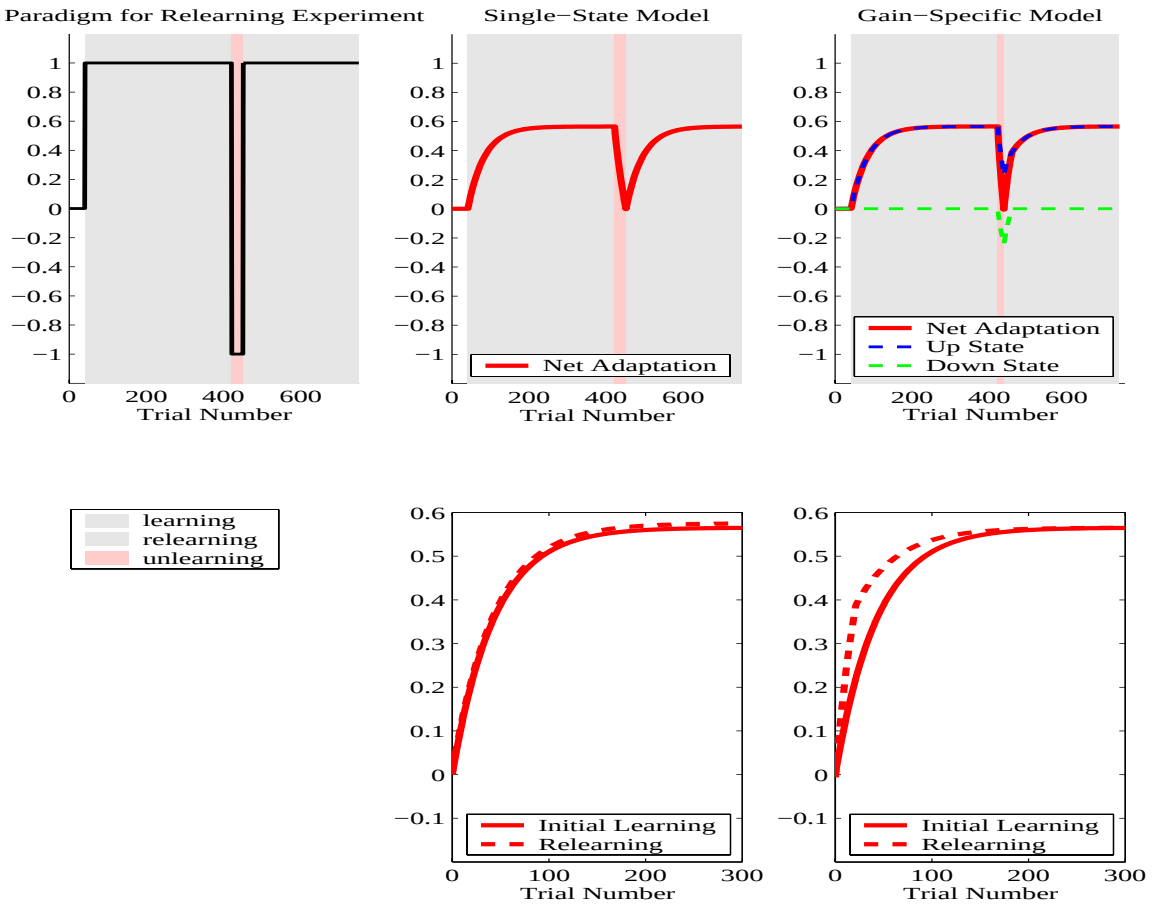

Figure 2.4

The gain-specific model predicts a faster relearning as compared to the initial learning block as can be seen in Figure 2.4 bottom figures. The gain-specific model also predicts that if one interleaves enough washout trials after the unlearning block the effects of savings can be wiped out, and that is what Kojima et al found next in their series of saccade experiments.

Saccade adaptation experiments with zero-error paradigm

In their next experiment, Kojima et al subjected the animals to saccadic steps not coupled with target jumps (zero-error trials) after the saccadic gain returned to ~1.0 by the de-adaptation block (unlearning block). Figure 2.5 shows the result of a gain-increase and a gain-decrease experiment. The fitted regression slopes during the relearning block were not significantly different from the learning block in both conditions in Figure 2.5 A and B.

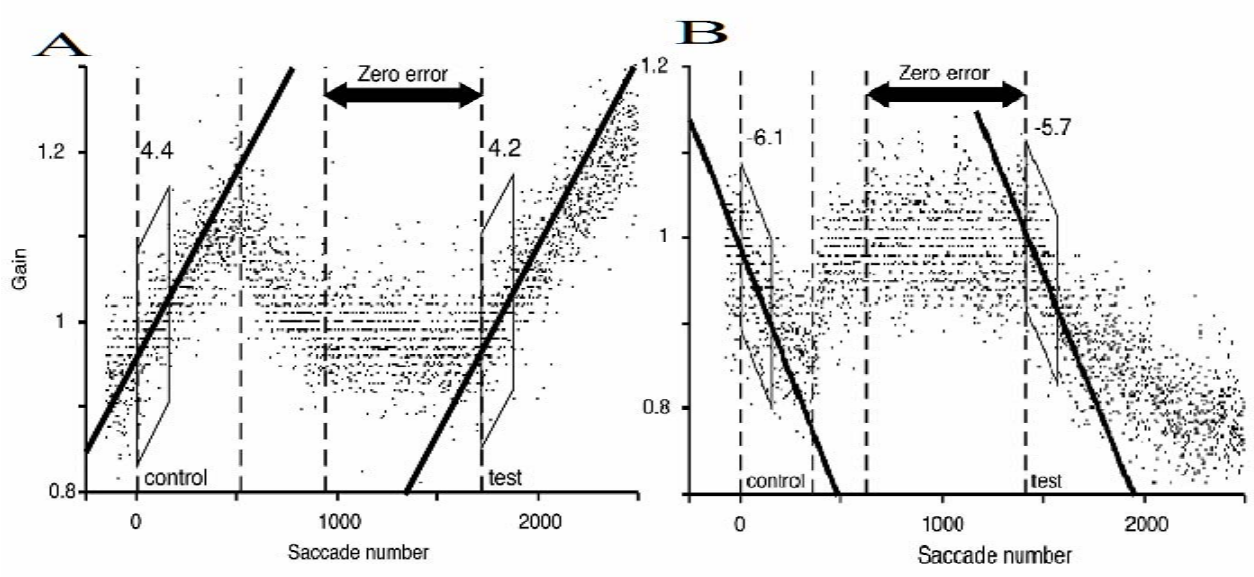

Figure 2.5

The results indicate that the re-adaptation is not accelerated after hundreds of normo-metric saccades (saccades with zero error). This result can be replicated under the gain-specific model as shown in simulation below.

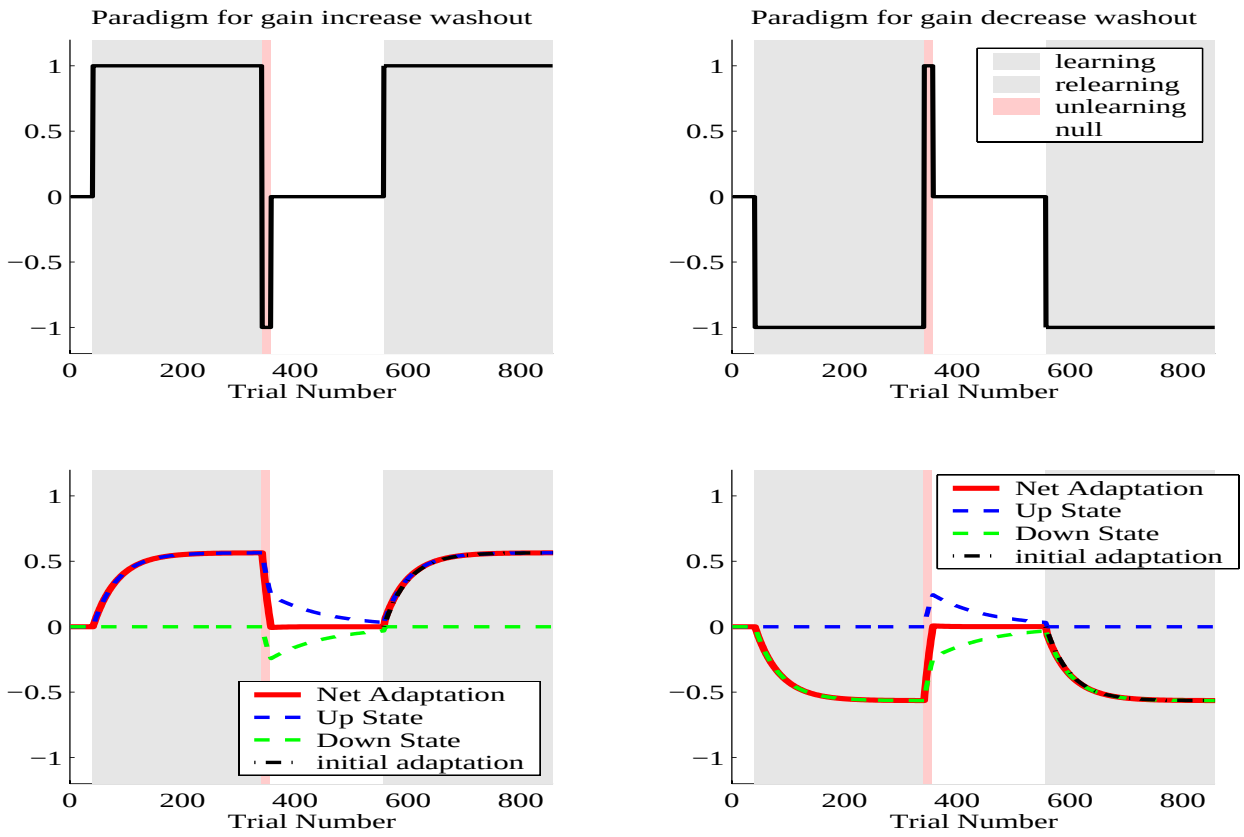

Figure 2.6

Figure 2.6 shows that the initial adaptation curve exactly matches that of relearning and as such it shows that after enough null trials the effects of savings has been washed away. The reason for the lack of savings in this case is simply because the state of gain-up and gain-down processes are back to their state for a naive system as shown in figure 2.6.

Gain-increase experiments with dark paradigm

The next question authors explored was to ask what will happen if the block of zero-error trials is replaced with a period of darkness during which any visual feedback is withheld. Any change in the output of the system in this case should only be mediated by

system's own internal processes without the external input. Figure 2.7 shows the experimental result.

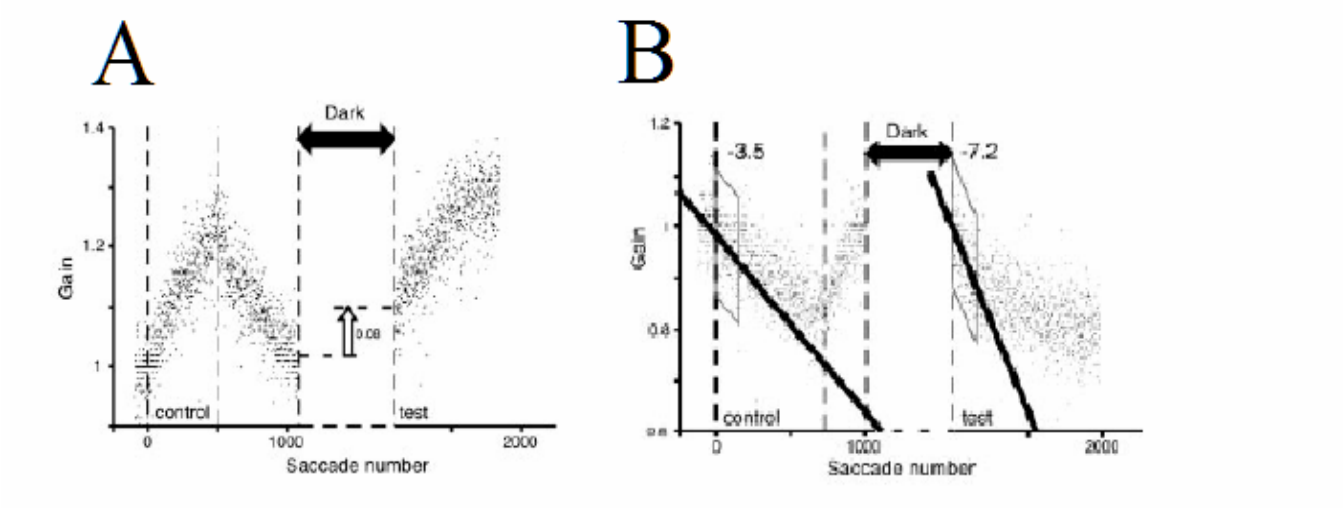

Figure 2.7

It can be seen that although the dark period has the same duration as the zero-error block, there is still facilitation in relearning block. This facilitation shows itself in Figure 2.7 B as the usual increase in the fitted regression slope, but the interesting phenomenon can be seen as a jump-up in performance after the dark period for the gain-increase paradigm in figure 2.7 A.

The gain-specific model proposed by Kojima et al. this time fails to predict the result presented in Figure 2.7. Precisely a symmetric gain-specific model cannot explain the jump-up seen in part A of Figure 2.7 while it may still account for an increase in slope after dark period depending on the dark period duration. Figure 2.8 shows the simulation results for the symmetric gain-specific model.

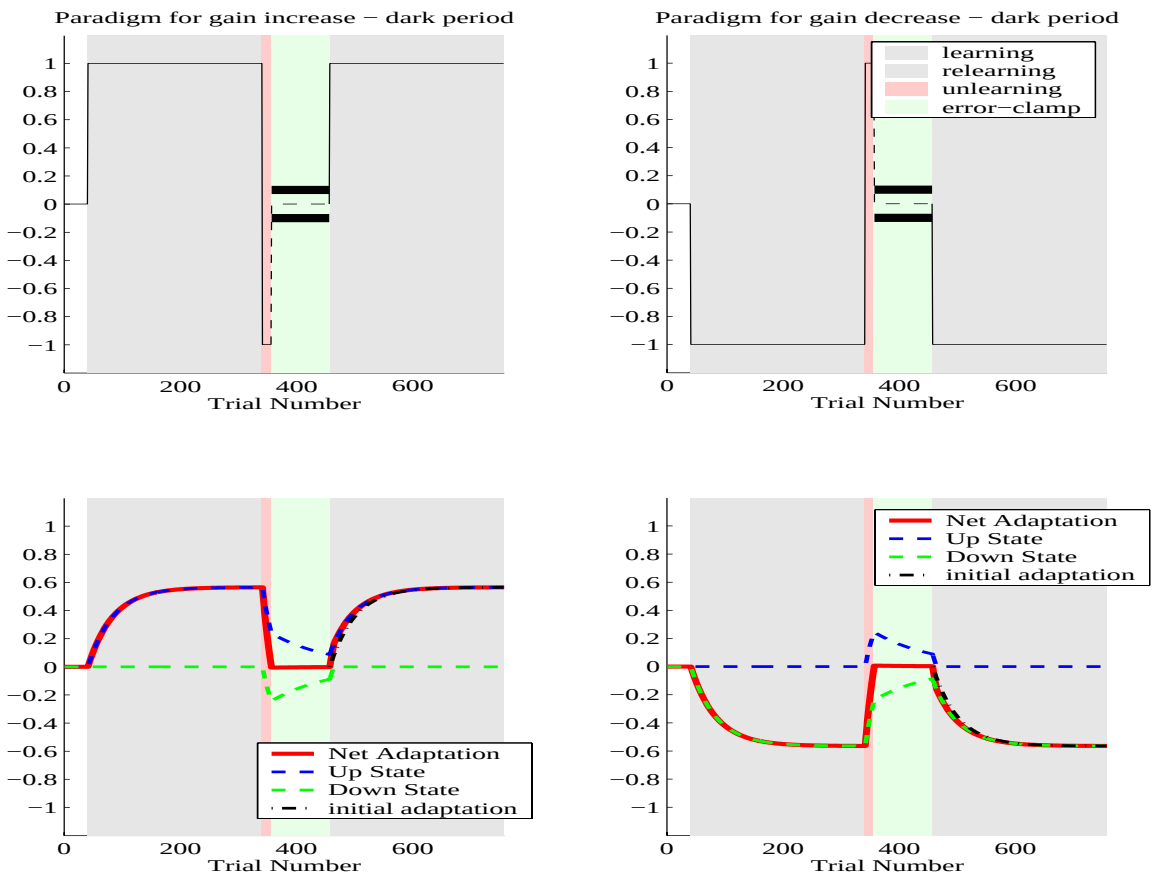

Figure 2.8

Due to the equal decay rate for both processes of gain-up vs. gain down, the sum of theses two processes will always cancel each other out during the zero-feedback after the gain is brought back to zero and thus a jump-up (or jump-down) will never occur using symmetric gain-specific formulation.

