Motor learning during reaching movements: model acquisition and recalibration

Sebastian Johannes Telgen

University College London (UCL)

PhD supervisor: Jörn Diedrichsen

Acknowledgements

First and foremost I want to thank my supervisor Joern Diedrichsen for his trust and support throughout my PhD. Joern has helped me to turn my intuitions into something more profound. His intellectual rigor and his attitude that almost anything can be done if one tries hard enough have shaped my thinking and I consider myself very lucky to have been taught by such an exceptional scientist and teacher.

My thanks go to the European Union for funding me for three years within a Marie Curie training program. I also want to thank my secondary Supervisor Brian Day for his advice and help throughout these years. I did not ask for his help very often, but when I did his input was very valuable.

Other colleagues who have supported me with their invaluable feedback on the work presented here are Maurice Smith, Jon Krakauer and Opher Donchin. Reza Shadmehr was one of the reviewers for Chapter 2 which is accepted for publication in the journal of Neuroscience. His comments significantly improved the manuscript.

I want to thank my colleagues Naveed, George, Katja, Alexandra, Sheena and Peter for creating this very energetic and fun environment that I worked in for the past three years. There are many more people at Queen Square who made working there a pleasure, among them Ben, Alessandro, Anna and the fantastic IT-support by Martin and Lambert.

Most of all my colleagues I want to thank Darius Parvin, an MSc student whom I supervised and who helped to collect data for every chapter in this thesis, and kept doing so long after he had his degree. It was a pleasure to work with him.

I also want to thank my close friends Eric, Janis, Simon and my brothers Alex and Marius for reminding me that there is life outside of the lab. I am indebted to my parents, Johannes and Caroline for always believing in me and for supporting me since the day I was born.

Above all I want to thank my partner Vera, who came to London on a one-way ticket without money, income or a plan, but full of love and confidence. You have made these past five years the best years of my life.

Declaration

I, Sebastian Johannes Telgen, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Date:

Contributions

I would like to clarify the contributions of my collaborators to the presented work.

General

Joern Diedrichsen supervised the work presented in this thesis. Darius Parvin helped with data collection in all chapters.

Chapter 4

The algorithms presented were jointly developed by Joern Diedrichsen, Darius Parvin and me. Darius Parvin, a M.Sc. student whom I supervised together with Joern Diedrichsen, presented some of the data (13 out of 16 participants) and analysis methods in his M.Sc. thesis. Therefore, the Methods section of Chapter 4 partly overlaps with the Methods section in the M.Sc. thesis of Darius Parvin.

Abstract

This thesis marks a departure from the traditional task-based distinction between sensorimotor adaptation and skill learning by focusing on the mechanisms that underlie adaptation and skill learning. I argue that adaptation is a recalibration of an existing control policy, whereas skill learning is the acquisition and subsequent automatization of a new control policy. A behavioral criterion to distinguish the two mechanisms is offered.

The first empirical chapter contrasts learning in visuomotor rotations of 40° with learning left-right reversals during reaching movements. During left-right reversals, speed-accuracy trade-offs increased and offline gains emerged, whereas during visual rotations, speed-accuracy trade-offs remained constant and instead of offline gains, there was offline forgetting. I argue that these dissociations reflect differences in the underlying learning mechanisms: acquisition and recalibration.

The second empirical chapter tests whether the dissociation based on time-accuracy trade-offs reveals a general property of recalibration or whether instead the interpretation is limited to the specific contrast between left-right reversals and visuomotor rotations. When the size of the prediction error– the difference between intended and perceived movement – was gradually increased participants switched from recalibration to control policy acquisition. This switching point can be derived by considering the role of internal models in recalibration: If the internal model that learns from errors and the environment are too dissimilar – e.g. in left-right reversal and large rotations– recalibration would cause the system to learn from errors in the wrong way, such that prediction errors would increase further.

To address this problem the final empirical chapter explores if the way the system learns from errors can be reversed.

In conclusion, the results provide behavioral criteria to differentiate between adaptation and skill learning. By exploring the boundaries of recalibration this thesis contributes to a more principled understanding of the mechanisms involved in adaptation and skill learning.

Contents

Figures

Tables

Symbols

A	retention rate of the state space model
B	learning rate of the state space model
CP	change point
e	prediction error of the motor system
ε_n	motor noise
m	mirror reversal factor
$P_{Modelbased}$	probability of a movement to originate from the state space model distribution
$P_{Outlier}$	probability of a movement to originate from the outlier distribution
$\sigma^2_{Modelbased}$	variance of the state space model output
$\sigma^2_{Outlier}$	variance of the outlier distribution
u	cursor displacement
$\dot{x}$	actual motor command
$\dot{x}$	future motor command
y	perceived visual hand location
$\dot{y}$	planned visual hand location
$\hat{y}$	visual hand location predicted by the motor system
$\hat{Y}$	hand location predicted by the state space model
z	state of the motor system

Chapter 1
Introduction

1.1 Adaptation and Skill learning

A search on Google Scholar for the terms “motor” and “skill learning” returns 32.000 results, 2710 of which were published in 2013 alone. While the frequent usage of these terms might suggest a mature scientific discipline, there is no commonly agreed on definition of what motor skill learning means (Shmuelof et al., 2012); rather it is often used as an umbrella term to denote a variety of motor learning tasks, such as finger sequence learning, finger chording, but also much more complex tasks such as dancing (Calvo-Merino et al., 2005). Skill learning is often contrasted with adaptation, the adjustment of movements to sensory prediction errors. Adaptation was first described by Helmholtz who noted that when the visual field was laterally displaced through prism goggles, e.g. to the left, goal-directed reaching movements towards a visual target would miss the target by reaching too far to the left (Helmholtz, 1866). After relatively few reaching movements however,

the nervous system had adapted to the shift such that the hand would accurately find the target. The compensation of the visual displacement might have been explained by aiming a bit further to the right of the visual target. However, Helmholtz noted that after the goggles had been removed, the hand would miss the visual target to the right. Such after-effects are the proof of concept for the idea that plastic changes instead of mere strategic compensation underlie the observed behavioral change.

Tasks that are typically used to study adaptation involve comparably simple movements such as saccadic eye, finger tracking and reaching movements. In contrast, tasks that are invoked to study skill learning in humans are more complex. They often also differ with respect to the effectors involved. Examples include the learning of sequential finger movements (Lotze et al., 2003; Meister et al., 2005), finger configurations (Waters-Metenier et al., 2014) and even learning to swing a golf club (Beilock et al., 2008).

Given these differences at the task level, do the terms *skill learning* and *adaptation* only refer to differences in task complexity and the effectors that are involved? If this was true, the distinction between adaptation and skill learning would be purely describing the nature of the task instead of distinguishing them on the basis of the underlying type of learning. Such a task-based definition is problematic because many tasks might be learned through a combination of different processes. Thus instead of classifying motor learning tasks as adaptation or skill learning, it is more promising to study the mechanisms involved and why they are involved.

In this thesis I ask whether it is possible to distinguish between different learning mechanisms based on behavioral criteria – and if so – how these criteria relate to the behaviors typically observed in adaptation and skill learning tasks.

To address these questions participants were instructed to perform center-out reaching movements using a robotic handle. Instead of the real hand location, they saw a cursor and the task was to move the cursor towards a target presented on the screen. The learning task consisted either of learning to compensate for a rotation of the cursor around the movement origin (visual rotation) or a left-right reversal across a mid-sagittal axis (mirror-reversal).

From the results it will become apparent that a) under certain circumstances visually rotated and mirror reversed feedback are learned by different mechanisms and b) that these mechanisms carry the behavioral signatures of learning as it is often observed in adaptation and skill learning tasks respectively. Moreover I will show that there are inherent limitations to the dominant mechanism in adaptation and that at its boundaries another mechanism (presumably the one that dominates most skill learning tasks) takes over.

In order to lay out my argument, I will first clarify a number of key concepts of motor control and how I use the corresponding terms in this thesis. I will start by describing two recent accounts of skill learning and explain how the current thesis extends them. Next I focus on the acquisition of internal models and control policies, the mappings that define sensorimotor control. Thereafter I will outline how single goal-directed movements are generated and controlled online, based on internal models. Finally I will revisit adaptation and skill learning, but this time from the perspective of internal models and control policies. From this we arrive at why different learning mechanisms must be responsible for skill learning and adaptation

and how we can distinguish between these mechanisms based on the observed behavior. After having focused on the active learning processes themselves I will give a brief account of how motor memories consolidate after training and how the signatures of consolidation can be used as additional hints at the learning process involved. I will conclude by summarizing the main hypotheses that are addressed in this thesis and preface how they are approached in the following chapters.

1.1.1 Definitions of adaptation and skill learning

In recent years the motor learning community has started to move away from a purely task based classification and towards more principled ideas of what adaptation and skill learning are.

For example Costa postulated what he called a "selectionist view" of de novo action learning (Costa, 2011). In this view all possible motor outputs are always available to the nervous system. The learning process then consists of associating the appropriate actions to external stimuli. As a result of training the selection process becomes automatized. As in other types of reward association learning the dopaminergic system plays a critical role by associating rewards with certain actions in response to external stimuli. Indeed patients with disorders connected to Basal Ganglia dysfunction, such as Huntington and Parkinson disease exhibit marked impairments in certain skill learning tasks, such as sequence learning (Boyd et al., 2009). While Costa proposes that action selection underlies skill learning, he neither defines what skill learning is nor does he speculate about the circumstances under which action selection is involved. It remains unclear if action selection

underlies adaptation as well, and if not, why action selection takes place in skill learning but not in adaptation.

A recent perspective on the dissociation between skill learning and adaptation focuses on the output goals of the learning mechanisms (Krakauer and Mazzoni, 2011; Shmuelof et al., 2012). Whereas the goal in adaptation is to return to the baseline performance while faced with a perturbation, the goal of skill learning is to improve performance beyond baseline levels, where baseline levels can also mean that the skill was literally non-existent before practice. In contrast to the action selection account, the output based definition assumes that motor outputs produced as a result of skill learning are truly novel outputs. Fortunately the distinction does not stop at the level of tasks but instead Krakauer and Mazzoni propose that "*skill learning is improvement in a controller through trial-and-error reinforcement*", whereas "*adaptation is updating of an internal model-based on sensory prediction errors"*(Krakauer and Mazzoni, 2011).

The results and their interpretation provided in this thesis largely agree with the intuition on the mechanisms proposed by Krakauer and Mazzoni. However, a distinction based on different learning goals of the underlying processes must be rejected based on the results presented in chapters 2 and 3. Throughout this thesis I argue that different mechanisms of motor learning can be distinguished based on whether the desired output can be derived from an existing control policy (or internal model), which will be referred to as recalibration (or adaptation), or whether a new control policy (or internal model) has to be acquired de novo.

1.2 Learning control: Internal Models and Control Policies

In this thesis the proposed dissociation between skill learning and adaptation builds on the notion of internal models, the mappings between the sensory and motor coordinate frames. Any goal directed movement requires some form of a mapping between the goal and the required motor commands. While some of these mappings might be innate, many are acquired throughout life. As an example for the acquisition of such maps consider the learning of reaching movements in infants. Here the goal is to move the hand to a desired target.

In motor control the rule that dictates a specific motor output in response to a visual input is called a control policy. In artificial intelligence and machine learning an optimal control policy is a rule that chooses an action X to produce an optimal outcome Y. The concept of policies as an element of learning originates from the field of reinforcement learning: "*A policy defines the learning agent's way of behaving at a given time. Roughly speaking, a policy is a mapping from perceived states of the environment to actions to be taken when in those states. It corresponds to what in psychology would be called a set of stimulus-response rules or associations. In some cases the policy may be a simple function or lookup table, whereas in others it may involve extensive computation such as a search process. The policy is the core of a reinforcement learning agent in the sense that it alone is sufficient to determine behavior"* (Sutton and Barto, 1998). Note that the environment can, depending on the definition, include the control plant – here the musculoskeletal system – and that states can include goal states as well.

Control policies can be learned in two ways. They can either be derived from an internal model or alternatively they can be learned directly by observing rewards associated with certain actions and then strengthening those circuits that produced the rewarded actions (Huang et al., 2011; Haith and Krakauer, 2013). The latter is also referred to as model-free learning.

1.2.1 **Model-free learning**

The only way a control policy can be learned directly is through reinforcement learning. In reinforcement learning the observed rewards are associated to the preceding motor commands. The stronger the observed reward, the more the synaptic weights that gave rise to the output are reinforced. In this way the rewarded outputs are more likely to reoccur in future movements. Direct reinforcement learning of control policies thereby establishes direct stimulus response pairs.

The alternative to direct or model-free learning is to establish a mapping that represents how a change in the goal state translates to a required change in the motor apparatus. With such a model the control policy could – at least mathematically – be derived easily.

1.2.2 **Inverse Models**

In control theory mappings from sensory goal states to the required motor commands are called inverse models (Jordan and Rumelhart, 1992; Wolpert and Kawato, 1998). They can be thought of as lookup tables that specify which motor command (x) needs to be issued to move from the current location (y) to the desired

location ($\dot{y}$) (Eq. 1.1). The mapping between a single joint and a single visual dimension can be captured by the following differential equation:

Equation 1.1 $$required\ motor\ command = \frac{\partial x}{\partial y} * (\dot{y} - y)$$

However, because the inverse model links three visual dimensions to a multitude of muscles, tendons, joints and their interactions, the mapping must be written as a Jacobian matrix, with one cell for each possible combination of *k* visual with *j* motor dimensions:

Equation 1.2

$$required\ motor\ command = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_j}{\partial y_1} & \cdots & \frac{\partial x_j}{\partial y_k} \end{pmatrix} * (\dot{y} - y)$$

There are essentially three alternative computational frameworks that explain how inverse models could be constructed. The first possibility is that inverse models are learned directly (Miller, 1987), without the help of an existing model of the environment. This is conceptually similar to learning multiple stimulus response mappings or control policies directly.

Since reinforcement is gradual rather than binary, especially when summed over repeated instances, an inverse model can be represented in the combined population output of several such stimulus-response mappings. However, direct inverse modelling might computationally not be feasible because it ignores a fundamental property of redundant systems: In linear non-redundant subspaces of the mapping, direct inverse modelling might lead to good results. Still, the high degree of redundancy and nonlinearity in the entirety of the mapping means that often the average output of two stimulus-response pairs with neighboring outputs

does not necessarily result in a movement that falls between the two individual outputs. (Kawato, 1995).

1.2.3 Forward models for inverse model acquisition

One alternative to the direct inverse modelling approach proposes that in the first step a forward model is learned by using the visual error as a teaching signal. Forward models predict sensory consequences based on the knowledge of the state of the environment, the plant and crucially, the controller and its commands or actions. For the purpose of learning inverse models for motor control, forward models predict how the visual hand location (y) changes depending on the difference between the current (x) and the future motor command ($\dot{x}$).

Equation 1.3

$$required\ motor\ command = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_j} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_k}{\partial x_1} & \cdots & \frac{\partial y_k}{\partial x_j} \end{pmatrix} * (\dot{x} - x)$$

Consider an infant who prior to producing accurate reaching movements seems clumsy while moving his arms in a way that might seem undirected. During exploratory movements, infants can observe the sensory consequences of different motor commands and thereby build a forward model through supervised learning. Once a forward model has been established, the nervous system can compare the actual sensory outcome to the outcome that was predicted by the forward model and use the error signal as a teaching signal. In a non-redundant linear system the forward model could then be used to derive the corresponding *inverse model* in the second stage. However, since the mapping from motor outputs to sensory inputs is

highly redundant and nonlinear the forward model cannot be inverted. The proposed solution is to back-propagate the error signals through the forward model. In this way the visual error signals can be transformed into motor error signals and used as distal teachers for the acquisition of an inverse model (Jordan and Rumelhart, 1992).

The third approach to inverse model construction is called feedback error learning. It proposes that feedback commands, which are generated online in response to visual error signals, can serve as teacher signals for the inverse model, without the need for a forward model (Kawato and Gomi, 1992; Kawato, 1995). In other words, the nervous system remembers the feedback correction that was issued to cancel out a visual error in a given movement and adds a time-advanced version of this feedback command to the initiation of the next movement.

1.3 Exerting Control: Feedforward and Feedback commands

Feedforward and feedback control have traditionally been used to denote two different phases of a movement and are therefore observed as well as possible in isolation in the empirical chapters.

Feedforward control describes open loop movement initiation. During goal directed reaching, it reflects the translation of visual coordinates into motor commands by a control policy. The word feedforward refers to the fact that during movement initiation, due to inherent delays in sensory feedback, movements are generated in an open-loop manner.

In contrast feedback control refers to control signals that are generated in response to sensory feedback after the movement has been initiated. For example,

the visual hand representation might be perceived to move too far to the right when reaching for a target. Feedback control, as understood in this thesis, then refers to the process that results in the online control signal that causes the hand to move further to the left. Feedforward and feedback control can under certain circumstances be dissociated (Gritsenko and Kalaska, 2010). For example, when reaching along the axis of mirror-reversal, the optimal feedforward command is the same as during normal reaching, while the feedback command needs to be reversed, such that when the visual feedback of the hand position indicates that the hand moves too far to the right, the actual hand must reach even further to the right to cancel out the error. Such feedback corrections are very fast and highly automatic.

Day and Lyon pioneered a paradigm in which participants make goal directed reaching movements where the target position is displaced early after movement onset (Day and Lyon, 2000). They found that feedback corrections towards the target occurred as early as 125-160ms after the target displacement. Moreover, when participants were instructed to move the hand in the opposite direction of the target displacement, the fast feedback response towards the target still occurred at very fast latencies, while the explicitly instructed response in the opposite direction occurred only at longer latencies. Although there is evidence that the fast feedback responses elicited by target and hand representation displacements are not equivalent (Reichenbach et al., 2014), similar findings have been obtained for cursor displacements (Franklin and Wolpert, 2008), when the cursor represents the hand position. Therefore, feedback corrections to cursor displacements offer valuable insights into the state of the motor system with relatively little interference from cognitive processes. In chapter 2 feedback corrections are used as a window into the temporal characteristics of the computation of motor commands.

1.3.1 **Forward models in online control**

The brain continuously adjusts its estimates of the current hand location and corrects its movements accordingly. It does so through the use of sensory feedback as well as the use of efference copies (Holst and Mittelstaedt, 1950; Sperry, 1950; Bell, 1989), but also through predictions from its internal models. The earlier introduced forward model is ideal for predicting sensory consequences after a motor command has been issued. The striking advantage of using predictions derived from internal models is that they allow for optimal state estimation, at minimal temporal delays as opposed to the integration of delayed sensory feedback. In other words the motor system can adjust suboptimal movements online without having to wait for the observation of the sensory consequences of its suboptimal control signals (Miall and Wolpert, 1996; Blakemore et al., 1998). Indeed patients with lesions to the cerebellum, the neural structure that is thought to implement forward models (Wolpert et al., 1998), suffer from endpoint ataxia. The explanation is that normally the motor system would predict when the hand will reach the target and thus issue a stop signal, before the target is reached (Miall and Wolpert, 1996; Desmurget and Grafton, 2000). If the brain cannot predict when the hand will reach the target, it must rely on delayed sensory feedback and therefore the stop signal is always issued too late.

The dissociation between feedforward and feedback control becomes blurred at this point, which until the advent of optimal feedback control (Todorov and Jordan, 2002) were often described as being the output of two distinct controllers (Kawato, 1995; Hay and Redon, 1999). One advantage of using predictions is that instead of two separate internal models, a combined internal model, consisting of a forward and an inverse model or control policy is sufficient

for motor control. In line with the computational equivalence of feedback and feedforward control, empirical evidence suggests that feedforward and feedback control rely at least partially on shared internal models and thus controllers (Wagner and Smith, 2008). Therefore in the current thesis I will only use operational definitions of the terms feedforward and feedback control: Here feedforward control refers to movement onset measured at delays that are too short for sensory feedback from the movement to interfere, whereas feedback control refers to online movement corrections in response to sensory feedback.

The earlier introduced concept of the forward model is not only relevant for the construction of inverse models and online control but it also plays an important role in maintaining accurate internal models for motor control through recalibration.

1.4 Maintaining Control: Sensorimotor Adaptation

Adaptation has been studied extensively since it was first discovered by Helmholtz almost 150 years ago (Helmholtz, 1866) (McLaughlin, 1967; Tseng et al., 2007; Shadmehr et al., 2010). It has been documented in a variety of species and behaviors ranging from walking (Prokop et al., 1995) to the vestibulo-ocular reflex (Shelhamer et al., 1994).

The data presented in the current thesis will exclusively describe reaching movements under manipulated visual feedback in humans. In most reaching movement adaptation tasks participants usually cannot see their hand directly. Instead they are presented with a visual representation of the hand location in the form a cursor shown on a screen. Movements are initialized at some start location

and the goal is to move the cursor to the target. Participants are not shown the veridical hand location, but instead the feedback of the hand position on the screen is perturbed such that moving the cursor to the target requires the motor system to compensate for the perturbation. For example the movement gain can be up or down regulated (Krakauer et al., 2000), the cursor position can be displaced by applying a constant translation (Wei and Körding, 2009) or the cursor trajectory can be rotated around the movement origin (Abeele and Bock, 2001a). The nervous system learns to adjust subsequent movements in such a way that they predictively cancel out the expected external perturbation (Shadmehr and Mussa-Ivaldi, 1994).

What do translations, rotations and gain scalings of the visual hand representation have in common? All of these examples can usually be learned by gradual approximation through a simple first order parametric learning process. After observing that the cursor did not move where it was intended to move, the nervous system will in the subsequent trial compensate for a fraction of the error signal. To learn from error signals in a goal directed fashion, the nervous system must have access to the gradient ($\frac{\partial x}{\partial y}$) that determines how a visual error signal, that is the difference between actual (y) and predicted visual outcome ($\hat{y}$), translates into the required update of subsequent motor commands:

Equation 1.4

$$required\ motor\ update = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_j}{\partial y_1} & \cdots & \frac{\partial x_j}{\partial y_k} \end{pmatrix} * (\hat{y} - y)$$

The knowledge about the underlying gradient constitutes a model in its simplest form and therefore learning along this gradient is by definition model-based (Haith and Krakauer, 2013).