What if we assume an asymmetry in the gain-up vs. gain-down processes?

Specifically speaking what if we have a gain-up process which has a slower forgetting rate than the gain-down process. Will such modification helps predict the result Kojima et al observed in their dark paradigm? Figure 2.9 shows the results of such simulation.

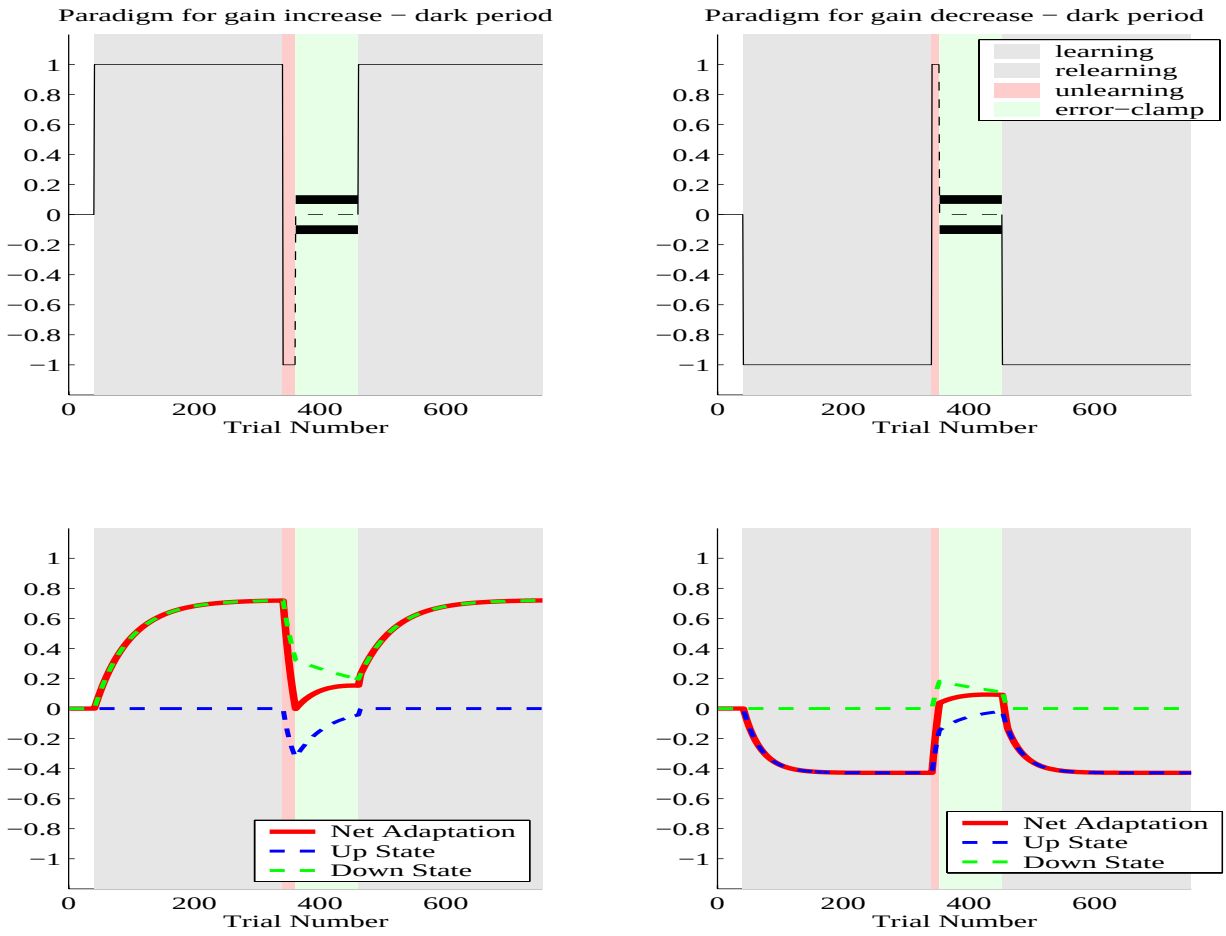

Figure 2.9

The asymmetric gain-specific model always predicts a jump in the direction of the process with a slower forgetting rate. Therefore this modification can go so far as to explain half of the results obtained in the dark paradigm. That is it explains the jump-up in the gain-increase condition but also falsely predicts the same jump-up to happen in the gain-decrease condition.

2.3 Multi-timecourse model for motor adaptation

We suggest an alternate model for the savings effects seen by Kojima et al. in their study of saccade adaptation. We hypothesize that during adaptation, performance is a reflection of two motor learning systems in the brain. The first process is a slow learner which has a high retention capacity and can work either to facilitate or delay learning based on the prior experience in the environment. The second process on the other hand has a high learning rate but a poor retention of what is being learned, providing the advantage of short term compensation for the rapidly changing environment. We first show that the proposed model (termed a Multi-timecourse model) like the model proposed by Kojima et al can account for the effects of savings observed in the saccade adaptation task. We then show that the two models, i.e., the multi-timecourse model and the gain-specific model, can make different and testable behavioral predictions.

Our two dimensional state-space model can be written as follows:

$$E_N = Zd_N - \overline{Z}_N$$
$$\overline{Z}_N = Z_N^f + Z_N^s$$

Multi-timecourse model 2.2

$$Z_{N+1}^f = A_f Z_N^f + B_f E_N$$
$$Z_{N+1}^s = A_s Z_N^s + B_s E_N$$

Z_N^f represents the fast process in the model which has a high learning rate B_f but a small retention capacity A_f. Z_N^s represents the slow process in the model which has a lower learning rate B_s but a high retention capacity A_s. In other words we have:

$$A_f < A_s \quad \text{and} \quad B_s < B_f$$

Figure 2.10 shows in simulation how our multi-timecourse model predicts savings in double reversal learning paradigm discussed before (learn-unlearn-relearn). Notice that the amount of savings observed is a monotonically decreasing function of the number of unlearn trials. This relationship is predicted by both our multi-timecourse model as well as by Kojima's gain-specific model. In fact both models here predict the pattern of savings as well as savings washout if enough base line trials are interspersed between learn and relearn blocks. The single state-space model does not predict any change in the learning rate as a function of learning history as illustrated below.

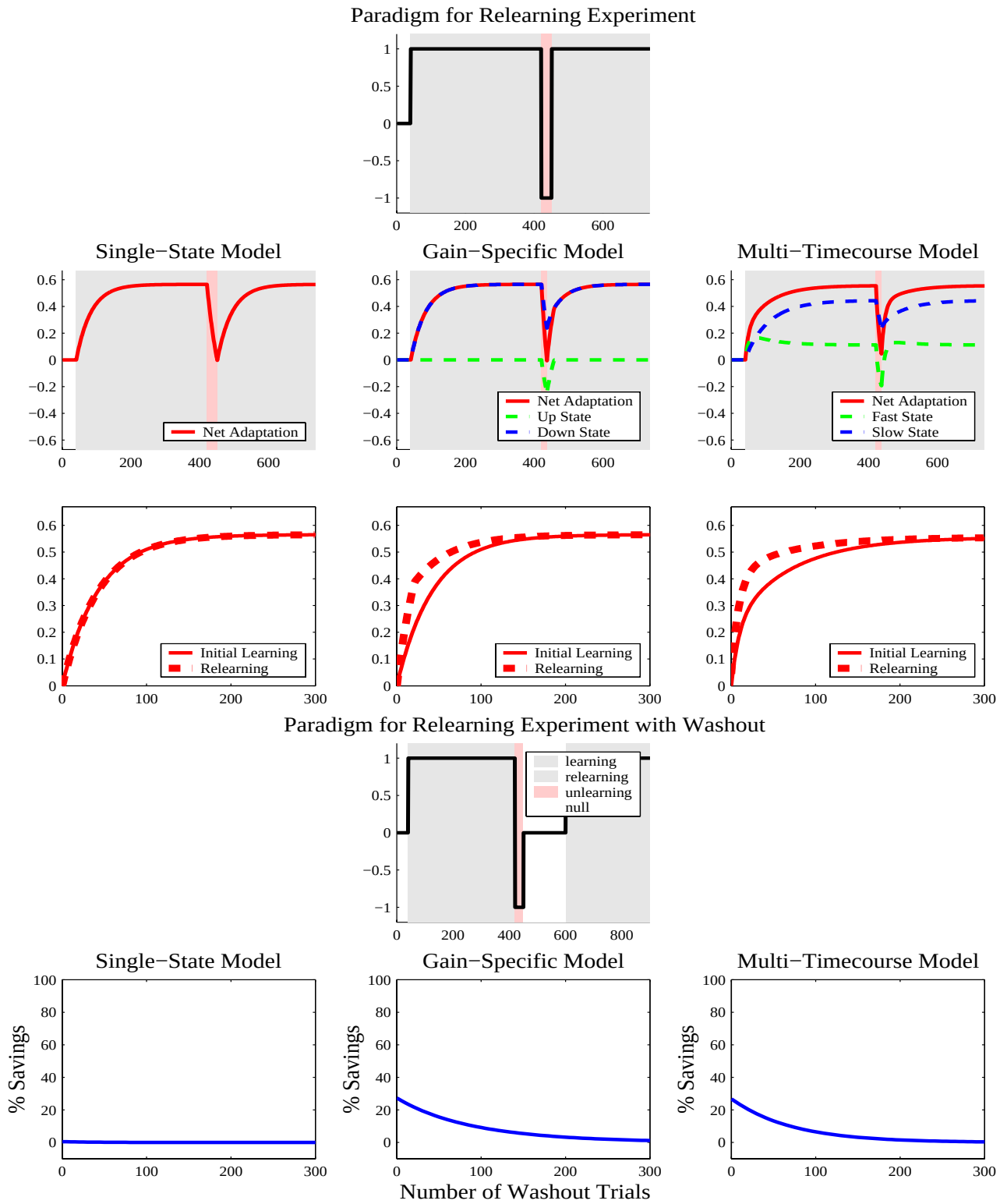
Paradigm for Relearning Experiment
Single−State Model
Gain−Specific Model
Multi−Timecourse Model
Net Adaptation
Up State
Down State
Fast State
Slow State
Initial Learning
Relearning
Paradigm for Relearning Experiment with Washout
learning
relearning
unlearning
null
% Savings
Number of Washout Trials

Figure 2.10

Also the multi-timecourse model correctly predicts the jump-up effect seen after zero-feed back trials termed as dark period. This model predicts the jump-up to always occur in the direction in which the subjects received most of their training. In our paradigm this is the gain learned during the long learning period i.e. a jump-up for gain increase and a jump down for gain decrease. This model as well as the asymmetric gain-specific model can only explain the result of gain-increase adaptation where there was a jump-up in performance.

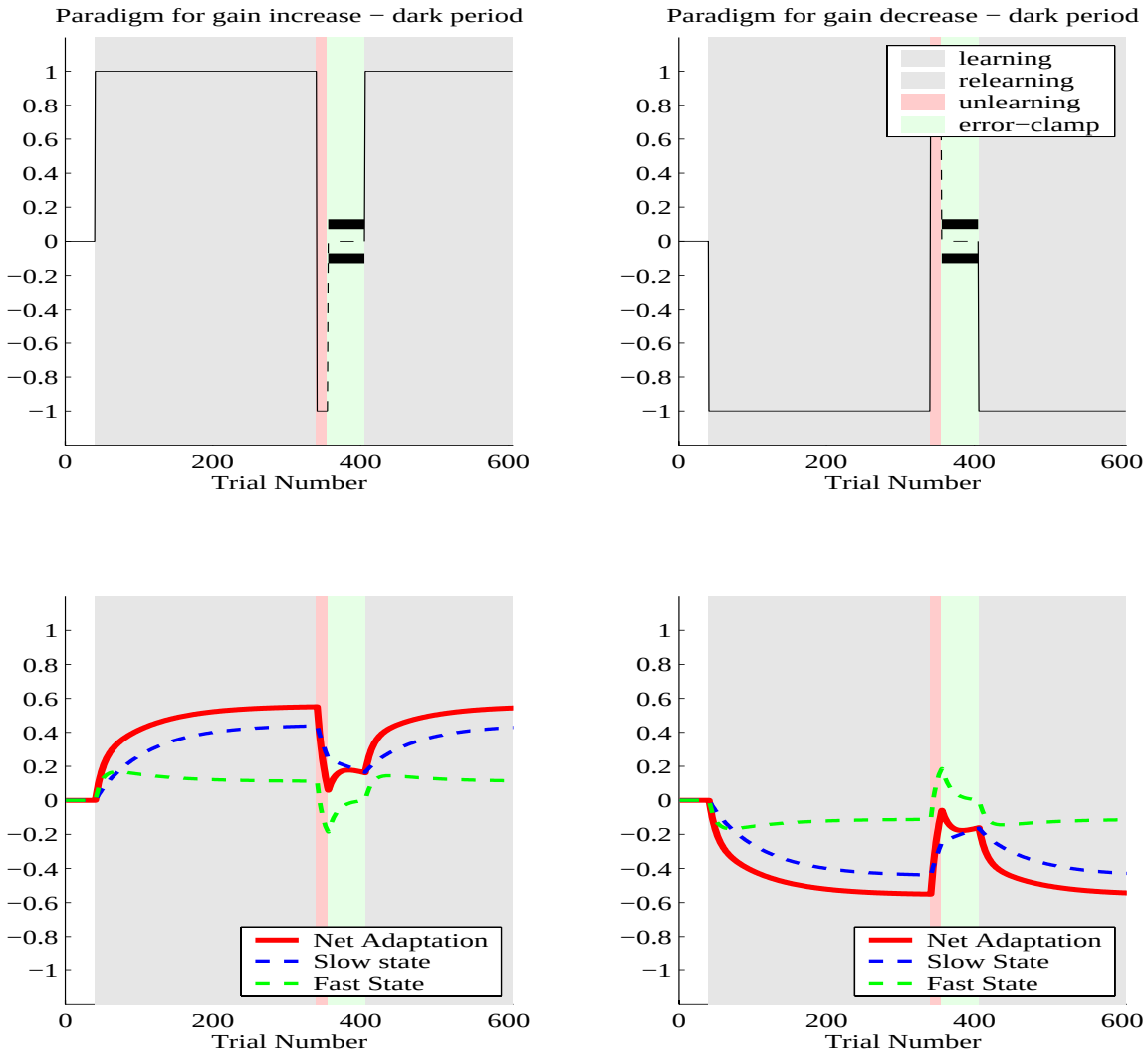

Figure 2.11

While the multi-timecourse model predicts a jump in both gain increase and gain decrease conditions, Kojima et al study of saccade adaptation reports such a jump in performance to happen only for the gain-increase paradigm. Therefore our multi-timecourse model of learning also cannot explain the result of the dark period in its full totality.

As yet both the multi-timecourse model proposed by us as well as the asymmetric version of the gain-specific model hypothesized by Kojima et al seem to be able to explain a large portion of the saccade adaptation data. In order to contrast these two models' behavioral predictions and to see which one can stand up to further experimentations, we considered an error clamp experiment discussed below.

2.4 Error clamp experiment

Using the experimental setup discussed in detail in previous chapter we asked 14 right handed subjects (9 females and 4 males) to participate in our error-clamp study. The subjects were trained first in two sets of "null" (60 out and 60 back movements each set) where robot motors were not turned on. The back movements throughout the experiments were performed in the force channel. The reason to have "null" trials was to train subjects to learn the passive dynamics of the robot arm and the basics of the task. We randomly included force channel trials (1 in 6) in the second set of "null" in order to obtain a baseline reading for the lateral forces subjects produced in the "null" trials. We then

divided our subjects into two groups of 7 subjects each. The first group of subjects was exposed to a clockwise velocity dependent curl force field while the second group had to learn a counterclockwise field. We included occasional (9 in 60) force channel trials on out movements to monitor the learning of the force field by the subjects. Both groups of subjects were trained in the same field for two sets. In the next set the subjects moved in the same field for another 10 trials before the direction of the force reversed for 15 trials. During these 15 trials-called the unlearning trials- the group that was trained in a clock-wise field experienced a counter clock-wise field and vise versa.