From a computational standpoint error-based learning along a gradient is conceptualized in the form of state space models (Eq. 1.5) (Thoroughman and Shadmehr, 2000). These equation systems are - in spite of their apparent simplicity - remarkably accurate in predicting behavior during adaptation, from the learning phase to the after-effects that persist after the perturbation has been removed (Smith et al., 2006).

Equation 1.5

(i) $$z_{n+1} = A * z_n - B * (y_n - \hat{y})$$

(ii) $$y_{n+1} = z_{n+1} + \varepsilon_{n+1}$$

The hand location (y) in trial *n+1* is determined by the state estimate (z_n) and motor noise (ε). The retention rate (A) determines how strongly the state in trial *n* is retained in trial *n+1*. B is the learning rate that determines how much the prediction error ($y_n - \hat{y}_n$) contributes to the state in trial *n+1*.

Note that I used the term prediction error instead of visual error to emphasize a fundamental property of recalibration: In one recent experiment (Mazzoni and Krakauer, 2006) participants were asked to reach for targets under a counter-clockwise 45° rotation of the visual feedback. Crucially the participants were explicitly debriefed about the nature and size of the rotation and instructed to counter the visual rotation by aiming at a point that was 45° clockwise relative to the target. Initially this strategy proved successful, however over time participants would increasingly miss the target clockwise. Finally performance deteriorated to a point where participants started to actively ignore the instruction to mentally aim for the neighboring point. Instead they successfully reached for the real target

because they had adapted to the cursor rotation. Thus parametric learning continued even when participants were instructed to actively use a reaiming strategy, such that the visual error was zero (Mazzoni and Krakauer, 2006). Since the visual error signal between cursor and target was zero, and therefore could not possibly have driven recalibration, what alternative mechanism might explain the clockwise drift? The most parsimonious explanation is that forward models continuously predict the sensory outcomes of motor commands. It is the mismatch between forward model predictions and visual outcomes that drives parametric learning.

Feedback error learning offers an elegant explanation of how adaptation proceeds. It suggests that just as in the case of the original acquisition of the inverse model, inverse models can be recalibrated by updating the internal model by a proportion of the online feedback correction (Kawato and Gomi, 1992). However, it has been shown that online corrections are not necessary for adaptation (Tseng et al., 2007). For example when saccades are executed towards a target and the target is briefly displaced and then returned to the initial location, such that feedback corrections are unnecessary and not executed, saccades will still adapt over time, even in the absence of feedback corrections (Noto and Robinson, 2001). Thus, taken together with the results from Mazzoni and Krakauer (2006), evidence strongly suggests that recalibration neither depends on feedback corrections nor on reward learning, but instead sensory prediction errors of the forward model drive recalibration.

The cerebellum is considered to be one of the key substrates in this type of learning. In particular its anatomical structure is considered ideal for implementing forward models for motor control (Miall and Wolpert, 1996), and even for higher

cognitive functions such as predictive language processing (Lesage et al., 2012). Moreover a plethora of studies from rodents, over monkeys to man show that lesions in the cerebellum result in the impairment and sometimes even the complete loss of the capacity for sensorimotor adaptation (Martin et al., 1996; Takagi et al., 1998, 2000). It has recently been suggested that while the cerebellum is needed for parametric adaptation, truly long lasting changes are induced in primary motor cortex (Galea et al., 2011). Indeed M1 neurons have repeatedly been shown to alter the center of their cosine tuning curves in response to visually rotated cursor feedback (Paz et al., 2003). Paz et al. found that only those neurons in M1 that were tuned in the direction of the required movement before the onset of visual-rotation learning changed their preferred direction. If this mechanism is required for parametric adaptation, then, given the approximately cosine tuning functions of M1 neurons, it would predict constraints on the maximal rotation error size that can be learned through recalibration. I will address this idea in Chapter 3.

1.4.1 Recalibration and control policy acquisition

The term adaptation is often used as if to imply first-order error-based learning. However as I will argue throughout this thesis, one of the most studied sensorimotor adaptation tasks cannot be learned through recalibration. In particular, when left-right reversing the location of the cursor over a mid-sagittal axis, also termed mirror-reversal, parametric learning using the old internal model leads to catastrophic results. While during normal reaching the internal model would correctly transform a leftward visual error into a rightward motor update, during mirror-reversal the same update would cause the leftward error to increase even further in the following trial. The idea that during mirror reversal, the internal

model learns from its prediction errors in the wrong way will be tested in great detail in chapter 4. If the same internal model generates feedforward commands and parametric state updates, then adaptation cannot be used to learn mirror-reversal. Thus instead of updating an existing internal model and with it an existing or "old" control policy a "new" control policy must be *acquired* to learn mirror-reversal. Note that in the nervous system it is unclear in how far a control policy that is derived from an internal model is equivalent with the internal model itself. Therefore I will use these terms synonymously throughout the remainder of this thesis. I will refer to adaptation as recalibration to stress that an old existing model and control policy are updated and to emphasize the departure from the task based distinction. In contrast, mirror-reversal learning relies on the establishment of a new control policy. But what is the neural analogue of recalibration of an old versus acquisition of a new control policy? The underlying idea here is that motor control is supported by a set of control policies that can be activated depending on the specific context. When an old control policy is recalibrated an existing neural circuit is modified, such that the automaticity of its computations can be inherited. In contrast, the acquisition of a new control policy is initially achieved by additional time intensive computations, which can with continuing practice become automatized. I here propose that the relative automaticity of a motor command can be used to characterize its neural implementation.

1.5 Time-accuracy trade-offs

In this thesis I will argue that the relationship between processing time and accuracy is the key criterion for the dissociation between recalibration and acquisition. One common observation that can be made across a multitude of skill learning tasks is that changes in the relationship between preparation time and accuracy are altered as a result of learning. In particular after training the same accuracy can be achieved at shorter latencies. The study of processing times is as old as the academic discipline of psychophysics itself (Donders, 1969). The concept of a trade-off between processing time and performance attributed to capacity in information processing has first been formalized by Fitts (Fitts, 1954). The underlying assumption is that processing times are proportional to the computational load of the nervous system. If the latency at which the nervous system generates an output is shorter than the required latency for producing an optimal output, then the output produced at the suboptimal latency will be suboptimal as well.

In this thesis I will use the terms speed-accuracy trade-off and processing time-accuracy trade-off synonymously for the following reason: Under the assumption that optimal feedback control theory is correct, an increase in movement speed can be understood as a decrease of available processing time per unit of output. Thus increasing movement speed and decreasing processing time should reduce accuracy for the same reason.

Interestingly a common theme that reverberates through many definitions of skill learning is that practice leads to shifts of the speed-accuracy trade-off function, such that the same performance can be achieved at shorter processing times (Reis

et al., 2009; Costa RM, 2011; Shmuelof et al., 2012). For example when playing a piano tune it might initially be very hard to produce the desired finger presses at the required speed. However, if the same piece is played at a slower speed fewer errors occur. Finally, after extensive practice, there will be equally many errors irrespective of whether the piece is played at a fast or at a slow pace. Practice affects the time-accuracy function in two ways. It shifts the curve as a whole and it decreases its overall steepness. Such changes in the time-accuracy trade-off will also be referred to as automatization throughout the remainder of this thesis.

The central hypothesis of this thesis is that in contrast to skill learning, sensorimotor adaptation does not result in changes of the speed-accuracy trade-off function. This idea can also be motivated by the concept of structural learning.

1.5.1 Time-accuracy trade-offs motivated by Structure Learning

The idea of structure learning originated in the field of artificial intelligence in the study of inductive learning algorithms (Dietterich and Michalski, 1981). The general idea is to reduce the dimensionality of a complex optimization problem, such that it can be solved more efficiently in a lower-dimensional space. Structure learning was only introduced into motor control very recently (Braun et al., 2010). As an example, consider a cyclist who learns to ride a motor-cycle. At first the cyclist might have trouble controlling the motorcycle, however he will be able to transfer some of his bicycle skills to the motorcycle and then improve performance very rapidly. The rapid learning is possible because the nervous system has a structural representation of balancing a two-wheeled vehicle condensed into a lower-dimensional space. It will not alter movement parameters at random in the hope that performance improves. Rather the nervous system can search the reduced

subspace that it has previously learned for an optimal solution (Braun et al., 2009a, 2010). This can also be expressed as having a set of Bayesian priors or beliefs about causal relationships of the environment. For example when participants made reaching movements in a position-dependent force field, the nervous system interpreted and learned subsequent velocity-dependent force fields initially as position dependent as well (Yousif and Diedrichsen, 2012).

As the system adapts, it can reuse the previously learned low dimensional structure and utilize the directional information contained in the prediction error signal to reset its reference points within the existing structure (Braun et al., 2010, 2010; Yousif and Diedrichsen, 2012). Thus, when translating the visual hand representation by a few centimeters to the left, the directional information in the prediction error automatically results in a translation within the established control structure. However, if no lower-dimensional structure has been established before, movements require computations in a higher-dimensional space. As a result movement preparation requires additional processing time.

1.5.2 **Time-accuracy trade-offs motivated by Internal Models**

From the perspective of internal models and control policies, during adaptation an existing control policy is updated. The error signal is translated into an adjustment of the old control policy by simply processing the visual prediction error through the old existing inverse model. Therefore after perceiving a directional error signal, no new control policy needs to be computed and thus no *change* of the speed-accuracy trade-off emerges. However this does not mean that during adaptation speed-accuracy trade-offs do not exist in general. If the movement which is being adapted had a speed-accuracy trade-off function before the onset of

adaptation, then the general form (i.e. the slope) of this function during and after adaptation will remain the same as before. Instead the function will shift as a whole, such that more accurate movements can be produced as a result of recalibration.

In contrast, skill learning does not rely on an existing internal model to compute a state update in response to an error signal. Thus a new policy needs to be learned. When a motor command cannot be computed from an existing inverse model, the computational demand will be relatively high. From a structural learning perspective higher processing times would be predicted because the space in which the system searches for the optimal output, has not been condensed into a lower dimensional structure yet. Therefore, if the system generates motor outputs under suboptimal temporal constraints we should observe a time-accuracy trade-off in addition to any trade-off that existed before.

Mirror reversed reaching is a prime example of a task where the existing internal model cannot be used for motor learning. In fact, almost 70 years ago Sperry (Sperry, 1947) trans-positioned agonist and antagonist nerves in the forearm of macaque monkeys and documented that: "*Not only did reversed movements appear during the early stages following nerve regeneration but,..., the reversed action persisted in some instances for months and even years.*" He noticed that especially during rapid instinctive movements the arm would be moved in the wrong direction, for example away from the food instead of towards it. Humans with paralysis to forearm muscle groups have been treated by transferring tendons from unaffected antagonist muscle groups. Although the forearm muscles could generally reverse their function, fast ballistic movements expressed unreversed EMG signatures even after years of living with transposed agonist and antagonist nerves (Illert et al.,

1986). Crucially the reversal of motor outputs that has to be learned after the nerve transposition does not require a truly novel motor output.

At this point it becomes clear why the goal-based part of the definition of skill learning by Shmuelof et al. (2012), does not hold. It is safe to assume that the primary goal in this task is to return to baseline levels of performance. Thus following a goal-based definition this task should be considered an adaptation task, while in fact the presence of speed-accuracy trade-offs suggests that this kind of learning might be similar to learning in many tasks that are traditionally considered skill learning tasks. The framework proposed in this thesis predicts the emergence of speed-accuracy trade-offs after nerve transposition: The old internal model is so strongly at odds with the external world that the state updates inferred from the prediction errors point in the wrong direction. Thus prediction errors cannot drive adaptation in this task.

1.6 Offline consolidation

Another characteristic of different motor learning mechanisms is how learning consolidates or is forgotten in periods in which the movement is not produced. As will be shown in this thesis – this criterion also appears to dissociate control policy recalibration from acquisition.

Consolidation is an umbrella-term that describes plasticity-related changes that take place after active practice has ended. Two phenotypes of consolidation can be distinguished: Stabilization and enhancement of the learned material or skill (Robertson et al., 2004a). Although it is not clear in how far stabilization and enhancement are the same or different processes, enhancements as opposed to

stabilization reveal an interesting property of learning. They show that a dynamic process continues after the end of training.

Performance enhancements between practice-sessions are frequently termed 'offline gains' (Robertson et al., 2004a, 2005; Korman et al., 2007; Doyon et al., 2009b). To this end offline gains have been found in declarative and cognitive tasks such as learning stimulus-response rules and insight (Wagner et al., 2004) as well as in sequence learning (Wright et al., 2010) and finger tracing tasks (Abe et al., 2011). Offline gains have also been shown for skill learning tasks, where for example participants are trained to produce certain sequences of finger presses (Fischer et al., 2002; Robertson et al., 2004a, 2004b; Doyon et al., 2009b).

In contrast, force field adaptation and visual displacement adaptation - in which for example during reaching movements the cursor is rotated - usually show forgetting (Debas et al., 2010; Trempe and Proteau, 2010). The same is true for saccadic and smooth pursuit adaptation (Kahlon & Lisberger 1996, Xu-Wilson et al. 2009).

Skill learning and sensorimotor adaptation tasks differ in many respects. Due to the many differences at the task level between usually relatively simple adaptation tasks and more complex skill learning tasks, it has to date been impossible to compare consolidation in skill learning with consolidation in recalibration. Therefore it is unclear whether differences at the level of tasks or at the level of mechanisms involved are responsible for offline gains or forgetting respectively. Interestingly offline gains have been reported in an arm movement adaptation task that required learning a 180° visual rotation (Doyon et al., 2009b). As I have argued earlier many adaptation tasks have in common that they can be learned by deriving the required control policy from an already existing internal

model, whereas most skill learning tasks have in common that the new control policy has to be established de novo. Although considered an adaptation task, a 180° rotation can, like mirror-reversal, not be derived from an existing internal model. I therefore explicitly tested whether the consolidation of control policies that are established de novo (e.g. in mirror reversed reaching) differs from the consolidation when the control policy can be derived from an existing internal model (i.e. in a 40° rotation task) in chapter 2.

Interestingly many of the studies that find offline gains also find that the gains are sleep-dependent. The currently prevailing view is that replay of neural activation patterns during slow wave sleep in areas of the hippocampus and Neocortex, strengthens the neural representations of the to-be remembered material (Diekelmann and Born, 2010; Oudiette and Paller, 2013). While the exact role of sleep is not clear yet, there is evidence that it does play a role in many types of learning, such as rule learning (Peyrache et al., 2009), bird song learning (Derégnaucourt et al., 2005) and a range of tasks used for studying motor skill learning (Walker et al., 2002; Stickgold, 2005; Diekelmann and Born, 2010). In particular, sleep has repeatedly been linked to offline improvements in finger sequence learning in humans (Walker et al., 2002; Korman et al., 2007; Doyon et al., 2009b).

Therefore we tested whether a night of sleep would benefit offline consolidation in a mirror reversed reaching task (see Chapter 2).

1.7 Hypotheses and predictions

We have now covered the relevant background to arrive at a testable framework for dissociating between the mechanisms behind adaptation and skill learning. In summary the most important hypotheses of this thesis are:

I) Recalibration relies on an inverse model that can generate beneficial state updates from the directional information contained in the visual error signal.

II) If the mismatch between the existing inverse model and the external sensorimotor mapping becomes so large that the state updates derived from the old inverse model cannot be used to reduce the size of the error, control policy acquisition takes over.

III) Recalibration reuses an existing control policy. Therefore the computational demands and the time required for an accurate output are unchanged.

IV) Acquisition cannot exploit the automaticity of an efficient existing control policy but instead a new control policy needs to be established. Therefore the computational demands and the time required for an accurate output increase.

V) Acquisition benefits from stronger consolidation than recalibration and sometimes shows offline gains.

In chapter 2 I will address hypotheses I through IV by comparing reaching movements under visual rotations of 40° to 60° with reaching movements under mirror-reversal. If recalibration requires an error signal that can easily be translated by an existing inverse model, then the 40° to 60° visual rotations should be learnable by recalibration of an existing control policy. Therefore no time-accuracy trade-offs should emerge for the visual rotations. In contrast, the old inverse model cannot be

used to produce useful state updates in mirror-reversal learning. A new control policy must be acquired and speed-accuracy trade-offs should become visible for mirror-reversal learning. The experiment was performed over two sessions and allowed a comparison of how different motor memories are consolidated between sessions. If the process underlying mirror-reversal learning really is similar to the dominant mechanism in other skill learning tasks, there should be relatively little forgetting and potentially even offline gains during control policy acquisition. In contrast adaptation tasks typically result in forgetting from one session to the next.

Chapter 3 will scrutinize hypotheses I and II even more carefully. First I ask whether any potential differences between learning rotations and mirror-reversal are due to the task itself (mirror-reversal vs. rotation) or whether time-accuracy trade-offs can be elicited in rotation learning as well. From hypotheses II one would predict that if the disparity between the existing inverse model and the rotation size is sufficiently large, recalibration cannot take place. Therefore I tested whether speed-accuracy trade-offs emerge if the size of the mismatch between inverse model and visual rotation is sufficiently large. In one condition the size of the rotation and thus the error was gradually increased. In this way I was able to determine the size of the visual error at which a potential switch from recalibration to control policy acquisition took place. The results also shed light on the question of whether the absolute size of the imposed rotation or the prediction error drives recalibration.

Chapter 4 explores if the way the system learns from errors can be changed. To test this idea, participants performed mirror reversed reaching movements for 4 consecutive days. Simultaneously we injected small perturbations to the endpoint feedback of the targeted reaching movements. By then studying how subsequent movements changed based on the perturbation of previous movements, it was

possible to estimate if the system had learned to mirror reverse the way it learned from errors.

Chapter 2
Adaptation and Skill learning

2.1 Abstract

Motor learning tasks are often classified into adaptation tasks, which involve the recalibration of an existing control policy (the mapping that determines both feedforward and feedback commands), and skill-learning tasks, requiring the acquisition of new control policies. We show here that this distinction also applies to two different visuomotor transformations during reaching in humans: Mirror-reversal (left-right reversal over a mid-sagittal axis) of visual feedback vs. rotation of visual feedback around the movement origin. During mirror-reversal learning, correct movement initiation (feedforward commands) and online corrections (feedback responses) were only generated at longer latencies. The earliest responses were directed into a non-mirrored direction, even after 2 training sessions. In contrast, for visual-rotation learning no dependency of directional error on RT emerged, and fast feedback responses to visual displacements of the cursor

were immediately adapted. These results suggest that the motor system acquires a new control policy for mirror-reversal, which initially requires extra processing time, while it recalibrates an existing control policy for visual rotations, exploiting established fast computational processes. Importantly, memory for visual rotation decayed between sessions, whereas memory for mirror-reversal showed offline gains, leading to better performance at the beginning of the second session than in the end of the first. With shifts in time-accuracy trade-off and offline gains, mirror-reversal learning shares common features with other skill-learning tasks. We suggest that different neuronal mechanisms underlie the recalibration of an existing vs. acquisition of a new control policy, and that offline gains between sessions are a characteristic of latter.

2.2 Introduction

Humans are experts in adjusting their movements to changing task demands (Helmholtz, 1866; McLaughlin, 1967; Gentilucci et al., 1995). Learning a new task requires a change in the functions that translate goals (and states) into motor commands. These functions have been synonymously referred to as visuomotor mappings, control policies, or inverse models (Sutton and Barto, 1998; Todorov and Jordan, 2002).

But are all new tasks learned the same way? Here we contrast the learning processes for two different visuomotor transformations: visual rotation and mirror-reversal. It has been suggested that mirror-reversal and visual rotations are learned using separate learning mechanisms (Werner and Bock, 2010). Here we hypothesize that visual rotation can be learned by a gradual recalibration of the existing control policy, while mirror-reversal requires the establishment of a novel mapping. This idea is motivated by how the motor system uses error to update future movements (Fig. 2.1). When confronted with visual rotations, the correction calculated under the old policy will be directed approximately (for rotations smaller than 90°) in the appropriate direction. The new policy therefore could be learned by updating the next motor command with the correction calculated following the outdated mapping (Kawato and Gomi, 1992). Repeated applications of this learning rule leads to the correct policy. During mirror-reversal, however, the update inferred from the old mapping points in the wrong direction and a novel policy would have to be acquired instead.

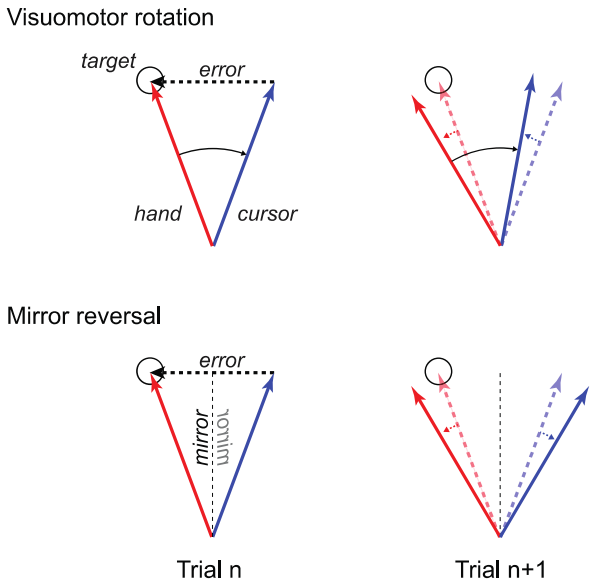

Figure 2.1: Schematic drawing of recalibration during mirror-reversal and visual rotation.

The dashed vertical line represents the mirror-reversal axis. In trial *n* hand (red) movements towards the -20° target (see Fig. 2.2 for coordinate frame) result in the cursor (blue) travelling to +20°, thus producing an error (dashed black arrow) of 40°. A fraction of this error vector is used to update the next motor command. On trial n+1 the hand movement direction (solid red arrow) is therefore shifted from the previous movement direction (dashed red arrow). During visual rotation (upper panel) this leads to error reduction between cursor (solid blue arrow) and target compared to the previous movement. During mirror-reversal (lower panel) the same update results in an increased error.