After this brief period of unlearning both groups moved out and back for the rest of the set (35 trials) and another full set (60 trials) in the force channel. We used the force channel trials to monitor the motor output of the subjects during a zero-feedback period which resembles the dark period in saccade adaptation.

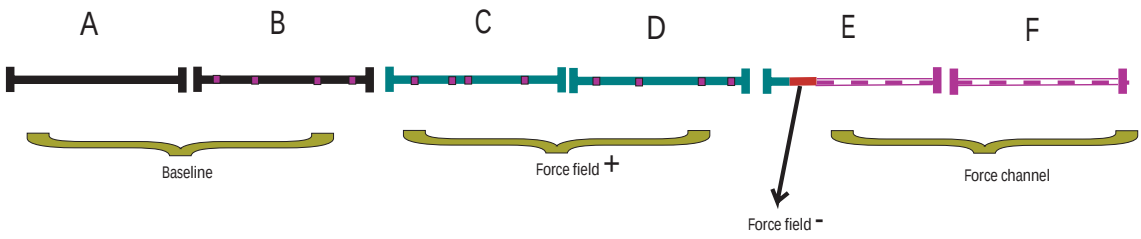

Figure 2.12 Experimental paradigm sets A to F ; black represent baseline trials, light green is the first force field, red is the washout using the opposite force field and purple represents force channel trials.

Before moving on to the experimental results lets look at the predictions of both multi-timecourse model and gain-specific model in this paradigm. Figure 2.13 shows the experimental paradigm and the predictions of both models.

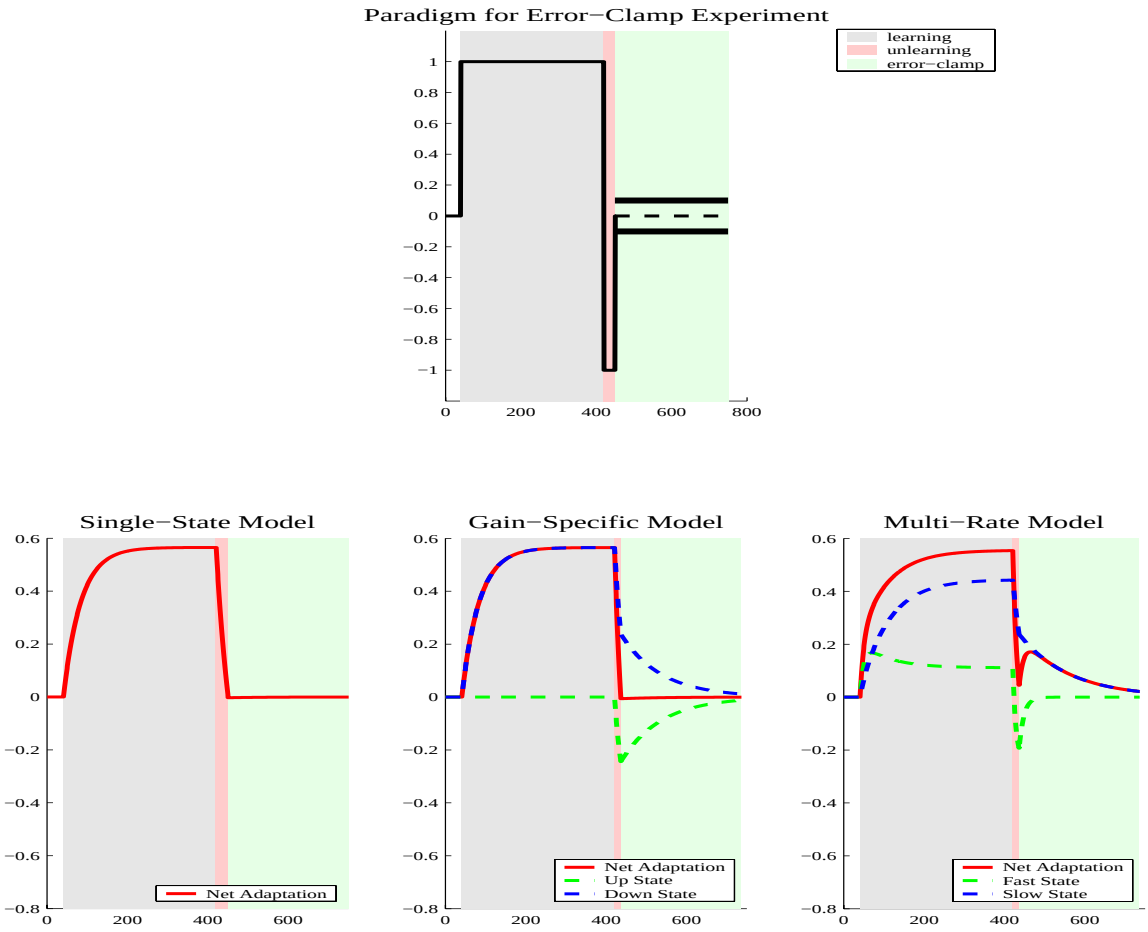

Figure 2.13

One can see that the gain-specific model maintains zero motor output throughout the error-clamp trials after the short unlearning period. On the other hand the multi-timecourse model for motor adaptation makes an interesting prediction about the

evolution of motor output during error-clamp trials. As seen in figure 2.13 our model predicts that the motor output starts from zero or a slightly negative value but shows a positive rebound toward the learned force field and then dies out slowly back toward baseline. It is as if our model shows a transient recovery of the prior training in spite of the apparent washout of the learned skill.

As such this experimental paradigm can provide a powerful way of distinguishing between the two models, and as a matter of fact it does. The following figure shows the combined results of our analysis for the two groups who participated in the study. Similarity index (SI) is used as the measure to quantify learning.

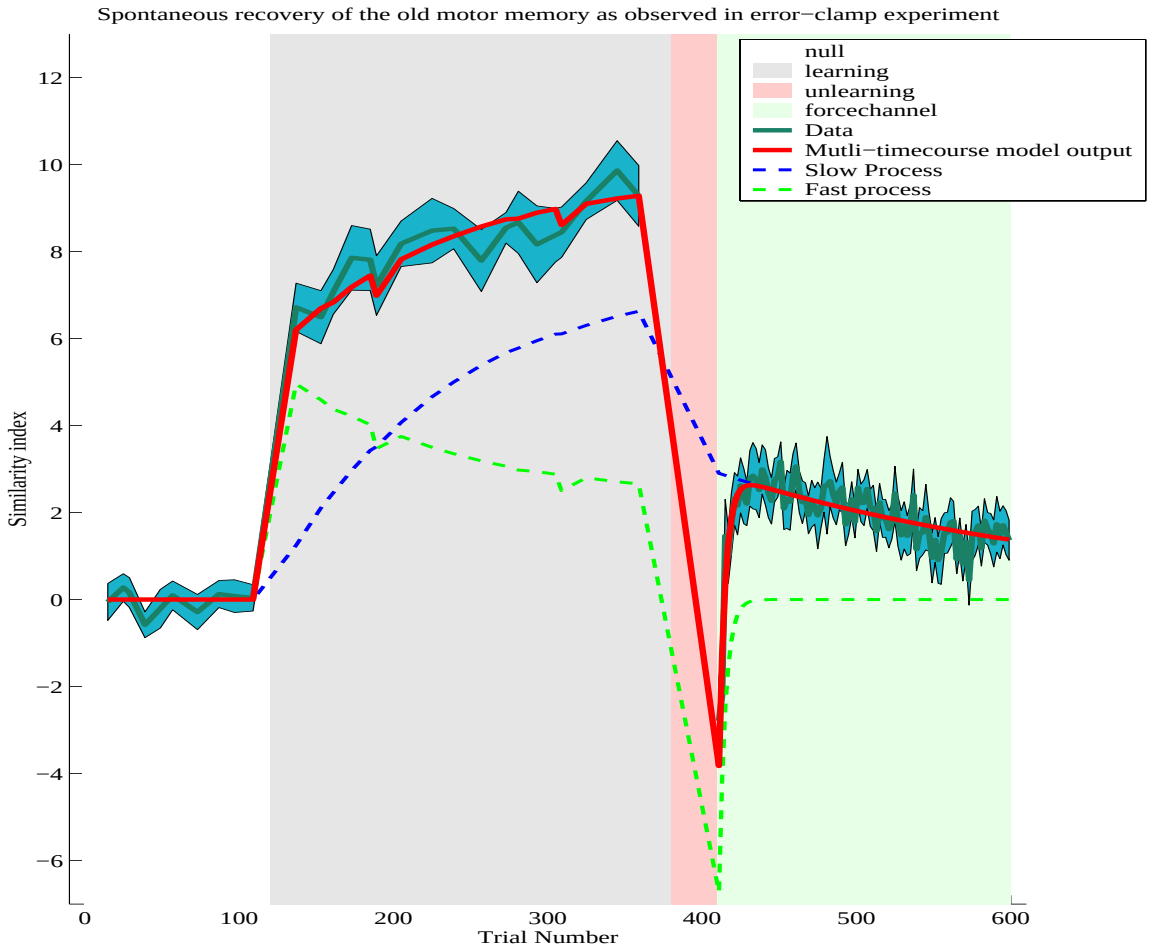

Figure 2.14

As predicted by our multi-timecourse model the data averaged across all subjects show a clear rebound toward the learned state which reaches its peak of about 20% recovery in about 25-30 trials into the error clamp block. We fitted the data with our multi-timecourse model as shown above. The fitted values are as follows:

$A_f = 0.611 \quad B_f = 0.275 \quad A_s = 0.992 \quad B_s = 0.019 \qquad R^2 = 0.97$

The multi-timecourse model fits the experimental data pretty well. One can see that the fast process has a learning rate which is ~ 15 times higher than the slow system. On the other hand the forgetting of the fast process is ~50 times higher than the slow system as well. Forgetting rate is defined as one minus the retention capacity that is 1-A for each process.

Figure 2.15 shows the data for two single subjects one from each group. It shows the actual forces of the two subjects during the movement as well as the force profiles for 100% compensation of the force field. One can observe the increasing similarity of the subjects' force output profile compared to the force field as subjects experience more trials in the field. This figure also shows the transient rebound of lateral forces toward the learned state following unlearning trials as subjects move in the force channel.

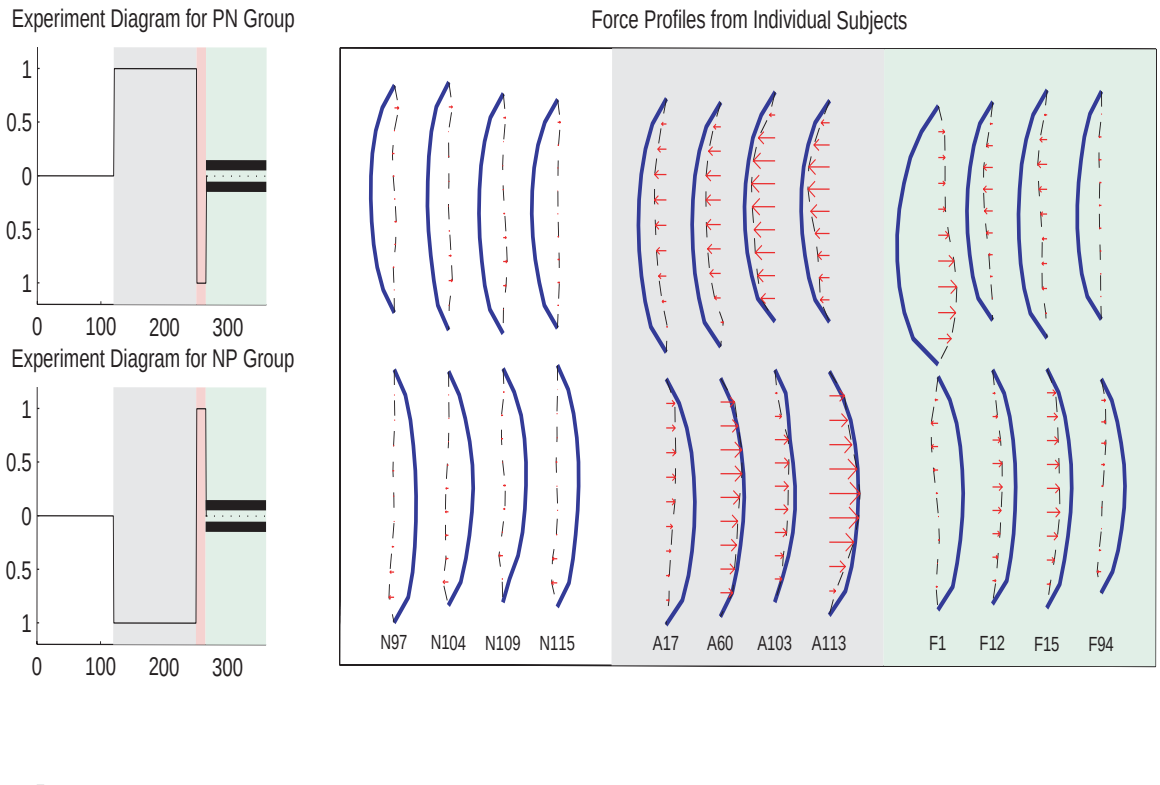

Figure 2.15

Figure 2.14 and 2.15 together show that the gain-specific model fails to correctly model the process of motor adaptation while the multi-timecourse model provides an intuitive explanation for the motor adaptation phenomena seen so far. But can we rescue the gain-specific model once again by assuming an asymmetry between gain-up vs. gain down processes?

The following plot shows the predictions of the asymmetric gain-specific model for both of our group conditions. The asymmetric gain-specific model predicts a rebound which is always toward the process with better retention capacity and is independent of the learning history and thus it predicts the same rebound for both groups. This is not

consistent with the data for the two single subjects shown in figure 2.16 where the rebounds are in opposite direction.

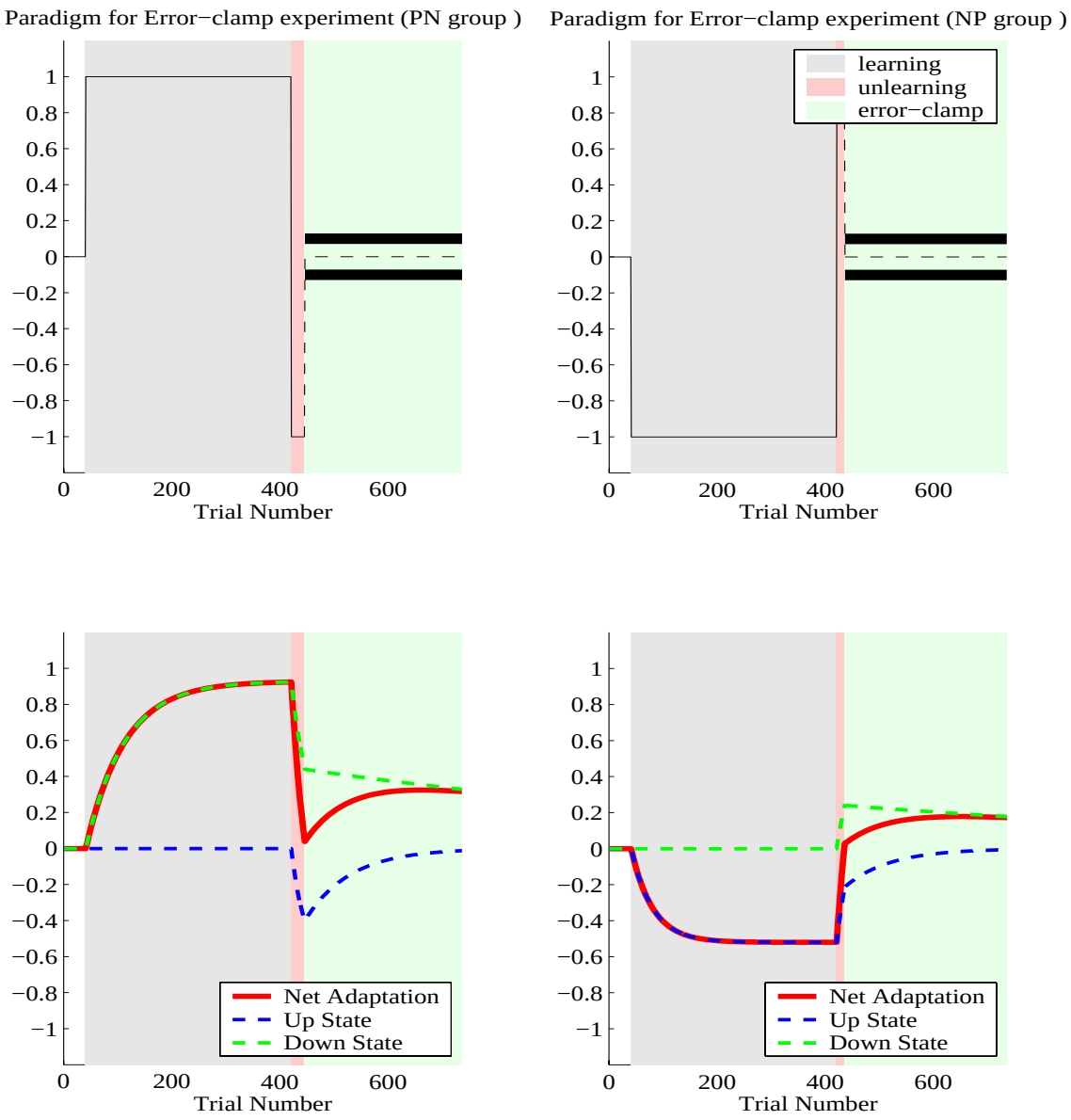

Figure 2.16

To further explore this prediction of the asymmetric gain-specific model we looked at the results of each group of subjects separately. Figure 2.16 shows the experimental results for both groups. The rebounds for the two groups are not in the same direction. They are oppositely directed and are toward their trained state. So gain-specific model in both symmetric and asymmetric format is unable to explain our data.