Krakauer and colleagues suggested that the difference between recalibration and acquisition is visible in speed-accuracy trade-offs (Reis et al., 2009; Shmuelof et al., 2012). Because fast sequential movements require the rapid generation of feedforward and feedback commands, this likely relates to the speed of the underlying computational processes: When the system recalibrates a well-learned control policy, it should be able to utilize existing fast automatic processes and

generate accurate responses even under time pressure. The establishment of a new control policy, however, should entail initially slower, and possibly more explicit components (Willingham, 1998; Hikosaka et al., 2002) requiring additional processing time. Only with long practice, it should become automatized and achieve equivalent performance at shorter time intervals. Thus, we expected that the acquisition of a control policy would be accompanied by a shift in time-accuracy trade-offs. We tested this idea by studying fast feedforward and feedback commands.

Finally, we also tested whether visual-rotation and mirror-reversal learning differ in how the memory consolidates between sessions. Adaptation tasks typically show forgetting between sessions (Tong et al., 2002; Klassen et al., 2005; Krakauer et al., 2005; Trempe and Proteau, 2010), whereas skill-learning tasks such as learning novel sequences of finger movements show little forgetting (Reis et al., 2009), and sometimes even offline gains (Wright et al., 2010; Brawn et al., 2010; Doyon et al., 2009a; Abe et al., 2011; Stickgold, 2005). Given that skill-learning tasks are also characterized by shifts in speed-accuracy trade-off (Reis et al., 2009; Shmuelof et al., 2012), we hypothesized that mirror-reversal learning may also show offline gains between sessions.

2.3 Materials and methods

2.3.1 Participants

All participants (N=112, 52 male) were right-handed according to the Edinburgh handedness inventory (Oldfield, 1971) and aged 18-30. None had a history of neurological illness and or were taking medication. Participants were recruited through online advertising, and received monetary compensation (£7/hour) at the conclusion of the study. Informed consent was obtained before the study started, and all procedures were approved by the UCL Ethics Committee.

2.3.2 General procedure

Participants made 15cm center-out reaching movements to targets displayed on a TFT LCD, while holding a robotic handle with the right hand. The robotic device allowed unrestrained movement in the horizontal plane and was able to exert forces to the participant's hand. Movements were recorded at 200Hz. Visual feedback was provided on a monitor (60Hz refresh rate) that was viewed via a horizontal mirror placed over the participant's hand. The delay of the visual display (65ms) was empirically measured using a photodiode and taken into account in the analysis of the data. Due to the mirror, the arm and hand were not directly visible. The position of the right hand was represented on the mirror by a cursor (2 mm diameter).

At the beginning of each trial the robot guided the participant's hand to the start location, a small rectangle, ~15cm in front of the participant's chest. After the hand remained inside the start rectangle for more than 400ms, a target (0.7x0.7cm^2 square) appeared on the screen. To probe the time-dependency of the forward command under the two visuomotor mappings, it was essential to enforce tight

bounds on reaction time (RT) - the time from target appearance to movement onset. Thus, participants were instructed that their first priority was to react quickly to the onset of the target. We played an unpleasant buzzing tone for slow reactions (RT>385ms), and an unpleasant high beep for anticipatory movements (RT<35ms).

A movement was considered started when the tangential velocity exceeded 3.5 cm/s and ended when it fell below 3.5cm/s. For offline analysis the velocity threshold for the movement start was set to 2.5cm/s. Participants were also instructed that their movements had to be fast and accurate to receive points. If the movement time (MT) – the duration from movement onset to termination - was too long or if the peak velocity was too low (<40 cm/s), all items turned blue; if the peak velocity was too high (>100cm/s), yellow. Green feedback indicated that the peak velocity was in the correct range but the movement was terminated outside of the tolerance zone around the target. Only when all criteria were met, did all items in the visual display turn red and a pleasant sound was played, signaling that the participants had gained a point. Participants were explicitly informed and then familiarized with these criteria over the first 4 practice blocks. The target zone in which the movement had to end was initially set to 1.2cm, and the maximum MT to 1200ms. These criteria were manually adjusted after each block to maintain a constant average success-rate: If a participant achieved over 50% of all points in the last block, both criteria were decreased by 0.1cm and 100ms, respectively, until they reached 0.7cm or 800ms. This adjustment ensured that the rate of reward stayed within a motivating range. Visually, the target always remained the same size (0.7cm), because changes of target size might have caused participants to alter their strategy. For offline analysis, we included all trials, irrespective of whether they satisfied the criteria described above (see data analysis).

2.3.3 **Experiment 1: Mirror-reversal, feedforward control**

The experiment consisted of two testing sessions, in which 15 participants were exposed to a mirror-reversed environment. The two experimental sessions took place between 4 and 10pm on two consecutive days for all participants. Participants reached from a central starting location to one of 6 possible targets located at -20°, 0°, +20°, +160°,180°, and -160° (Fig. 2.2).

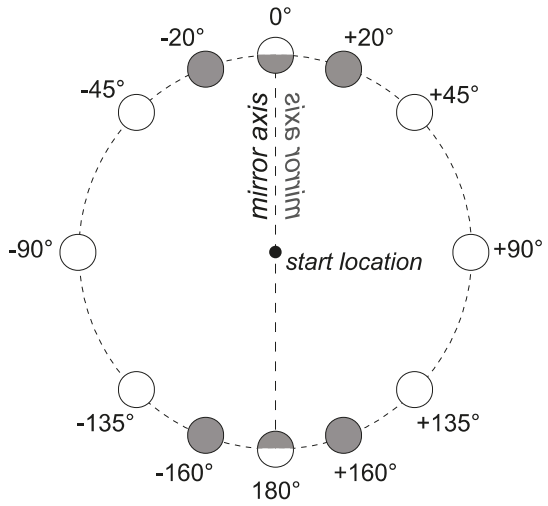

Figure 2.2: Target arrangements in experiments 1 & 2.

Grey circles indicate target locations in Experiment 1, whereas white circles indicate target locations in Experiment 2. Targets at 0° and 180° are half-grey half-white because they were presented in both experiments. The dashed vertical line indicates the mirror-reversal axis in Experiment 1. In Experiment 2 the rotations were applied relative to the start location.

Each session consisted of 16 blocks, each comprising 72 trials. The first session started with 4 training blocks to familiarize participants with the performance feedback (not included in the analysis) followed by 4 baseline blocks (blocks 1-4). Visual feedback was mirrored during the following 8 blocks of the first session (blocks 5-12); e.g. to reach to the right target, one had to generate a reaching movement to the left. In the second session visual feedback was mirrored during the

first 12 blocks (blocks 13-24). In the last 4 blocks of the second session visual feedback was returned to normal (blocks 25-28). Each block contained a total of 72 trials consisting of 12 reaches towards each of the 6 targets. Note that the 4 lateral targets (-160°, -20°, +20° and +160°) were chosen so that the required change in the motor command equaled 40° and would match the required change in the visual rotation condition (see below). To assess the state of the feedforward command in all experiments, we measured the initial movement direction, the angular hand position averaged from 100 to 150ms after movement onset. This early measure is relatively uninfluenced by possible feedback corrections (Franklin and Wolpert, 2008).

In Experiment 1-4, participants were informed in the break between block 4 and 5 that a visuomotor transformation would be imposed, and the nature of the transformation (visual rotation or mirror-reversal) was explained to them. We then stressed that their first priority should be to initiate their movement within RT limits, even if it meant that they missed the target. These restrictions largely prevented participants from consciously re-planning their movement endpoint (Georgopoulos and Massey, 1987; Mazzoni and Krakauer, 2006; Neely and Heath, 2009; Taylor et al., 2010; Taylor and Ivry, 2011).

2.3.4 **Experiment 2: Visual rotation, feedforward control**

Experiment 2 had generally the same structure as Experiment 1, with two testing sessions taking place on consecutive days. This time the participants (N=15) were exposed to a 40° visual rotation instead of a mirror-reversal of the cursor. As noted above, the required change in the motor command from the original to the new mapping in Experiment 1 was also 40°, such that the magnitude of the mapping

change was equal in both experiments. Center-out reaching movements were executed towards 8 circularly arranged targets (Fig. 2.2). Feedback regarding movement performance was given following the same criteria that were used for Experiment 1. Each session consisted of 16 blocks, and each block contained 72 trials, with each target appearing 9 times in random order. Again the first 4 of the 16 blocks in the first session were training blocks and were excluded from all further analyses. This was followed by 4 baseline blocks, and 8 blocks in which a +40° visual rotation was imposed. The second session began with 12 visual-rotation blocks, followed by 4 blocks without rotation.

2.3.5 **Experiment 3: Mirror-reversal, feedback control & sleep**

Whereas Experiment 1 and 2 assessed learning of feedforward control, Experiment 3 was designed to also assess learning of fast feedback commands with mirror reversed visual feedback, by laterally displacing the cursor on a fraction of trials. Additionally, we tested the hypothesis that consolidation between sessions depended on sleep, motivated by the finding that sleep has been reported to benefit offline consolidation (Walker et al., 2002; Stickgold, 2005). Experiment 3 had generally the same structure as experiment 1 and 2, using identical feedback procedures, number of trials per block, and the number of blocks per day. We tested feedback control only for the 0° target, as here no change in the feedforward command was required that could possibly confound the measurement. To increase the number of reaches to each target, we only tested targets at -20°, 0°, and 20°. Each block was divided into 9 miniblocks and each miniblock consisted of 8 different trials (Table 2.1), designed to test either feedforward or feedback control. The trials within each miniblock were ordered randomly, with each trial type occurring once.

To test changes in feedforward commands, reaching targets in trial types 1 and 2 were presented at an angle of 20° or -20° from straight-ahead. As in experiments 1 and 2, the angular hand position averaged from 100 to 150ms after movement onset was measured for studying feedforward control. In the remaining 6 trials in each miniblock participants reached to the straight-ahead target and we tested fast feedback mechanisms. For trial types 4, 5, 7 and 8 we displaced the cursor by 1.5cm to the left or right after the hand had travelled more than 1cm from the origin. Cursor displacements elicit an automatic corrective response in the opposite direction with the aim of bringing the cursor back to the initial trajectory. This response has shorter latencies than voluntary response initiation (Franklin and Wolpert, 2008) and cannot be voluntarily suppressed.

Trial type	1	2	3	4	5	6	7	8
Force channel	×	×	×	×	×	✓	✓	✓
Target location	-20	20	0	0	0	0	0	0
Cursor displacement	× ← →	× ← →	×	←	→	×	←	→

Table 2.1: Trial types within every miniblock in Experiment 3.
Note that for trial types 1 and 2 each of the three cursor displacements (none, left, right) occurs only once for every 3 miniblocks. Crosses indicate the absence of cursor displacements or force channels, whereas ticks and arrows indicate the presence of force channels and direction of cursor displacements respectively.

To obtain a sensitive measure of the feedback response, we clamped the hand to a straight-line trajectory towards the target using a force channel for trial types 6-8. These channels exerted a spring-like force of 6000N/m. When a cursor was displaced, participants pushed into the channel wall attempting to correct for the

displacement. The hand force was immediately counteracted by an equal amount of force from the robotic handle, which could then be used as a reliable measure of correction. On force channel trials the cursor was displaced back to the original trajectory after the hand had moved more than 10cm in the channel to allow the participants to reach the target. Because the automatic return of the cursor can cause attenuation of feedback responses (Franklin and Wolpert, 2008) we also added trials without channels (trial types 4 and 5) in which the cursor was not returned. These trials therefore required a correction to reach the target. For the same reason, we also displaced – and did not return - the cursor on 2 out of 3 trials in which the movement was directed at lateral targets (trial types 1, 2).

Blocks/ Groups	4 normal	8 MR	break	12 MR	4 normal
ME	Morning		12 h	Evening	
EM	Evening		12 h	Morning	
EE	Evening		24 h	Evening	
MM	Morning		24 h	Morning	

Table 2.2: Experimental groups in Experiment 3 with testing sessions at different times of day.

Note that both days consisted of 16 blocks, each containing 72 reaching movements. The first 4 blocks of day 1 were training blocks with normal visual feedback and are not listed in the table. MR= Mirror-reversal.

To determine whether performance changes between the sessions (forgetting or offline gains) depended on sleep, we assigned participants to one of four groups (table 2.2). The first group (morning-evening, ME; 16 participants) had the first session in the morning and the second session 12 hours later on the same day. The second group (EM; 15 participants) had the first session in the evening and

the next session 12 hours later after a night of sleep in the morning of the next day. To control for the effect of the time of day on performance, we included one control group that did both sessions in the evening (EE; 13 participants) and one that did both sessions in the morning (MM; 17 participants). For both groups the sessions were separated by a 24-hour break and a night of sleep. There were no significant age or gender differences between the 4 groups. Morning sessions took place between 7:30 and 10:30am and evening sessions between 7:30 and 10:30pm. Note that the role of sleep was only tested for mirror-reversal, but not for visual rotation, because no offline improvements were found for the latter.

2.3.6 **Experiment 4: Visual rotation, feedback control**

Experiment 4 was designed to assess changes in fast feedback control during visual-rotation learning, and was again similar in length and structure to Experiment 1-3. Movements were executed towards 8 targets. Instead of a +40° rotation, we imposed +60° or -60° rotations (balanced across 18 participants), to achieve sufficient power to detect changes in the direction of feedback corrections. On 48 of 72 trials the cursor position was displaced by 1.5cm once the hand had travelled more than 1 cm from the origin. Because force channels are only suitable to measure feedback corrections orthogonal to the movement direction, we assessed fast feedback responses using the direction of the initial corrective response in free movements. This was measured by computing the difference in instantaneous velocity of the hand on trials with and without displacements. The cursor displacement was applied after the hand had travelled 1cm from the start at an angle of -90° or +90° relative to the initial movement direction of the cursor, and therefore always at an angle of -30° or +150° relative to the movement direction of the hand

(Fig. 2.6b). An unadapted feedback response would yield an initial hand direction exactly opposing the visual displacement. For example if the cursor was displaced -90° relative to the cursor direction (or -30° relative to the hand, dashed dark blue arrow) the correction should be directed towards 150° (Fig. 2.6b, solid light blue arrow). A fully adapted feedback response would be rotated by 60° opposite to the imposed visual rotation, thus resulting in a +90° correction if the cursor was displaced -30° relative to the hand (Fig. 2.6b, solid dark blue arrow).

2.3.7 **Experiment 5: Control experiment for feedback response**

Experiment 4 relies on the assumption that the feedback response is always opposite to the cursor displacement, independent of the direction of hand movement. That is, we assumed that the visuomotor system corrects equally for displacements parallel and orthogonal to the direction of movement. To test this assumption, 3 participants performed reaching movements over 16 blocks towards 8 different targets without a visual rotation. We then displaced the cursor by 1.5 cm at angles of -150°, -90°, -30°, +30°, +90° and +150° relative to the initial hand and cursor movement direction (Fig. 2.6a, dashed colored arrows). If both orthogonal and parallel displacement components are corrected equally, the correction should always be exactly opposed to the displacement (Fig. 2.6a, solid colored arrows). In addition each block contained two movements without displacement towards each target.

2.3.8 **Data analysis**

The data were analyzed using custom-written MATLAB routines. For all 5 experiments we excluded movements where the angle between the first and the

second 100ms segment after movement onset was bigger than 60°, as a large difference between the two segments indicates that the movement was initially not directed at the target and only corrected online thereafter. Trials with peak movement velocities <40 or >100cm/s or RTs <50ms or RTs> 730ms were excluded in Experiment 1-3. For Experiment 3, we further excluded channel trials where force responses exceeded 5 Newton (N) at any point in time between 150 and 400ms after the cursor displacement. Because the main variable of interest in Experiments 4 and 5 was the corrective velocity vector, we excluded for these experiments trials where the peak velocity deviated by more than 25 cm/s from the median in the respective block, but included all trials independent of their reaction time. Combined, these criteria led to an exclusion of 5.4% of the trials in Experiment 1, 5.5% in Experiment 2, 4.5% in Experiment 3, 4.8% in Experiment 4 and 4.4% in Experiment 5.

In Experiment 1, trade-offs between preparation time and accuracy of the feedforward command were quantified by the slope of the simple linear regression between RT and error. A trade-off would show up as a negative relationship between these two variables. Assessing this relationship is complicated by the fact that both RT and error reduce over the course of learning, leading to a positive relationship that could obscure existing time-accuracy trade-offs. To account for this effect, we first removed - within each subject and block - any linear trend across the block for RT and error independently. The movements towards the peripheral targets were then assigned to 1 of 5 bins according to this relative RT. This was done for each block, each participant, and each target separately. To obtain more stable estimates, we then combined the data across all 4 lateral targets by mirroring results towards the -20° and +160° onto the +20° and -160° targets. Furthermore, we averaged the data across 4 blocks for each participant. As a measure of the relationship between

RT and error, we performed a simple linear regression analysis with the mean RT of each bin as the independent, and the mean signed error as the dependent variable, separately for each subject and block. The slope values were then compared using paired t-tests. The time-accuracy trade-off for visual rotations in Experiment 2 was assessed using a similar analysis, while rotating the data to combine results across all 8 targets.

In Experiment 3, we compared the state of the feedforward command across days. Because of the possible RT-dependency of the feedforward command, and because mean RTs could change from session to session, we determined the expected initial error for a RT of 250ms. For this, the relationship between RT and error was fitted for each participant, each block and each target separately. Because this relationship was slightly non-linear, we used Gaussian Process Regression (Rasmussen, 2006), which can accommodate any smooth relationship between two variables. The values of the length scale, variance and noise variance hyper parameters were determined by fitting the data from all subjects together for each mirror reversed block and then taking the median values.

For Experiment 4 and 5, data was combined across all targets by rotating the movement data such that the movement direction 1cm into the movement was located at 0°, because the cursor displacements were always performed at an angle relative to this initial movement direction. We then used the difference between the average instantaneous velocity vector of trials with and without displacements to compute the velocity component that was due to the corrective response.

2.4 Results

2.4.1 Time-accuracy trade-off in feedforward commands

We hypothesized that the learning of mirror-reversal would be associated with a new time-dependent process that maps targets to actions, whereas visual-rotation learning would be supported by the recalibration of an existing control policy, and should therefore require no extra processing time.

We tested this idea by enforcing fast RTs in all reaching tasks. For mirror-reversal learning (Experiment 1, Fig. 2.3a), RTs increased at the onset of mirror-reversal by 145ms (±18ms standard error), $t_{(14)} = -8.232$, $p < 9.8*10^{-7}$.

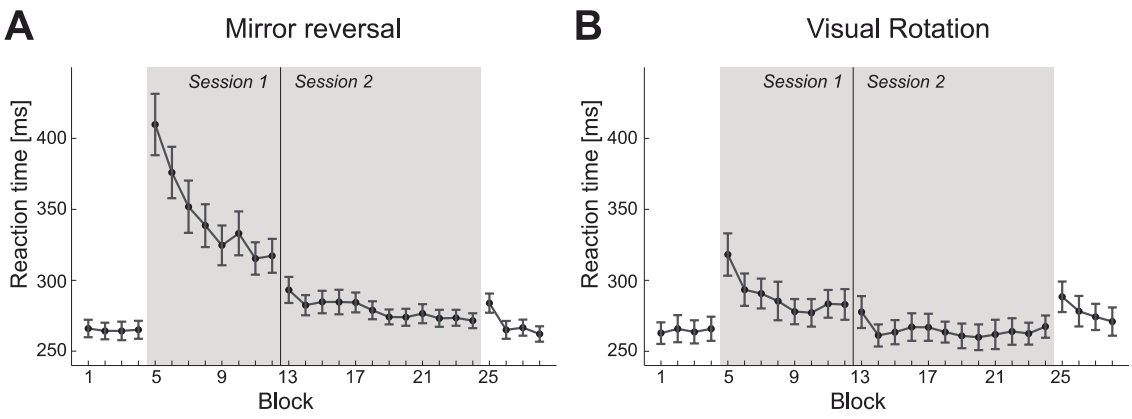

Figure 2.3: Group-average reaction time across Experiment 1 & 2.

White background indicates reaching under normal visual feedback, while grey background indicates reaching during mirror reversed or rotated visual feedback. The vertical line indicates the break between sessions. (A) RT for -160°, -20°, 20° and 160° targets during mirror-reversal learning (Experiment 1). (B) RT for reaching towards 8 targets during visual rotation (Experiment 2). Error bars indicate between subject standard error.

RTs reached a plateau in the late mirror reversed blocks of the second session and approached the levels of the baseline performance. However, when the visual feedback switched back to the non-reversed mapping in block 25, RTs increased at

first but subsequently decreased to 272ms (±5ms) in the last block, yielding almost significantly shorter RTs than the last mirror reversed block ($t_{(14)}$ = 2.123, p = .052). Thus, even after two days of training, movements in a mirror-reversed environment required slightly more preparation time than in the normal environment.

For the equivalent visual-rotation experiment (Experiment 2, Fig. 2.3b), we expected RT to increase to a lesser degree, if at all. Average RT increased by 45ms (±8ms) when the rotation was first introduced ($t_{(28)}$ = -2.918, p = .007) (Fig. 2.3b). Thus the increase of RT during visual-rotation learning was considerably smaller than the increase during mirror-reversal learning ($t_{(28)}$ = -5.170, p = $1.74*10^{-5}$). During the second day of training, none of the visual-rotation blocks differed significantly from baseline anymore (block13: $t_{(14)}$ = -1.683, p = .114). After the rotation had washed out (last block), the RTs were not significantly shorter than in the 4th block of training ($t_{(14)}$ = -1.256; p = .23). Thus, we found that visual rotations induced less than a third of the RT increase as compared to mirror-reversal.

Our main prediction, however, was that the difference between the two learning mechanisms should become visible in a time-accuracy trade-off, i.e. the fact that – for a given adaptation state - trials with longer RTs show smaller errors. Since reaction times as well as movement errors decreased over the course of the experiment, we first subtracted out any possible linear relationship between trial number and error and between trial number and reaction time for each participant and block separately in the mirror-reversal and the visual-rotation conditions. We then plotted the initial movement direction of the hand (averaged from 100 to 150ms after movement onset) as a function of RT for different groups of 4 blocks (Fig. 2.4). For mirror-reversal learning (Experiment 1, Fig. 2.4a) baseline reaching angles were offset from zero by approximately +5°, indicating that participants

showed a bias towards moving in the straight forward or backward direction (see caption of Figure 2.3 on how angles were combined across targets), an effect likely caused by the unequal distribution of targets around the circle.