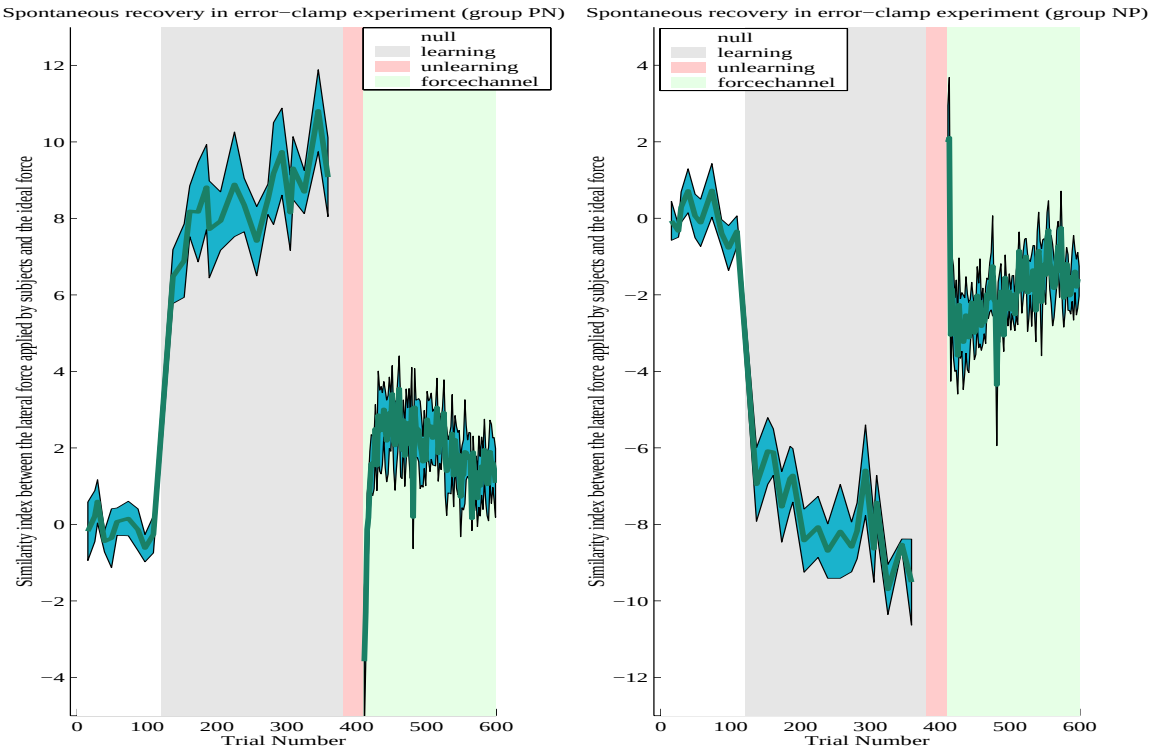

Figure 2.17

The bar-plot below shows that the effect of rebound is significant for each group separately as well as both groups put together.

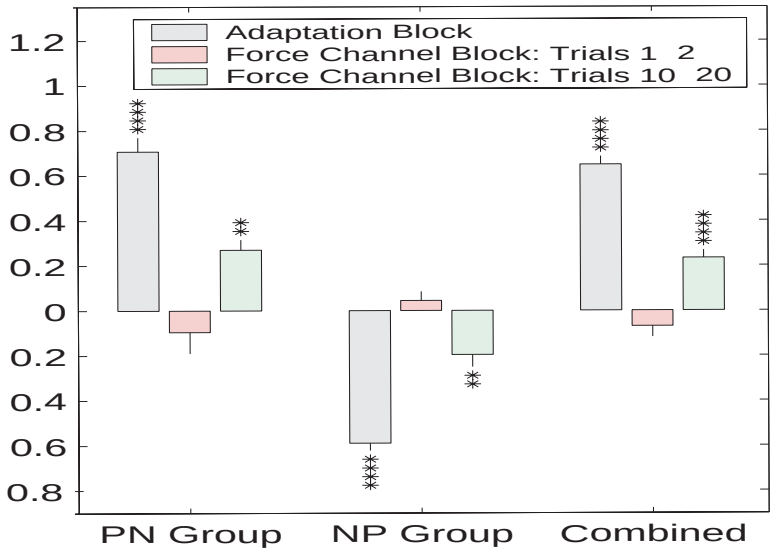

Figure 2.18

2.5 Additional results explained

Our multi state-space model can be used to explain some of the other important motor behaviors in human subjects which are hard to understand in the framework of present models of trial to trial motor adaptation.

One such well-known effect is “Anterograde interference” which refers to the reduction of the learning rate in adaptation to a novel field after adaptation to an opposite field has already taken place. Several studies have shown that the time constant for an initial motor adaptation is faster than the time constant for subsequent adaptation to the oppositely directed adaptation stimulus (Shadmehr and Brasher-Krug 1997, Thoroughman and Shadmehr 1999, Davidson and Wolpert 2004). Figure 2.19 shows how prior adaptation to

a clock-wise field make subjects slower in their subsequent learning of a counter clock-wise field when compared to naive subjects learning the same counter clock-wise field.

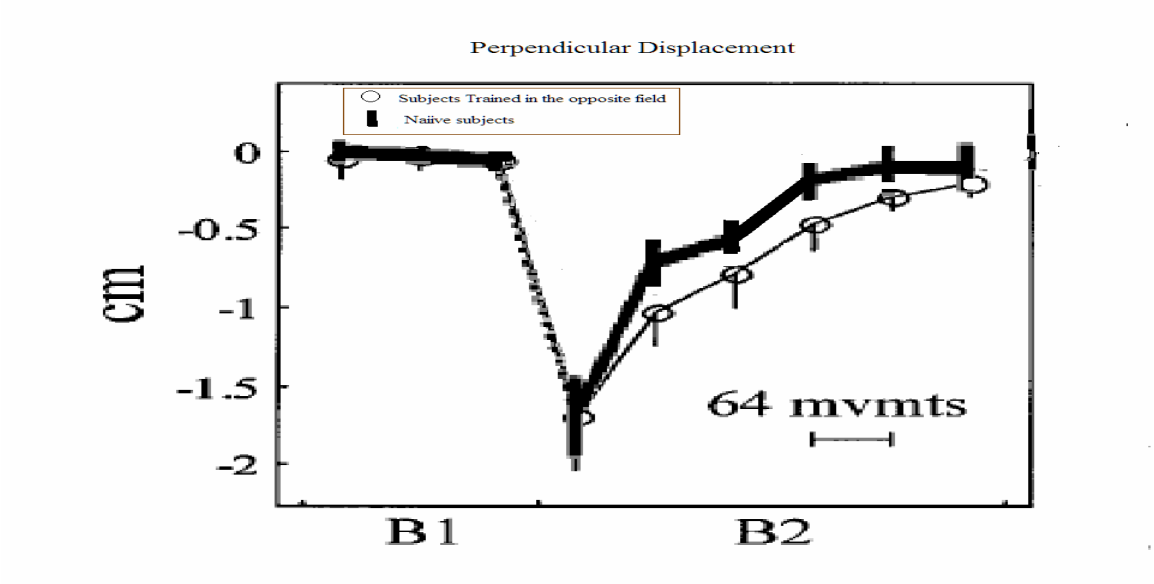

Figure 2.19

We have adjusted for the errors of naive subjects in the first block of 64 movements in the new field equalizing it to the first block of errors for the subjects trained in the opposite field in order to make the difference in the adaptation time constants easily observable.

The multi-timecourse model can easily explain this slowing down in adaptation caused by prior training in the opposite field. In fact the effect of the anterograde interference is nothing but the opposite of the effect seen in savings. Below we show how multi-timecourse model produces the effect seen in anterograde interference. Gain-specific model proposed by Kojima et al cannot predict the elongation of the time constant in this

case but instead predicts that we should observe the same facilitation in learning caused by savings in the case of adaptation to the opposite field which is inconsistent with the experimental evidence.

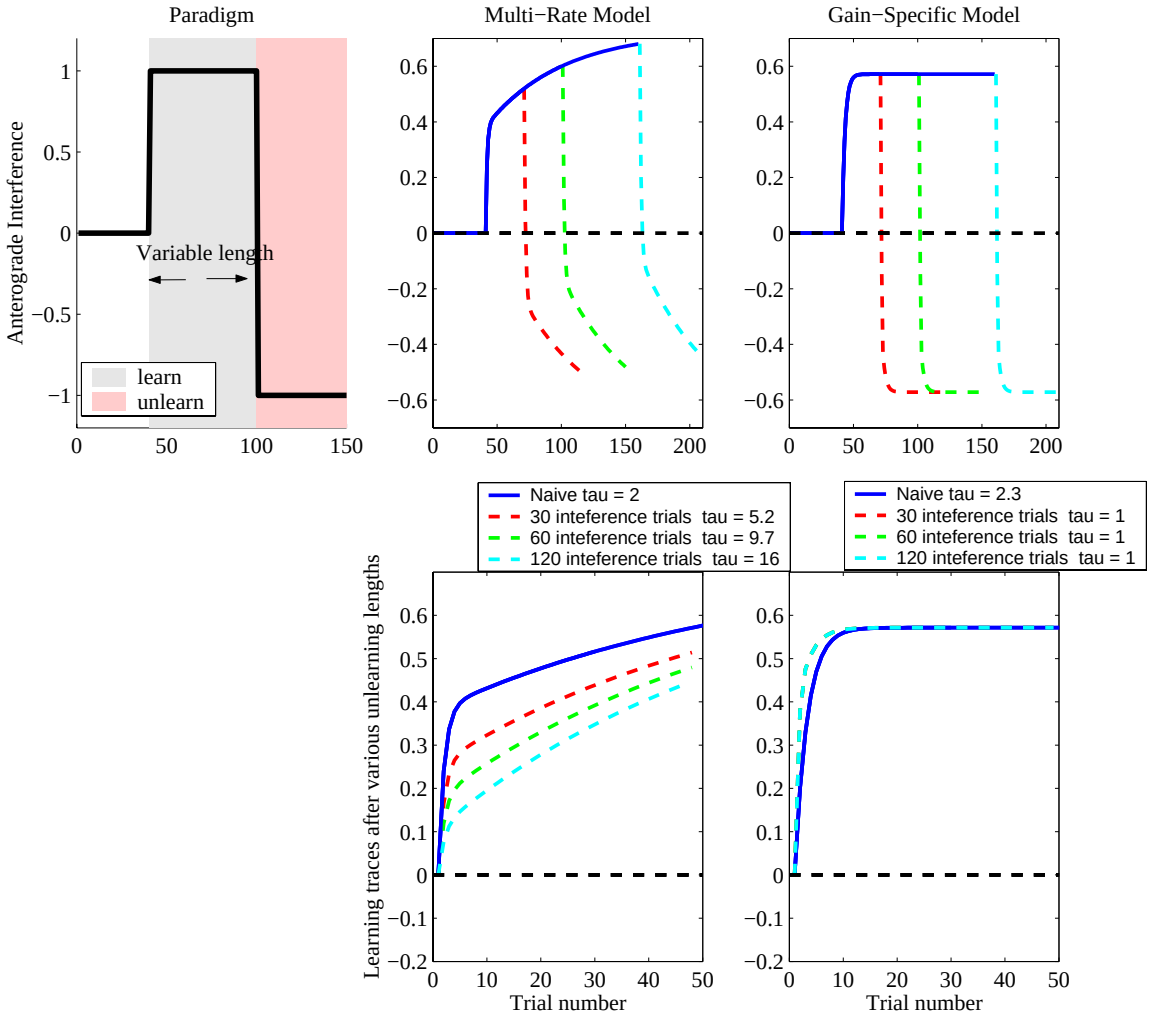

Figure 2.20

The reason for the slower adaptation seen in anterograde interference is explained by our multi-time course model of motor learning. Training in the opposite field biases the slow

module of the multi-timecourse model against learning the new field. Therefore it will take longer for the system to reach its asymptotic state in which both processes are at steady state. It also predicts that the slowness of relearning is dependent on the number of unlearning trials, that is the longer the interference trials are; the slower the process of relearning becomes (see fig 2.20).

The same argument can be used to explain the effects of savings, that is biasing of the slow system in favor of the learning is the underlying mechanism for savings and its various manifestations.

It is also known that the process of de-adaptation to the baseline or adaptation to a reduced version of a force field happens faster and with smaller time constants than the initial adaptation to a novel field (Davidson and Wolpert 2004). This effect also can be accounted for in the light of our new formulation for motor adaptation. Figure 2.20 shows the simulations using the multi timecourse as well as the gain-specific model. Both models seem to be able to explain the general trend of faster de-adaptation and down-scaling of the force field. Yet Wolpert's group observed that down-scaling happens even faster than de-adaptation and this subtlety of the experimental results can only be captured by the multi-timecourse model and not by the gain-specific model. Looking at the τ fitted to the learning curves one can see that the down-scaling in the multi-timecourse model happens faster than the de-adaptation which is consistent with experimental data while for the gain-specific model it is the other way around.