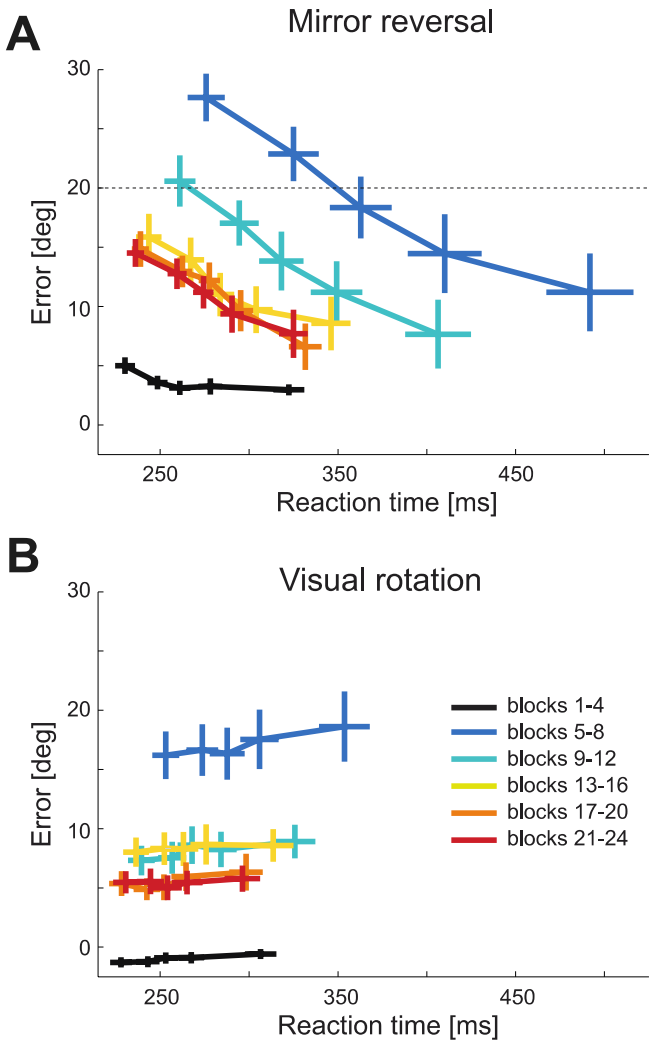

Figure 2.4: Relationship between RT and directional error in Experiments 1 and 2.

Blocks 1-4 were collected during baseline and blocks 5-24 during mirror-reversal or visual rotation. The trials were binned by RT for each target, participant and block. Visual feedback was veridical during blocks 1-4 and mirror reversed or rotated during blocks 5-24. Blocks 1-12 were measured during the first, blocks 13-24 during the second session. (**A**) Mirror-reversal: Visual Errors from movements towards the -160° and +20° target were flipped to allow averaging with errors from the -20° and +160° targets. Visual Errors larger than 20° indicate that the hand reached into the wrong (unmirrored) direction. Completely unadapted responses would yield an error of +40°. (**B**) Visual rotation. A completely unadapted response would result in an error of +40°. Error bars indicate between-subject standard error.

To determine whether there was a time-accuracy trade-off, we calculated the regression slope between error and RT across bins (see methods) (Fig. 2.4). In the mirror-reversal experiment (blocks 1-4), there was a small, but significant negative slope, ($t_{(14)}$ = -4.477, p = .001) during baseline. With the beginning of mirror-reversal learning (blocks 5-8), the slope became significantly more negative compared to baseline ($t_{(14)}$ = 5.004, p = 1.93*10^{-4}). For long RTs, participants produced the correctly mirrored movements. However, for the fastest RT bin, movements started in the direction of the visually presented target, rather than in the opposite, correct direction; the error was significantly larger than 20°, where a 20° error signifies a movement towards the mirror-reversal axis ($t_{(14)}$ = 3.812, p = .001). As training proceeded, the relationship between RT and movement error retained similar slopes across all groups of 4 blocks (repeated measures ANOVA with groups of 4 blocks as within-subject factor: $F_{(4, 56)}$ = .588, p = .673). Even in the end of training in Experiment 1, the difference in the RT-error relationship was still significant compared to baseline ($t_{(14)}$ = 3.995, p=.001). However, the time-accuracy curve shifted sideways, such that higher accuracies could be achieved at shorter RTs. To quantify this observation, we calculated the RT necessary to reduce the error to 12° - as this time point allowed for assessment for all groups of 4 blocks of the experiment (Fig. 2.4a) - by assuming an approximately linear relationship between error and RT in the range tested here and linearly predicting the reaction time for an error of 12° for each participant and quadruple of blocks. We found significant differences between blocks 5-8 and blocks 9-12 ($t_{(14)}$ = 2.405, p = .031), blocks 13-16 ($t_{(14)}$ = 4.836, p = 2.64*10^{-4}), blocks 17-20 ($t_{(14)}$ = 3.769, p = .002), and blocks 21-24 ($t_{(14)}$ = 3.860, p =0.002). Likewise we found significant horizontal shifts between blocks 9-12 and blocks 13-16 ($t_{(14)}$ = 2.806, p = .014), blocks

17-20 ($t_{(14)}$ = 3.405, p = .004), and blocks 21-24 ($t_{(14)}$ = 3.353, p = .005), meaning that each curve on day 2 was significantly shifted compared to each curve on day 1. In other words, mirror-reversal training led to automatization of the new target-to-movement mapping, visible in a shift of the time-accuracy trade-off.

In contrast, we hypothesized that visual-rotation learning (Experiment 2) is achieved by the recalibration of an existing control policy. Participants should therefore be able to exploit the automaticity of the old mapping even during learning, and should thus not require additional time for processing. Hence, we predicted that for visual-rotation learning, longer reaction times should not result in lower errors. This is indeed what we found (Fig. 2.4b). At baseline there was a small but significant positive relationship between error and RT ($t_{(14)}$ = 3.453, p = .004). However, with the introduction of the visual rotation, this relationship did not change (t-test between the slopes of blocks 5-8 and blocks 1-4: $t_{(14)}$ = -1.442, p = .171). Thus, although angular errors increased as soon as the visual display was rotated (blocks 5-8), longer RTs did not result in smaller errors. In subsequent blocks, the error reduced further, but no change in the dependency on RT was observed (t-test between the slopes of blocks 21-24 and blocks 1-4: $t_{(14)}$ = .503, p = .623).

Although the range of RTs between Experiment 1 and 2 were slightly different, the RT distribution overlapped considerably, especially for the later learning phases. To compare the mirror-reversal and visual-rotation conditions directly, we recalculated the slopes between RT and reach angle for the fastest 4 bins during mirror-reversal and the slowest 4 bins during visual-rotation learning, such that the average reaction time used for calculating the slopes in mirror-reversal (292ms ±9ms) and visual-rotation (279ms ±9ms) were not significantly different,

$t_{(28)}$ = 1.053, p = .301. After subtracting the baseline slopes from all other phases we found that in all phases, there was a significant difference between the time-accuracy slope of the mirror-reversal and visual-rotation conditions (blocks 5-8: $t_{(28)}$ = 4.429, p = 1.4*10^{-4}; blocks 9-12: $t_{(28)}$ = 5.101, p = 2.1*10^{-5}; blocks 13-16: $t_{(28)}$ = -4.781, p = 5.05*10^{-5}; blocks 17-20: $t_{(28)}$ = 3.420, p = .002; blocks 21-24: $t_{(28)}$ = -4.401, p = 1.4*10^{-4}). Thus, over a comparable range of RTs, the mirror-reversal group clearly showed a significantly stronger dependency of accuracy on RT than the visual-rotation group.

2.4.2 **Adaptation of fast feedback responses**

A second window of insight into how computations in the motor system unfold over time is to investigate fast feedback responses. If a new control policy requires more time to compute a motor command, then the feedback responses after learning should also be delayed – or possibly the early responses should be dominated by the old policy. If, however, an existing policy was recalibrated, then both early and late components of the feedback response should adapt simultaneously.

To address this question for mirror-reversal learning, Experiment 3 probed the reactions of the arm to sudden displacements of the cursor (Sarlegna et al., 2003). We then calculated the difference between force responses to left and rightward cursor jumps and halved it to inspect the temporal evolution of the feedback correction in different groups of 4 blocks of the experiment (Fig. 2.5 shows the results averaged across the 4 consolidation conditions).

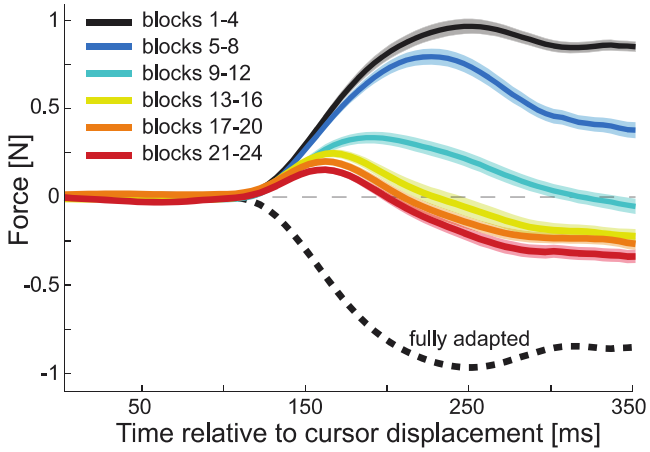

Figure 2.5: Relationship between time and feedback response during mirror-reversal learning (Experiment 3).

Shown is the force measured in the channel produced in reaction to a 1.5cm cursor displacement. Blocks 1-4 were collected during baseline and blocks 5-24 during mirror-reversal. The dashed line shows the reversed baseline response to serve as an illustration of what a perfectly mirror reversed feedback response would have looked like. Shaded area indicates between-subject SE.

During unmirrored baseline movements the corrective response began about 110ms after the onset of the displacement, and reached about 1N after 250ms. In the first 4 mirror-reversed blocks (blocks 5-8) it still reached around 0.8N in the same direction, but became less sustained thereafter; in the time window 250-350ms, it was significantly lower than during baseline, $t_{(60)} = 8.35$, $p = 1.2*10^{-11}$. Note that this unreversed response would increase the visual error, rather than compensate for it (Fig. 2.1). In blocks 9-12 the force response further decreased, but still did not reverse. Only during the second day, (blocks13-24) did we observe a reversal of the force response in the time window 250-350ms (blocks 13-16, -0.14N ±0.038N, $t_{(60)} = -3.695$, $p = 4.8*10^{-4}$). Yet, even in blocks 21-24, the initial incorrect force response was not fully abolished: in the time window between 130-

200ms, it remained significantly positive (0.13N ±0.018N, $t_{(60)}$ = 8.028, $p = 4.3*10^{-11}$).

In sum, feedback responses during mirror-reversal learning provide a very similar picture as feedforward responses. While the system generates correct movements after additional processing time, the fast and automatic responses remained unadapted even after 2 training sessions (Gritsenko and Kalaska, 2010). The data clearly showed a progression of learning in which the correct response was progressively generated at shorter delays, suggesting that the new control policy, which was initially rather slow, became automatized.