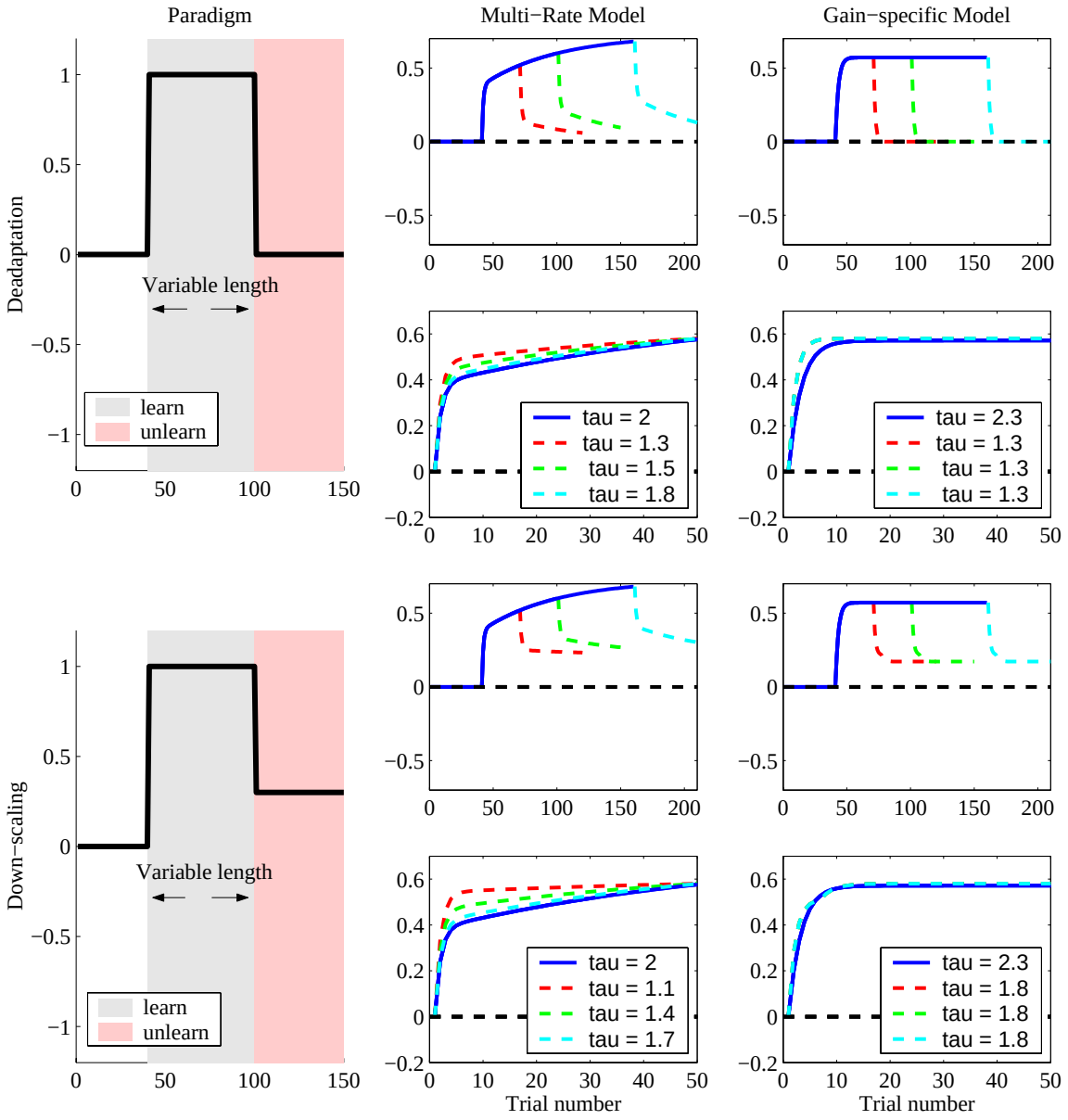

Figure 2.21

It is needless to say that a state-space model with a single learning process can not reproduce any of these motor learning effects which involve changes in the time course of learning.

2.6 Multi-timecourse model: Current version and the Alternatives

In this section we are going to take a closer look at the multi-timecourse model and study its properties from a theoretical point of view. In our current formulation of the model we have:

$$E_N = Zd_N - \bar{Z}_N \qquad \text{Output equation}$$
$$\bar{Z}_N = Z_N^f + Z_N^s$$

Multi-timecourse model 2.3

$$Z_{N+1}^f = A_f Z_N^f + B_f E_N \qquad \text{Update equation}$$
$$Z_{N+1}^s = A_s Z_N^s + B_s E_N$$

We can solve for the solution to equation 2.3 directly by rearranging it into the standard format for the first order discrete differential equations (See Chapter I section 1.4). We have:

$$\Delta\vec{Z} = \mathrm{H} \times \vec{Z} + \Phi \times Zd$$
$$\Delta\vec{Z} = \begin{bmatrix} Z_{N+1}^f - Z_N^f \\ Z_{N+1}^s - Z_N^s \end{bmatrix} \qquad \mathrm{H} = \begin{bmatrix} A_f - 1 - B_f ; -B_f \\ -B_s ; A_s - 1 - B_s \end{bmatrix} = \begin{bmatrix} A_f - B_f ; -B_f \\ -B_s ; A_s - B_s \end{bmatrix} - \begin{bmatrix} 1;0 \\ 0;1 \end{bmatrix} = \alpha - I$$
$$\vec{Z}_N = \begin{bmatrix} Z_N^f \\ Z_N^s \end{bmatrix}$$

$$\Phi = \begin{bmatrix} B_f \\ B_s \end{bmatrix} \qquad \alpha = \begin{bmatrix} A_f - B_f ; -B_f \\ -B_s ; A_s - B_s \end{bmatrix} \qquad 2.4$$

The solution is the same as for the single state model except that it is represented in matrix format:

$$\vec{Z}_N = \begin{bmatrix} Z_N^f \\ Z_N^s \end{bmatrix} = -(I - (I + \mathrm{H})^N) \times \mathrm{H}^{-1} \Phi Z d = (I - \alpha^N) \times \vec{Z}_\infty \qquad 2.5$$
$$\vec{Z}_\infty = -\mathrm{H}^{-1} \Phi Z d$$

If we further write $\vec{Z}_\infty$ in terms of the eigenvectors of α, we get:

$$\vec{Z}_\infty = \eta \vec{V}_1 + \mu \vec{V}_2$$

$$\lambda_{1,2} : Eigenvalues \quad \vec{V}_{1,2} : Eigenvectors$$

$$\vec{Z}_N = \eta \times (1 - \lambda_1^N) \times \vec{V}_1 + \mu \times (1 - \lambda_2^N) \times \vec{V}_2 \qquad 2.6$$

$$\eta = \vec{Z}_\infty \cdot \vec{V}_1 \quad \text{and} \quad \mu = \vec{Z}_\infty \cdot \vec{V}_2$$

The trial constants of the system consist of the two eigenvalues of the matrix $\alpha = \begin{bmatrix} A_f - B_f ; -B_f \\ -B_s ; A_s - B_s \end{bmatrix}$. As long as the forgetting rates and retention capacities of the two processes remain between zero and one, these eigenvalues remain between -1 and 1 and therefore system maintains stable behavior.

As with the single state model in Chapter I, we can illustrate the mechanisms of our two-state model with an electrical circuit:

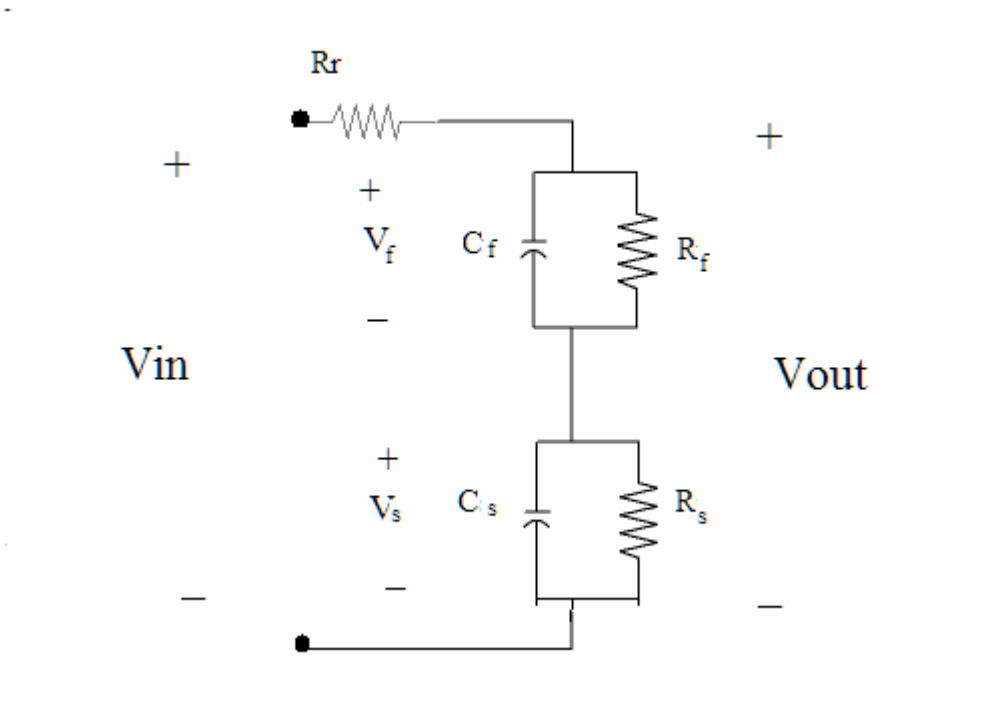

Figure 2.22

Using Kirchhoff's voltage and current equations and rearranging the variables we have:

$$V_f(t+\Delta t) = (1-\frac{\Delta t}{C_f R_f}) \times V_f(t) + \frac{\Delta t}{C_f R_r}(V_{in} - V_f(t) - V_s(t))$$
$$V_s(t+\Delta t) = (1-\frac{\Delta t}{C_s R_s}) \times V_s(t) + \frac{\Delta t}{C_s R_r}(V_{in} - V_f(t) - V_s(t)) \qquad 2.7$$

Comparing 2.7 with 2.3 we have

$$A_f = 1-\frac{\Delta t}{C_f R_f} \qquad B_f = \frac{\Delta t}{C_f R_r} \qquad 2.8$$

$$A_s = 1-\frac{\Delta t}{C_s R_s} \qquad B_s = \frac{\Delta t}{C_s R_r}$$

At steady state the transfer function of the circuit in figure 2.22 is:

$$V_{out}(t) = \frac{R_f + R_s}{R_s + R_f + R_r} \times V_{in} = \frac{\frac{R_f}{R_r} + \frac{R_s}{R_r}}{\frac{R_f}{R_r} + \frac{R_s}{R_r} + 1} \times V_{in} \qquad 2.9$$

$\lim t \to \infty$

Using the relationship in 2.9 and 2.8 we have:

$$\overline{Z}_{\infty} = \frac{\frac{B_f}{1-A_f} + \frac{B_s}{1-A_s}}{\frac{B_f}{1-A_f} + \frac{B_s}{1-A_s} + 1} \times Zd = \frac{\frac{\beta_f}{1-\beta_f} + \frac{\beta_s}{1-\beta_s}}{\frac{\beta_f}{1-\beta_f} + \frac{\beta_s}{1-\beta_s} + 1} \times Zd \qquad 2.10$$

Using the same analogy for the steady state values of the fast and slow system we have:

$$Z_{\infty}^{s} = \frac{\frac{B_s}{1-A_s}}{\frac{B_f}{1-A_f} + \frac{B_s}{1-A_s} + 1} \times Zd \qquad Z_{\infty}^{f} = \frac{\frac{B_f}{1-A_f}}{\frac{B_f}{1-A_f} + \frac{B_s}{1-A_s} + 1} \times Zd \qquad 2.11$$

Using the fitted values of A and B to our data and substituting in 2.11 and 2.10 we have:

$$\overline{Z}_{\infty} = 0.755 \times Zd \qquad Z_{\infty}^{s} = 0.582 \times Zd \qquad Z_{\infty}^{f} = 0.173 \times Zd \qquad 2.12$$

The results in 2.9 shows that using our fitted values subjects can at most compensate for ~ 75% of the force field and of this amount ~ 58% is learned by the slow system while the fast process contributes only to ~ 17% of total learning. This means that by the end of a long training in a force field the slow process learns more than 3 times more than the fast process.

In our current model both fast and slow processes learn independently from the external errors and the final motor output is considered to be the sum of their values together. This is shown in the output and update equations in 2.3. For reasons that become clear later in this section we call this version of model, the “multi-timecourse model with parallel updating”.

People who have studied motor learning in cerebellum often think of fast and slow processes in learning to have a sequential updating rule. In a recent study of savings in classical eye blink conditioning in rabbits by Medina et al, the authors hypothesized two sites of plasticity one in the cerebellar cortex and the other in cerebellar nucleus. During eyelid conditioning, the first changes involved decreased activity of Purkinje cells during the CS . This decrease was produced by the US activating the climbing fiber input and lead to the induction of LTD at CS-activated granular to Purkinje synapses as soon as acquisition training was started. In contrast, induction of plasticity at the mossy fiber to nucleus synapses was not directly under the control of inputs activated by conditioning stimuli but was instead under the control of Purkinje cells (see Figure 2.23).

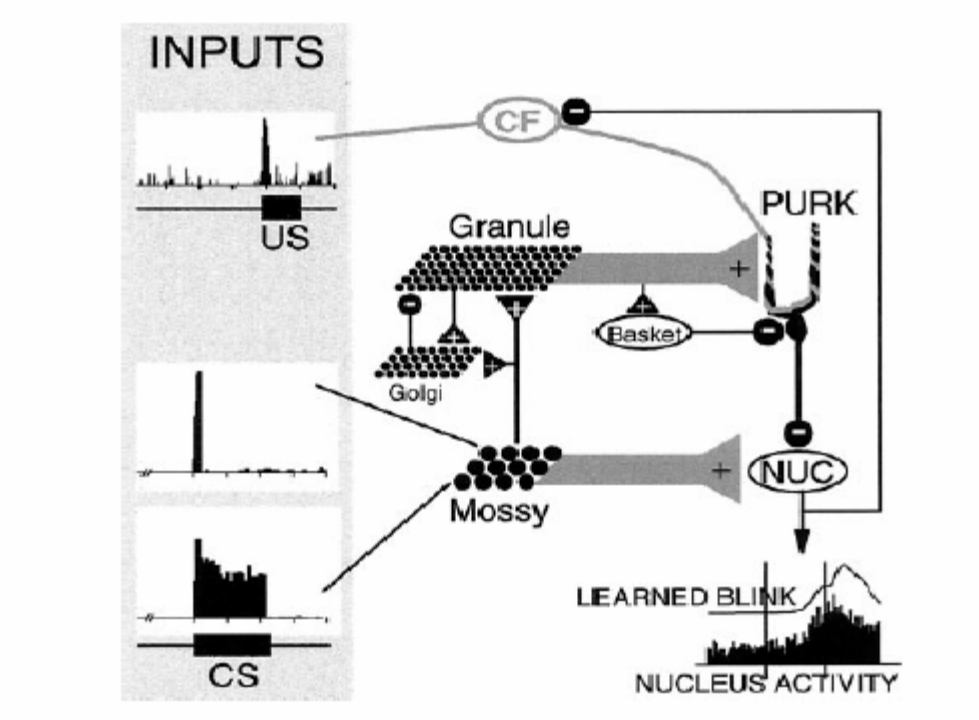

Figure 2.23

Therefore, during the eye lid conditioning, the induction of LTP at mossy fiber to nucleus synapses begins after plasticity in the cerebellar cortex have started to produce transient decreases in Purkinje activity during the CS. This necessarily makes the induction of plasticity at simulated mossy fiber to nucleus synapses lag behind plasticity at granular to Purkinje synapses. This interaction between cortex and nucleus can make the induction of plasticity somewhat sequential. In this learning scheme the fast learning module which is dependent on the cerebellar cortex (Purkinje cells) learns directly from input while the slow learning system in the nucleus learns from the input indirectly through the fast process.

We can write a new state space equation for the sequential updating of the fast and slow process, we have:

$$E_N = Zd_N - \overline{Z}_N$$
$$\overline{Z}_N = Z_N^f + Z_N^s$$

Output equation

2.13

$$Z_{N+1}^f = Z_N^f + K_f(Z_N^s - Z_N^f) + DE_N$$
$$Z_{N+1}^s = CZ_N^s + K_s(Z_N^f - Z_N^s)$$

Update equation

This is the multi-timecourse model with sequential updating. Here the slow process learns from the fast system with learning rate being K_s and it has a retention capacity C. The fast process learns with the learning rate D. Instead of forgetting toward zero we assumed that the fast system forgets toward the slow process with rate K_f.

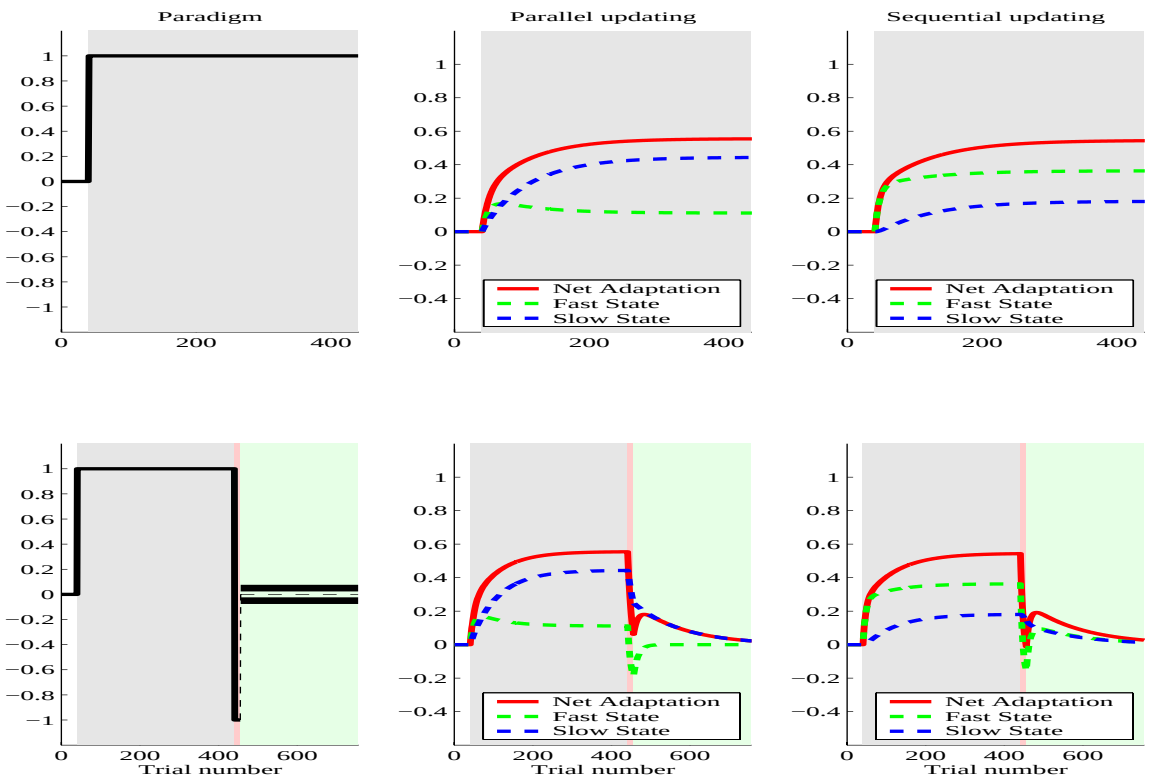

Figure 2.24

Figure 2.24 shows the step response of this version of the multi-timecourse model as well as the parallel version discussed before. The sequential updating can also predict the transient recovery of the old memories observed in our error-clamp experiment as shown in Figure 2.24. One can see that although both model versions have very similar behaviors their internal states are evolving quite differently from each other. This difference is most evident in comparing the relative levels of the fast and slow processes in both updating versions. Yet it seems that this internal difference has not resulted in a measurable net adaptation difference in performance for the paradigms shown in figure 2.24.

Shown below is the experimental data together with the sequential updating model as fitted to the data:

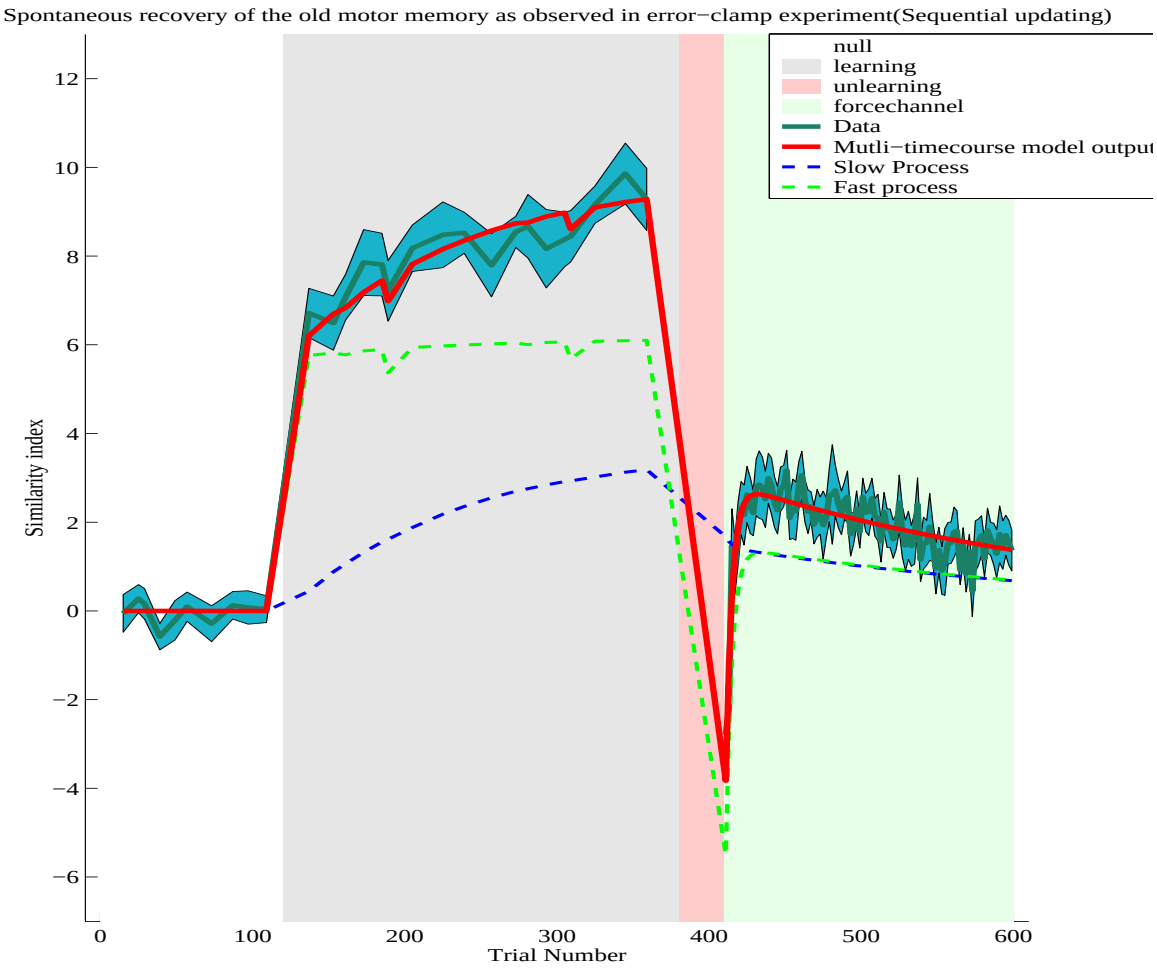

Figure 2.25

The fitted values for the sequential model are as follows:

$K_f = 0.375 \quad D = 0.295 \quad C = 0.991 \quad K_s = 0.012 \quad R^2 = 0.97$

We can see that the quality of the fit using the multi-timecourse model with sequential updating is as good as the fit with the parallel version of the model.

One can easily show that this model can also explain the other motor memory behaviors explained by the parallel version including savings, anterograde interference, rapid unlearning and rapid down scaling. So the question is which of the two versions of the model are closer to the reality of what happens in the brain? Do we have two learning processes that learn independently from errors or is it the case that the learning of the errors propagates sequentially between the two?

Because these two representations can have very similar input-output behavior, behavioral experiments alone in animals or people with normally functioning motor learning systems may not be able to distinguish between them. However the combination of behavioral experiments with neurophysiology and lesion studies may be able to extract the neural architecture of this multi-rate system, by assigning the properties of partially functioning systems and accessing intermediate internal representations.

2.7 Summary

We hypothesized a motor adaptation model with two learning processes with distinct time courses of learning, and showed how this model can be used to explain various puzzling phenomena observed in motor skill learning so far. The successes of this model open a new direction for investigating the properties of the fast and slow processes in motor learning. There are a series of interesting questions one might want to explore in relation to the multi timecourse model:

1. Apart from the different time constants of trial to trial learning, how the contribution of the fast vs. slow system may be different to the total motor output during a trial?

2. How does each of these two processes generalize across arm configurations as well as between arms?

3. How can we relate these two learning systems to regions of brain involved on learning?

4. How can we expand the present multi time-course model to account for formation of multiple motor memories which can act as priors to facilitate or inhibit new learning?

We will try to look at the generalization patterns of the two processes between two arms in the next chapter. A complete treatment of each of these questions can provide guidelines for future research.

Chapter III

Interlimb Generalization of Learning

In this chapter we are going to study how the generalization of the fast and slow processes might be different in the interlimb transfer of learning.

3.1 Interlimb transfer paradigm: Generalization properties of fast and slow processes?

The motivation to study this topic came form an experimental report by Malfait and Ostry 2004 in which the authors claimed that with long gradual introduction of the curl force field there was no significant transfer of learning from trained to untrained hand while adaptation to a sudden introduction of load for a short time with one hand shows *50%* generalization to the untrained hand if the trained hand was dominant. They explained this data by proposing that the cognitive factor of seeing large errors makes subjects aware of the fact that they are actively adapting to a new environment and that is this cognitive awareness which indeed is causing an inter-limb transfer of learning to happen. They then argued that for the same reason in the case of gradual increase of the field strength, subjects are mostly unaware of the small trial-to-trial adjustments they make to maintain a straight and smooth trajectory to the target because they do not observe noticeable and consistent visual errors.

We thought that maybe our multi-timecourse model of learning can provide a better and less presumptive explanation for the transfer experiment by Malfait et al. Figure 3.1 shows the adaptation of the multi-timecourse model[1] to a step change in input which is the same as a sudden turning on of the force field. One can see as a response to the abrupt change in the input the fast system changes quickly to compensate for the errors while the slow system is lagging behind and slowly building. Therefore by the end of a short training period the fast system is carrying 68% of the total learning while the slow process is only contributing to 32%. On the other hand for the gradual input the slow process has enough time to catch up with the input. So for the same level of total performance this time the slow system carries up to 61% of the learning which is considerably higher than the sudden input turn on. The vertical black line in the lower left plot of figure 3.1 shows the trial number in which the total performance of the system given the gradual input is equal to the total performance given the sudden input on the right plot.

[1] Here we assume a multi-timecourse model with parallel updating. The sequential updating in this case cannot make the same predictions as the parallel updating. (See figure 2.24 for relative levels of the fast and slow systems in both updating regimes)

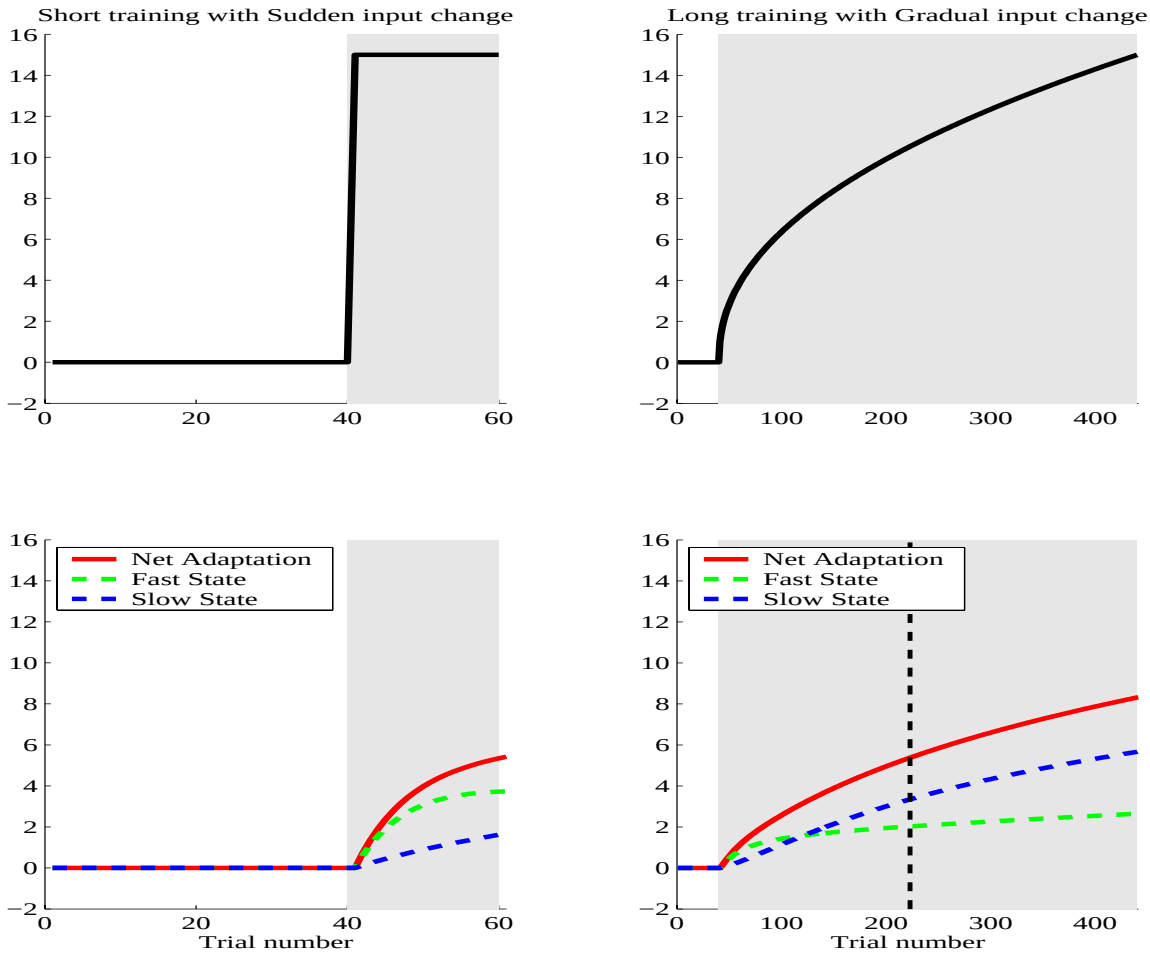

Figure 1.3

But could this difference in the contribution of the fast vs. slow system to the total amount of learning explain the different generalization pattern seen for the abrupt vs. gradual input change seen by Malfait et al.?

We can explain Malfait et al. data if we assume that the slow system has a poor generalization between arms while the fast process generalizes well across limbs. This way one can easily see that since during a short sudden training most of the learning is

being carried by the fast system naturally the transfer of this learning will be much better than the transfer with a gradual input training where the slow process is mostly responsible for learning.

One of the interesting properties of the linear time invariant (LTI) systems like the multi-timecourse models is that regardless of the history of the input, the final steady state of the system and its processes are only dependent on the final input steady state value. This means that no matter if the input is applied gradually or suddenly as long as it maintains a steady state level long enough, the steady state values of the fast and slow systems are going to be the same for the same final input levels. Figure 3.2 illustrates this situation.

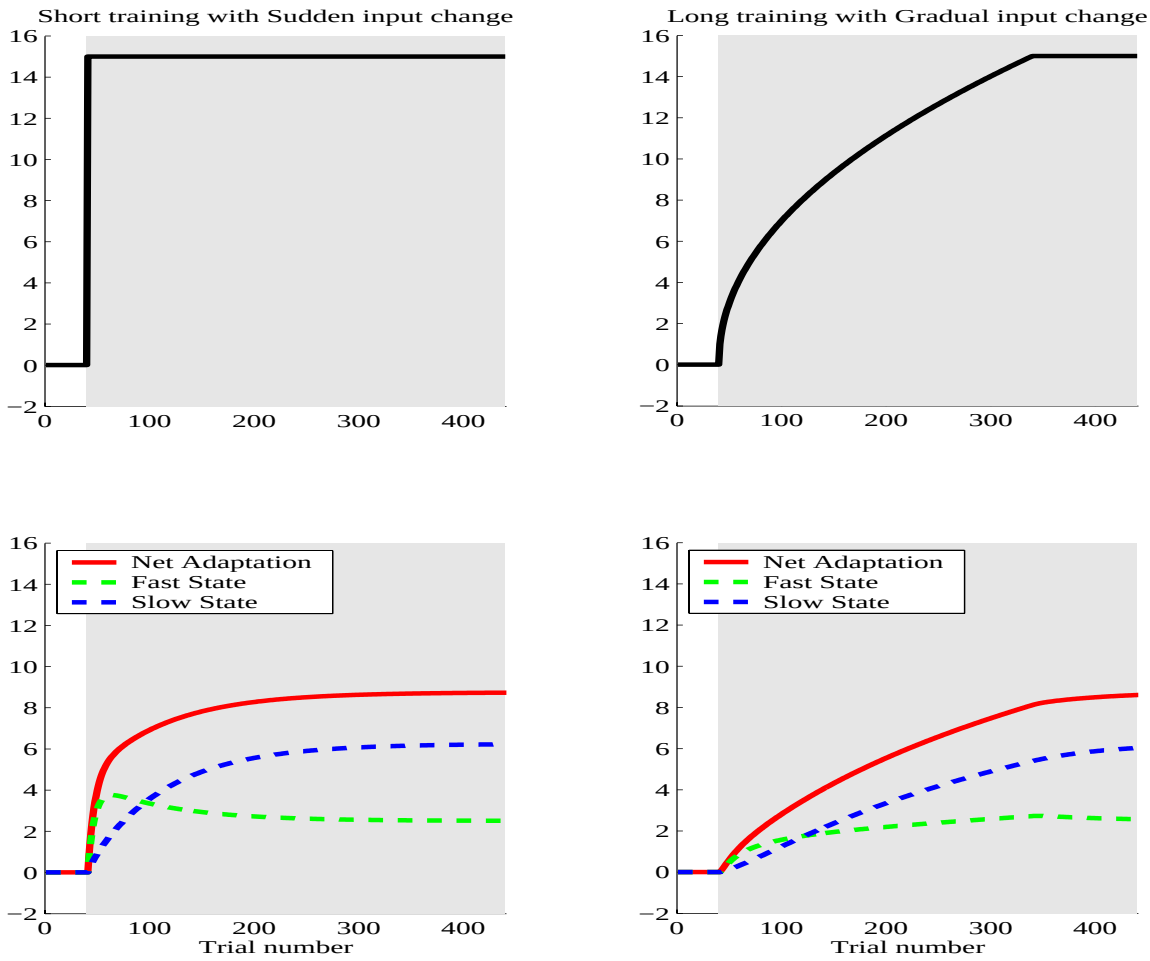

Figure 3.2

This means that if our hypothesis about generalization differences for the fast and slow systems across arms is right we should be able to see transfer results similar to the long gradual input case if we apply a sudden input but train the subjects long enough so that the slow process takes over. In other words our theory makes a somewhat counter intuitive prediction that with short amounts of training in an abrupt field there will be bimanual transfer while with long periods of training this transfer would be diminished.

We can use this phenomenon to design a simple experimental paradigm to contrast our hypothesis about different generalization properties of the fast and the slow system against Malfait's idea about involvement of cognition in the bimanual transfer of learning. Notice that the cognitive transfer of learning should not depend on whether subjects are trained for a long or a short period as long as the introduction of the load is sudden; subjects observe noticeable errors in the training which will make them aware of their adaptation. This means that according to Malfait et al. we should see the same pattern of transfer for the long training as with the short training for the sudden input change.

3.2 Experimental design

16 healthy (10 female 6 male) right handed subjects participated in our study. The experiment consisted of three training paradigms that we were interested to study, namely: Short training with Sudden force field introduction *(SS)*, Long training with Sudden force field introduction *(LS)* and Long training with gradual force field introduction *(LG)*. We wanted to study how the transfer of learning is different for these three paradigms. To be able to do within subject comparison we decided to do all three input conditions in each subject. In their study Malfait and Ostry compared *SS* with *LG* in each subject and they found a difference in the amount of generalization for each condition, that is *SS* training transfers to the untrained hand while *LG* training does not. In order to be able to draw sensible conclusions comparing our data with Ostry's group we designed the experimental procedures to be exactly like theirs:

Subjects made 12 cm point-to-point reaching movements to 1 cm diameter targets. They were trained to have movement times within ±50 msec of the desired time which was set to be *500ms*. Two sets of targets were defined, one set for each hand (see Fig. 3.3). Both arms had roughly the same configuration; initial elbow angles were set at 90°, and shoulder angles were 50°. For each arm, target 1 corresponded to a movement away from the body, and target 2 corresponded to a movement toward the body. The robot produced a force field in which the force f was a function of the velocity of the hand v; specifically $f = Bv$, where $B = \begin{bmatrix} 0; \alpha \\ -\alpha; 0 \end{bmatrix} N \cdot \sec \cdot m^{-1}$, with $0 < \alpha \leq 15$.

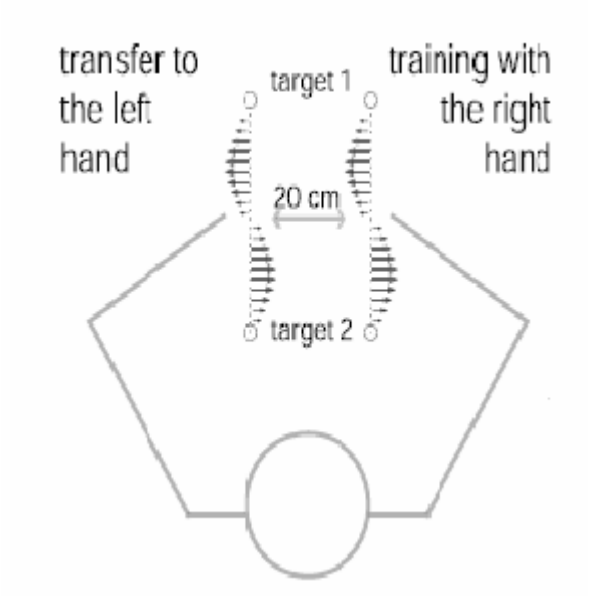

Figure 3.3

At the end of each movement, subjects moved back toward the origin in a force channel which was supposed not to have any considerable effect on the motor adaptation (In Malfait's experiment robot brings the hand back). Subjects were trained to move in the

force-field environment with their right hand and were tested for transfer of adaptation to their left hand.

Before training with the right hand all the subjects did a familiarization set of 30 trials and then a single set of 60 movements in null with their left hands. Three pseudo randomly selected "force-field catch trials" during which the motors of the robot were turned on, were introduced in this set to evaluate the effect of the force field before any learning.

After these initial movements with the left hand the following two training conditions were defined: a "sudden training," in which the load was introduced suddenly, and a "gradual training," in which it was gradually introduced. The sudden training itself had two subcategories.

The short sudden paradigm (*SS)*, consisted of a familiarization set of 30 trials and a single set of 60 trials (60 out and 60 back in force channel). The long sudden paradigm (*LS)* consisted of a familiarization set, two sets of 60 trials and a set of 85 trials. The gradual training paradigm (*LG)* had the same set structure as *LS,* one familiarization set, two set of 60 trials and a set of 85 trials.

In the sudden-training condition, after 45 movements were performed with the motors of the robot turned off ("null field"), the force field was unexpectedly and abruptly turned on; that is, the value of α flipped from 0 to 15 $N \cdot \sec \cdot m^{-1}$ between the 45th and the 46th trial and remained at this value for the final 15 trials (from 46^{th} trial to the 60^{th} trial). In

contrast, in the gradual-training condition, the force field was gradually increased; the value of α changed smoothly from 0 to 15 $N \cdot \sec \cdot m^{-1}$ over 145 trials after 45 movements in null. Specifically, the change in α was nonlinear: $\alpha = n^x$, with n = trial number and x = log(15)/log(145), to obtain $\alpha = 145^{\log(145)/\log(15)} = 15 \ \ N \cdot \sec \cdot m^{-1}$ on the 145th trial. As in the abrupt-training condition, the amplitude of the field remained constant for the final 15 trials. Figure 3.4 depicts the experimental paradigms and the set designs as explained above.

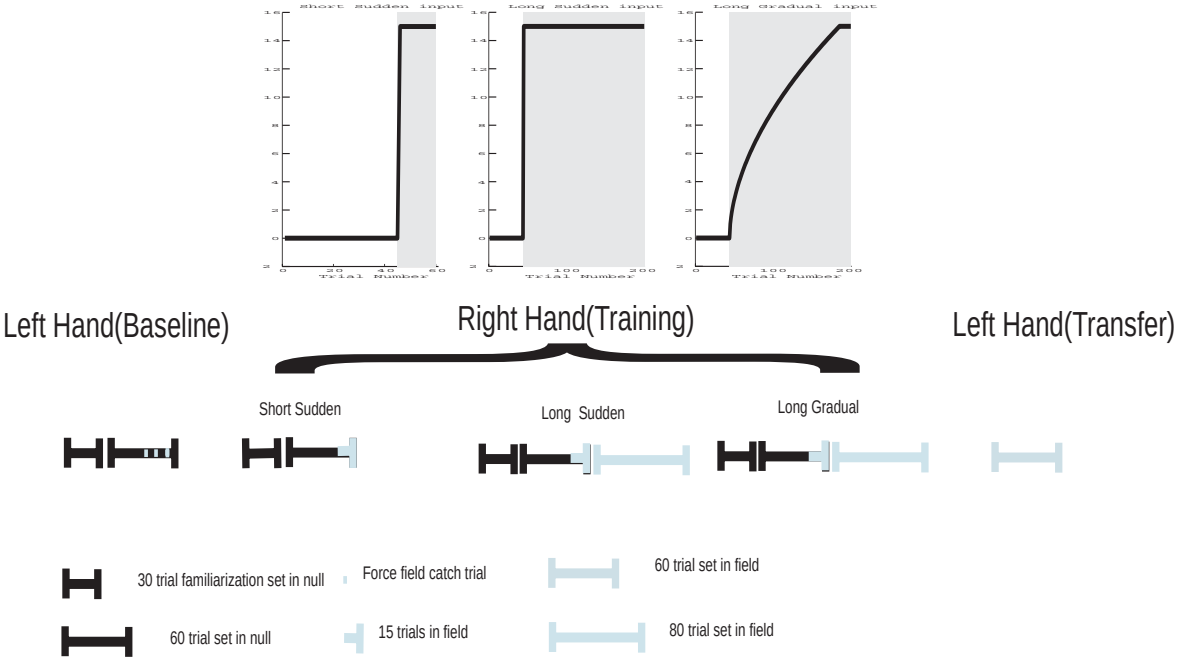

Figure 3.4

It should be noted that subjects performed an unequal number of trials in the three training conditions (160 trials in field for *LS* and *LG* and 15 trials in field for *SS*). Subjects received more training in the *LS* and *LG* conditions than in the *SS*. These differences can only favor transfer of learning for the *LS* and *LG* condition.

Figure 3.5 shows the experimental results obtained by the Ostry's group. In their experiment they only studied the differences in the transfer of *LG* vs. *SS.* To be able to do within subject comparisons, each subject first moved to one of the two targets in one of the two paradigms (i.e. *LG* or *SS)* then after finishing the first paradigm, moved to the other target 180° away in the second paradigm. The order of targets and paradigms were randomized into 4 groups as shown in table 3.1.

Table 1. Experimental design

	First training		Second Training	
Group	Condition	Target	Condition	Target
1	Abrupt	1	Gradual	2
2	Gradual	2	Abrupt	1
3	Abrupt	2	Gradual	1
4	Gradual	1	Abrupt	2

Table 3.1

As shown the *Short Sudden (SS)* training shows a significant reduction of the errors compared to the naïve group as measured by the angular difference from the straight line. The *Long Gradual (LG)* training shows no transfer and its performance seems to fall on top of the performance of the naïve subjects who were doing the task for the first time with their left hand.

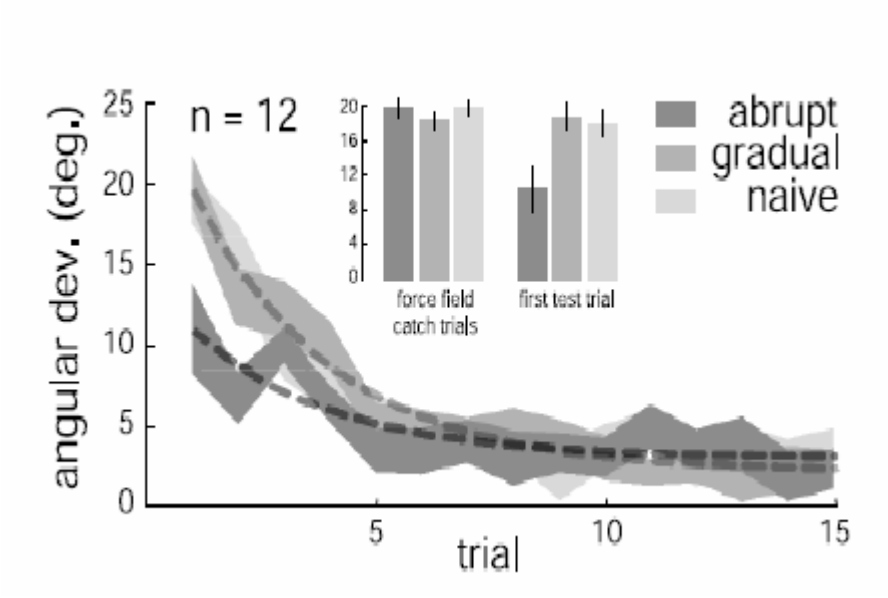

Figure 3.5

Comparing the performance of the first transfer trial with subjects' force field catch trials in the *Short Sudden (SS)* paradigm also shows some *50%* improvement in the angular deviation. This effect is absent for the *Long Gradual (LG)* paradigm.

In order to test our hypothesis we needed to compare the performance of the *Short Sudden (SS)* and *Long Gradual (LG)* paradigm with *Long Sudden (LS)* paradigm. We asked each subject to do two experiments. In one of the experiments we had a comparison of *LS* vs. *SS* transfer properties while in the other experiment we compared the transfer properties of *LS* and *LG*. Each experiment took approximately 1 hour. To randomize all the orderings in the experiment we needed to have a total of 32 groups of subjects which seemed impractical. Instead we decided to have 8 groups of two subjects each. Table 3.2 shows our group trainings. In each experiment training #1 was done in a clockwise field while training #2 was done in a counter clockwise field.

Table 2. Experimental design

Group	Experiment # 1				Experiment #2			
	Training #1		Training #2		Training #1		Training #2	
	Condition	target	Condition	target	Condition	target	Condition	target
1	*LG*	1	*LS*	2	*LS*	1	*SS*	2
2	*LS*	1	*SS*	2	*LG*	1	*LS*	2
3	*LS*	2	*LG*	1	*SS*	2	*LS*	1
4	*SS*	2	*LS*	1	*LS*	2	*LG*	1
5	*LG*	2	*LS*	1	*LS*	2	*SS*	1
6	*LS*	2	*SS*	1	*LG*	2	*LS*	1
7	*LS*	1	*LG*	2	*SS*	1	*LS*	2
8	*SS*	1	*LS*	2	*LS*	1	*LG*	2

Table 3.2

3.3 Experimental results

Figure 3.6 shows our experimental results for all groups of subjects combined using angular deviation from the straight line to quantify learning. As can be seen all our training paradigms result in similar performance indexes in the transfer trials. As shown below we did not see any significant transfer in any of the three training paradigms, comparing the first transfer trial with the force field catch trial. This data seems really puzzling when contrasted to what Ostry's group has reported which showed a clear and significant transfer up to *50%* of learning in the *Short Sudden (SS)* paradigm. The difference between our results and Ostry's group becomes more striking when one bears in mind that our experimental procedures were designed to exactly match Ostry's group.

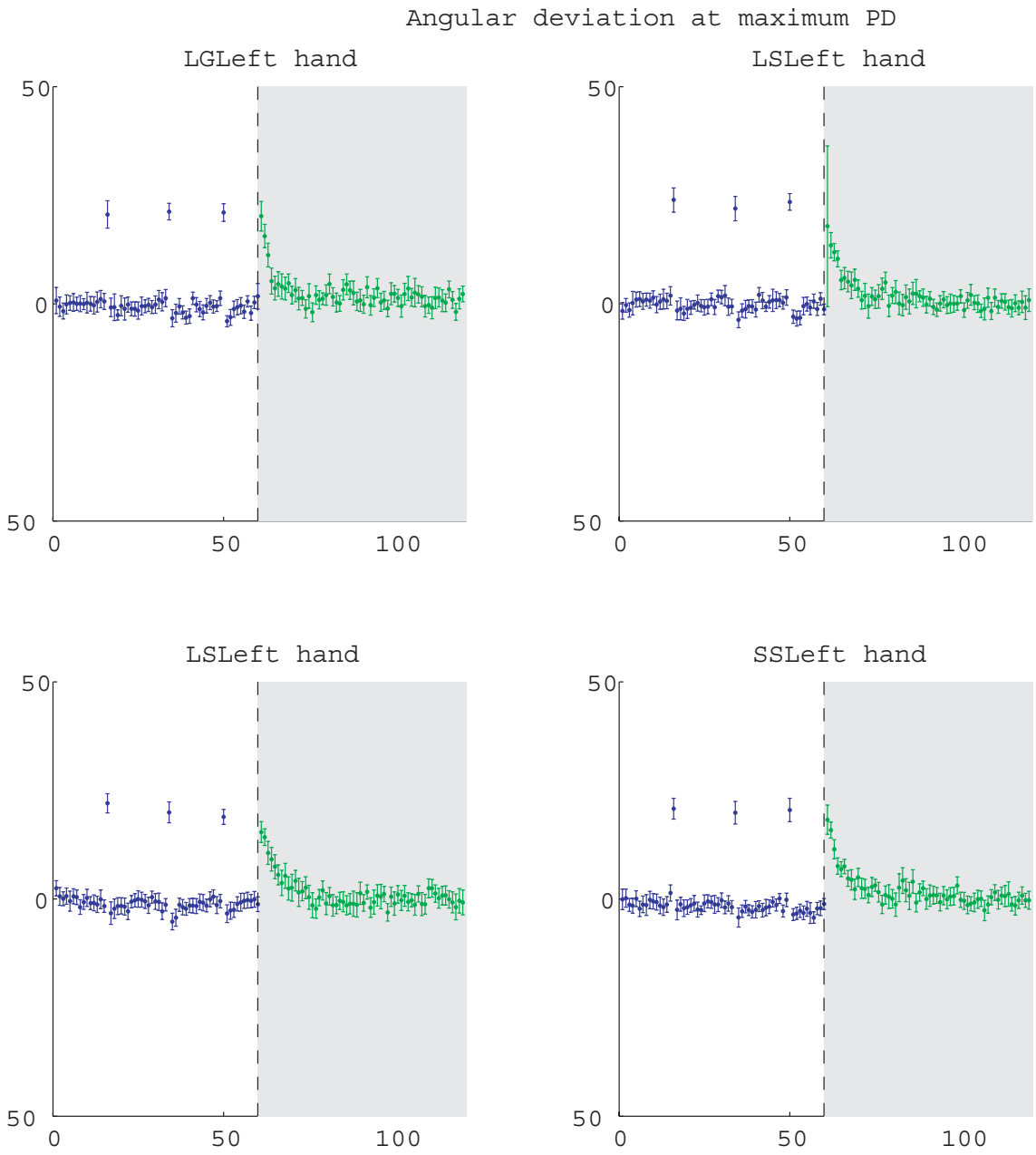
Angular deviation at maximum PD
LGLeft hand
LSLeft hand
LSLeft hand
SSLeft hand
50
0
50
0
50
100

Figure 3.6

Figure 3.7 shows that learning is taking place normally during training with the right hand.

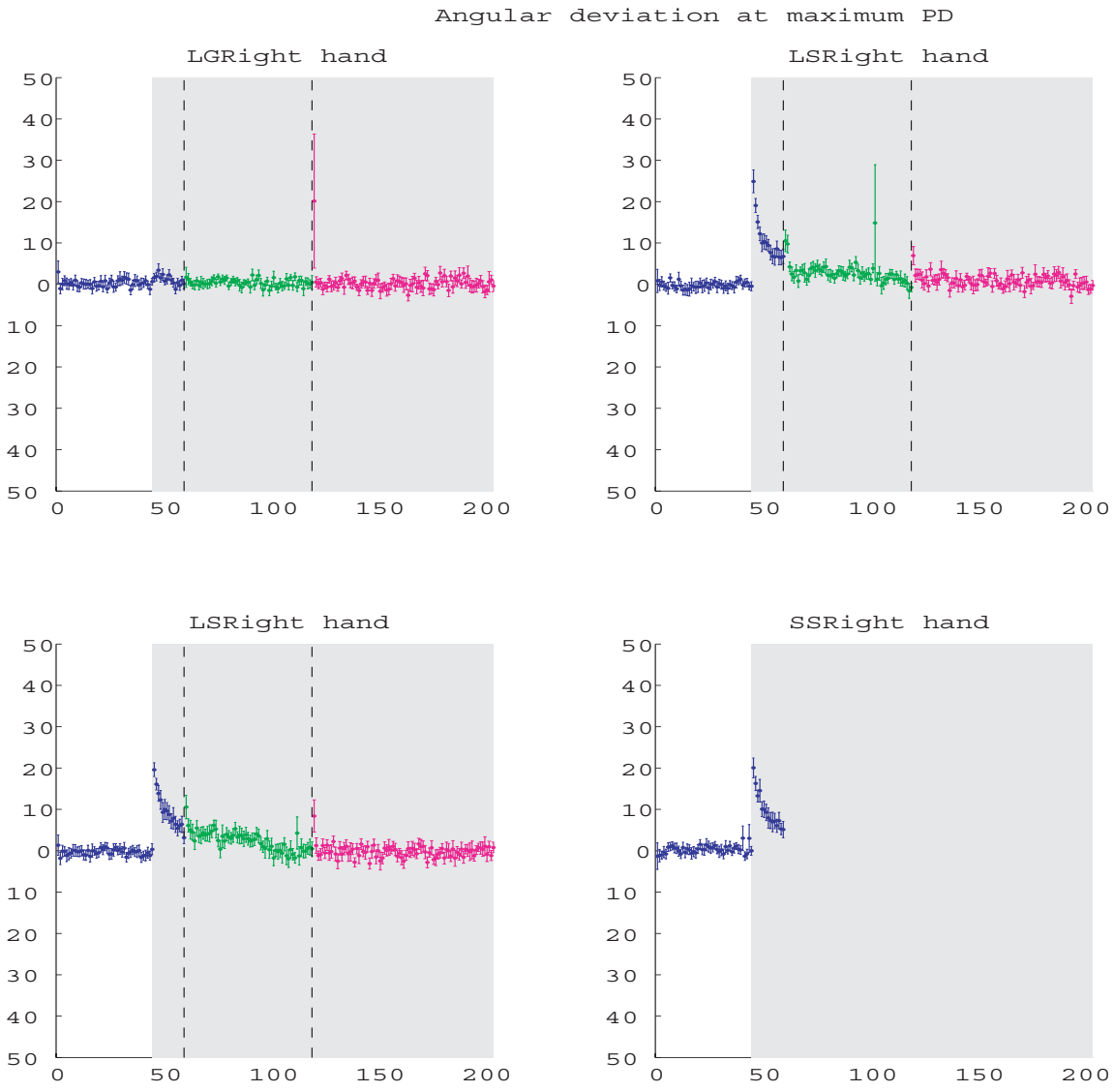

Figure 3.7

To have a more direct comparison of the transfer of the learned force field to the left hand with naïve condition, we further conducted a control experiment using 4 healthy individuals. The control group always moved in the null condition with the right hand therefore their performance on the transfer trials with the left hand can be considered as naive movements in the force field. Since there is no force field applied to the right hand, the control subjects had only two different training conditions: a short training condition during which subject did a familiarization set of 30 trials and then a single set of 60 movements in null with their right hand and a long training condition in which subjects did a familiarization set of 30 trials, two sets of 60 trials and a set of 85 trials. The set design for the left hand was the same as the main experiment. We divided our four subjects into two groups each consisting of two subjects. First group received short training in one direction first and the long training in the other direction while the other group had first the long training and then the short training. Table 3.3 shows the experimental conditions for both groups.

Table3 Experimental design: Control groups

	Experiment = 1			
Group	Training =1		Training =2	
	Condition	target	Condition	target
1	Long	1	Short	2
2	Short	2	Long	1

Table 3.3

Shown below are the results of the transfer trials with the left hand for both groups of control subjects combined.

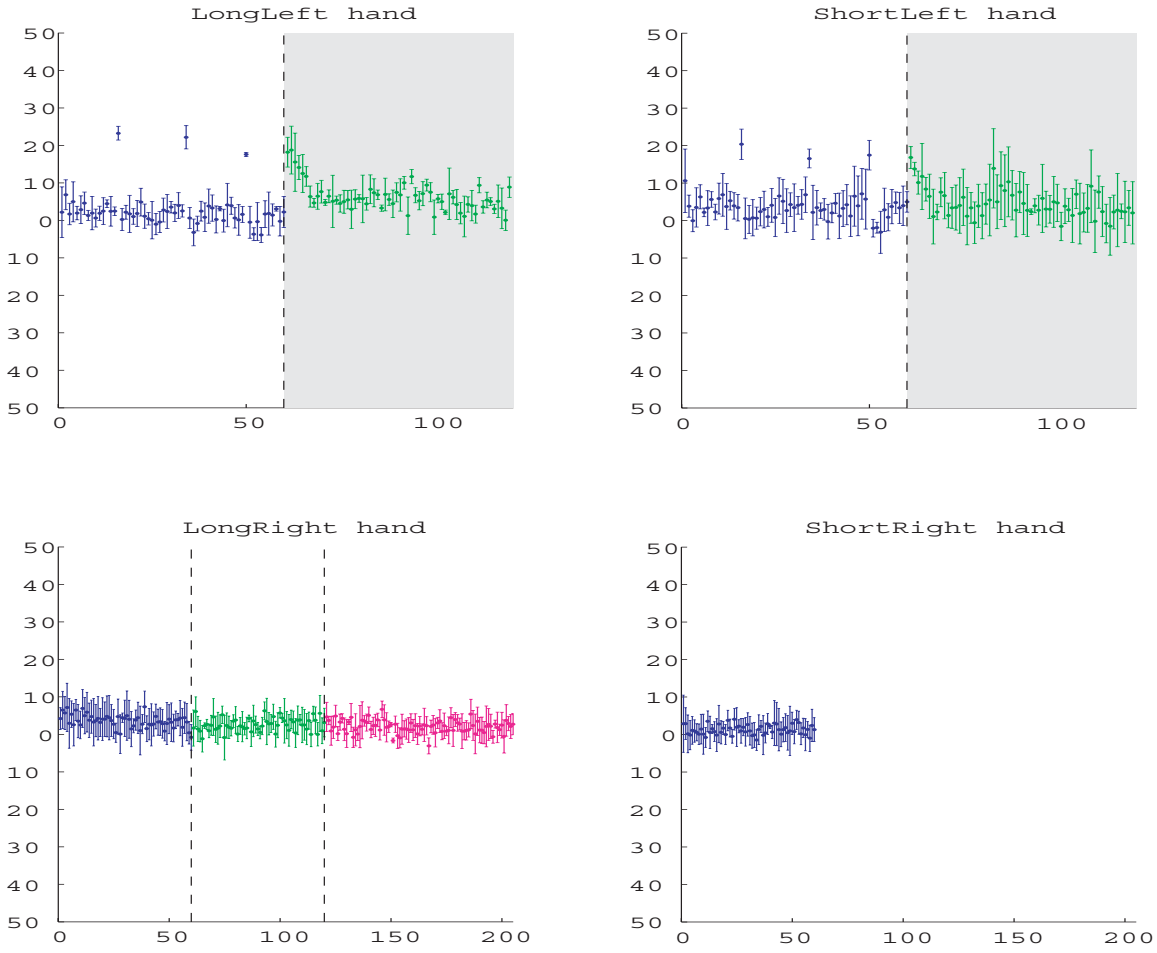

Figure 3.8

If we superimpose the performance of the control groups with our main experimental groups, one can see that there is hardly any evidence that the performance of the left hand in the field might be affected by training of the right hand since the error sequence for the control group seems to overlap with the data from our main experimental paradigm. Notice that we have used the Long control trials as the control for *LS* and *LG* conditions and the Short control trials as the control for the *SS* condition.

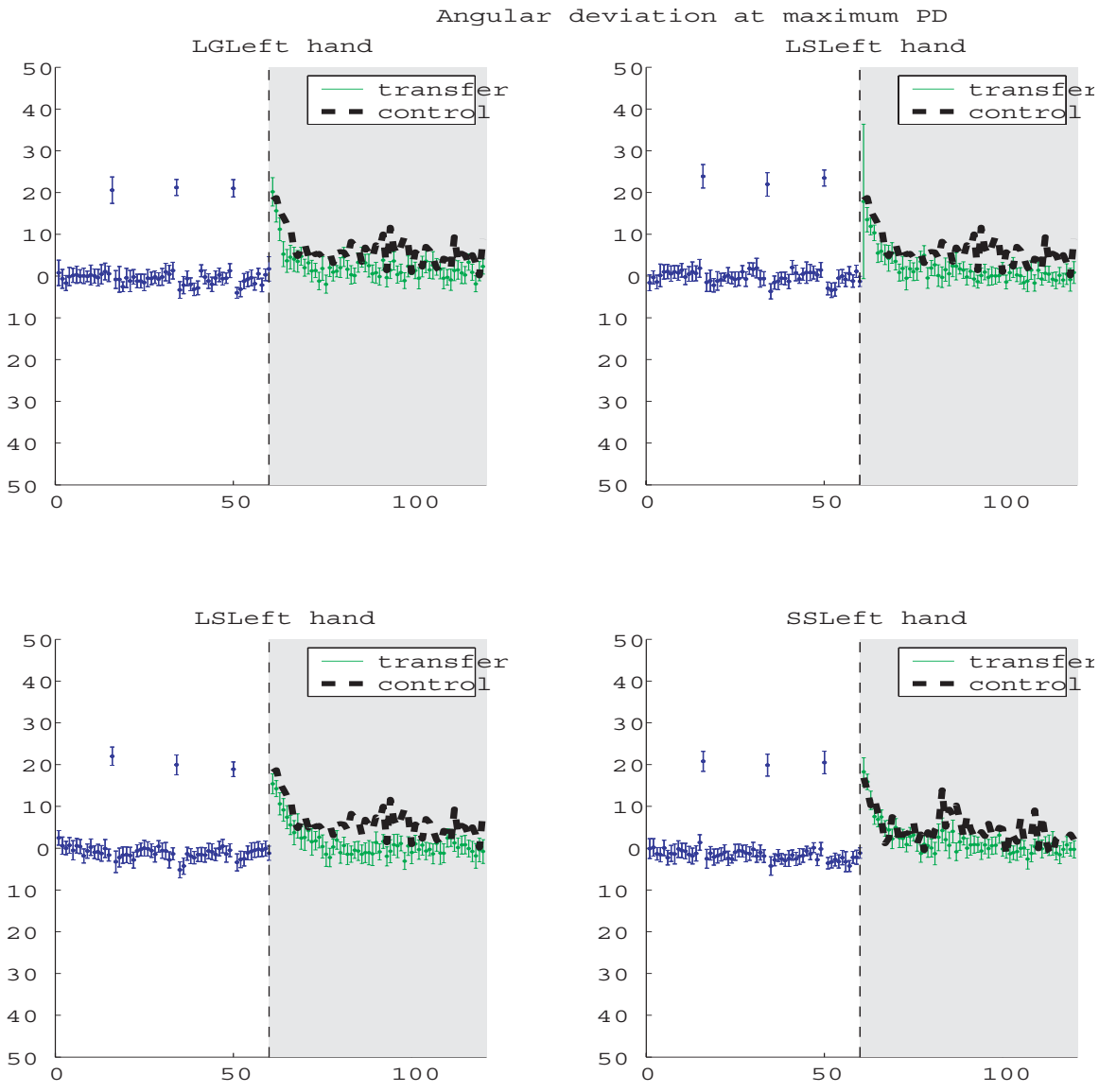
Angular deviation at maximum PD
LGLeft hand
LSLeft hand
LSLeft hand
SSLeft hand
transfer
control
50
40
30
20
10
0
10
20
30
40
50
0
50
100

Figure 3.9

3.4 Summary

Our goal in this chapter was to provide an alternative explanation for the results of the interlimb transfer experiment done by Malfait and Ostry based on our Multi-timecourse model. We hypothesized that by assuming considerable bimanual transfer for the fast process but not for the slow process, one can easily explain the advantage of the short sudden training over the long gradual training in transfer of learning to the untrained arm. This explanation can be contrasted with the Ostry's theory which emphasizes the role of cognitive awareness in the amount of transfer to the naïve arm, using the experimental paradigm described in this chapter. Unfortunately we were not able to see any transfer to the untrained arm in any of our conditions. The source of this failure to reproduce the transfer results is not evident for us at the moment especially since we tried to copy our experimental paradigm exactly the same way as it was designed by Ostry's group.

Although we have not been able to show any difference in bimanual generalization properties of the fast and slow system, the questions raised in the previous chapter related to the differences in the neural substrates and output properties of these two processes as well as the possible existence of other motor memory systems that work on longer time scales can provide us with new directions for future research in motor learning.