Determining how feedback commands adapt during visual rotation is more challenging, as the adapted and unadapted response are not opposite to each other, but differ only by the imposed rotation angle. To amplify the contrast, we conducted another study (Experiment 4) in which participants adapted to either a +60° or a -60° rotation, and probed feedback responses by displacing the cursor orthogonally to the cursor movement (±90°, Fig. 2.6b, dashed dark blue and red arrows). In the condition in which the cursor was rotated by +60°, the effective visual displacement was in a direction -30° and +150° relative to the hand movement. For a fully adapted feedback response, the hand should correct orthogonally to the hand trajectory as before (Fig. 2.6b, solid red & dark blue arrows). In contrast, if the feedback response is unadapted, the correction should be opposite to the visual displacement, i.e. +150° or -30° relative to the hand movement direction (Fig. 2.6b, solid orange & light blue arrows).

The latter prediction, however, relies on the assumption that participants would correct their hand movement opposite to the visual cursor displacements, even if the displacement were not orthogonal to the movement direction. Because it

is possible that the motor system reacts less to the component of the visual displacement in the direction of the movement, we tested our assumption in an additional experiment. In Experiment 5, we displaced the cursor by 1.5 cm at an angle of ±30°, ±90° and ±150° relative to hand and cursor movement (Fig. 2.6a). Even for the oblique angles, the initial correction should be exactly opposite to the cursor displacements.

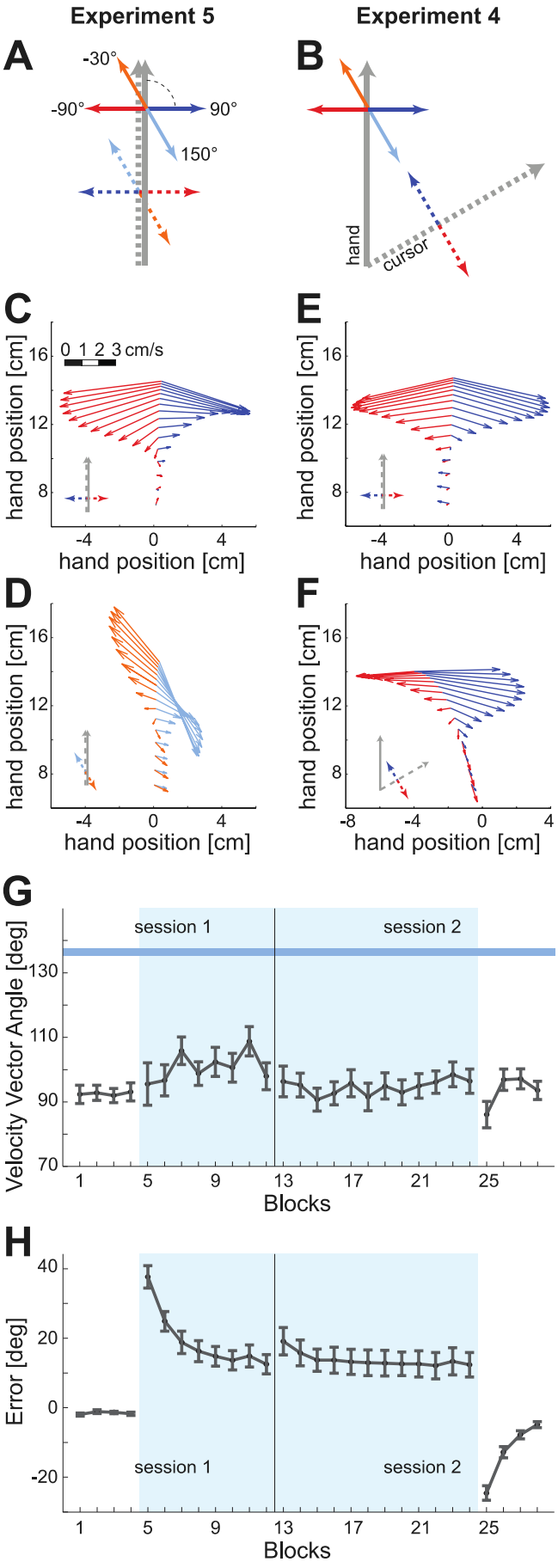
Experiment 5
Experiment 4
A
-30°
-90°
90°
150°
B
hand
cursor
C
0 1 2 3 cm/s
hand position [cm]
16
12
8
-4
0
4
hand position [cm]
E
hand position [cm]
16
12
8
-4
0
4
hand position [cm]
D
hand position [cm]
16
12
8
-4
0
4
hand position [cm]
F
hand position [cm]
16
12
8
-8
-4
0
4
hand position [cm]
G
session 1
session 2
Velocity Vector Angle [deg]
130
110
90
70
1
5
9
13
17
21
25
Blocks
H
Error [deg]
40
20
0
-20
session 1
session 2
1
5
9
13
17
21
25
Blocks

Figure 2.6: Feedback responses in Experiment 4 & 5.

(**A**) In Experiment 5, the cursor (dashed grey line) and the hand (solid gray line) moved in the same direction. The cursor was displaced (dashed colored arrows) at an angle of -90° (dark blue), -30° (light blue), +90° (red) or +150° (orange) relative to the movement direction. Displacements also occurred in +30° and -150° directions (not shown). The hand movements that cancel out the cursor displacements are shown as solid arrows of the same color. (**B**) In Experiment 4, the cursor (dashed gray line) was rotated by +60° or -60° (only the +60° is shown in the schematic) from the hand movement (solid gray line). Displacements were -90° (blue dashed) or +90° (red dashed) relative to the movement direction of the cursor. The solid red and dark blue arrows indicate the required hand movement directions that cancel out the corresponding displacement (dashed arrow with the same color). The orange and the light blue arrows show what an unadapted response would look like. (**C**) Quiver plot of feedback responses in Experiment 5 to -90° (dark blue) and +90° (red) cursor displacements. The vector origin represents the average hand position at time points from 75 to 375 ms after the cursor displacement (20ms resolution), and the vector the difference in instantaneous hand velocity between trials with and without displacement. (**D**) Feedback responses to -30° (light blue) and +150° (orange) cursor displacements in Experiment 5. (**E**) Response to -90° (dark blue) and +90° (red) cursor displacements during baseline reaching, i.e. before cursor rotation in Experiment 4, and (**F**) with rotated cursor (blocks 5 to 8). Results are shown averaged over the +60° and -60° rotation groups, by right-left flipping the results for the -60° group. (**G**) Mean angular direction of feedback correction (±SE) 250 to 350ms after the displacement plotted over all blocks of Experiment 4. Responses are combined across cursor displacements and rotation groups. Light blue background: blocks with visual rotation. Blue line and shading: prediction of fully unadapted feedback response, based on mean and SE of responses to oblique cursor displacement in Experiment 5. (**H**) Mean angular error of the feedforward command (±SE) averaged from 100 to 150ms after movement onset while adapting to the 60° rotation in experiment 4 for comparison.

We used the difference between the instantaneous velocity vectors between trials with and without displacements at different time points after the displacement as a measure of the corrective response. We found, that for the 90° displacements under the natural mapping, the velocity difference vectors were slightly tilted downwards, meaning that the hand not only corrected in the appropriate direction, but also decelerated along the main direction of movement (Fig. 2.6c). To summarize the effects across displacement directions offline, we rotated the correction vector for the -90° displacements by 180°, effectively canceling out any decelerating effect.

For oblique displacements, we found that the corrections were approximately opposite to the displacement (Fig. 2.6d). To analyze the responses together we inverted the horizontal component of the responses to the +150° and +30° displacements, and the vertical component of the responses to the ±150° displacement, such that all corrections would superimpose with the correction for the -30° displacements (which requires a +150° correction for full cancellation). The angle of the resulting correction was +136.4° (±9.1°), slightly less than the ideal response of +150°, indicating that participants reacted to displacements in movement direction slightly less than to displacements orthogonal to it. Thus, based on these results we would expect that a fully unadapted feedback response to an anticlockwise (-90°) cursor displacement under a +60° cursor rotation should be +136.4°.

In Experiment 4, we averaged the results of the +60° and -60° rotation groups, by flipping the trajectories for the group that underwent the -60° rotation. The average feedback responses during visual-rotation learning (Fig. 2.6f) did not resemble the feedback responses observed in the control experiment (Fig. 2.6d). Rather, the corrections were oriented -90° and +90° relative to the movement

direction. In other words, the feedback response in visual-rotation appeared to be immediately oriented in the correct direction (Fig. 2.6g). Although we cannot directly compare the forces measured in Experiment 3 with the velocity vectors measured in Experiment 4, these results contrast starkly with the slow and incomplete adaptation of fast feedback responses during mirror-reversal learning.

Our results therefore suggest a fundamental difference in the way in which mirror-reversal and visual-rotation are learned. Mirror-reversal learning initially requires extra processing time to compute accurate feedforward and feedback commands, indicating that it may involve the establishment of a new control policy. Although the new motor commands could be generated more quickly after 2 days of training, it remained dependent on processing time. In contrast, visual-rotation learning did not show such dependency even early in learning - consistent with the idea that here a fully automatized control policy was recalibrated.

2.4.3 **Offline gains in performance between sessions**

With the shifting time-accuracy trade-off, mirror-reversal learning shares an important feature with other motor learning tasks (Beilock et al., 2008). It has been recently suggested that such shifts should be considered the defining feature of "skill learning" (Reis et al., 2009; Shmuelof et al., 2012). Another characteristic of many tasks that are considered "skill" tasks concerns consolidation between sessions: For example, for learning of sequential movements, performance levels typically deteriorate very little overnight (Rickard et al., 2008), and sometimes even appear to show offline gains (Stickgold, 2005; Wright et al., 2010; Abe et al., 2011). In contrast, adaptation tasks that require a recalibration of an existing control policy nearly universally show some decay of the motor memory during an intervening

interval (Tong et al., 2002; Klassen et al., 2005; Krakauer et al., 2005; Trempe and Proteau, 2010). If this different temporal dynamic of consolidation can be attributed to the suggested distinction of automatization of a new control policy vs. recalibration of an existing control policy, then mirror-reversal learning should show offline gains in the break between the two sessions, whereas visual-rotation learning should show offline forgetting.

Offline gains in skill learning experiments are often reported to depend on sleep (Walker et al., 2002; Robertson et al., 2004a; Cohen et al., 2005a; Stickgold, 2005; Rickard et al., 2008). For mirror-reversal learning in Experiment 3 we therefore randomly assigned the participants to one of four groups. The ME group had the first session in the morning and the second session in the evening of the same day, and therefore did not have a night of sleep between the two sessions. The EM group had the first session in the evening and the next session in the morning of the next day. Both of these groups had a break of 12 hours between their two sessions. To test whether potential differences depended on the time of day of the first or second session, rather than on the presence or absence of sleep, we included two additional groups which performed the experiment either on the mornings (MM) or on the evenings (EE) of two consecutive days. If consolidation really depended on sleep but not time of day, then only the ME group (the only group without sleep) should show worse consolidation than any of the other three groups, while the other three groups should not differ from each other.

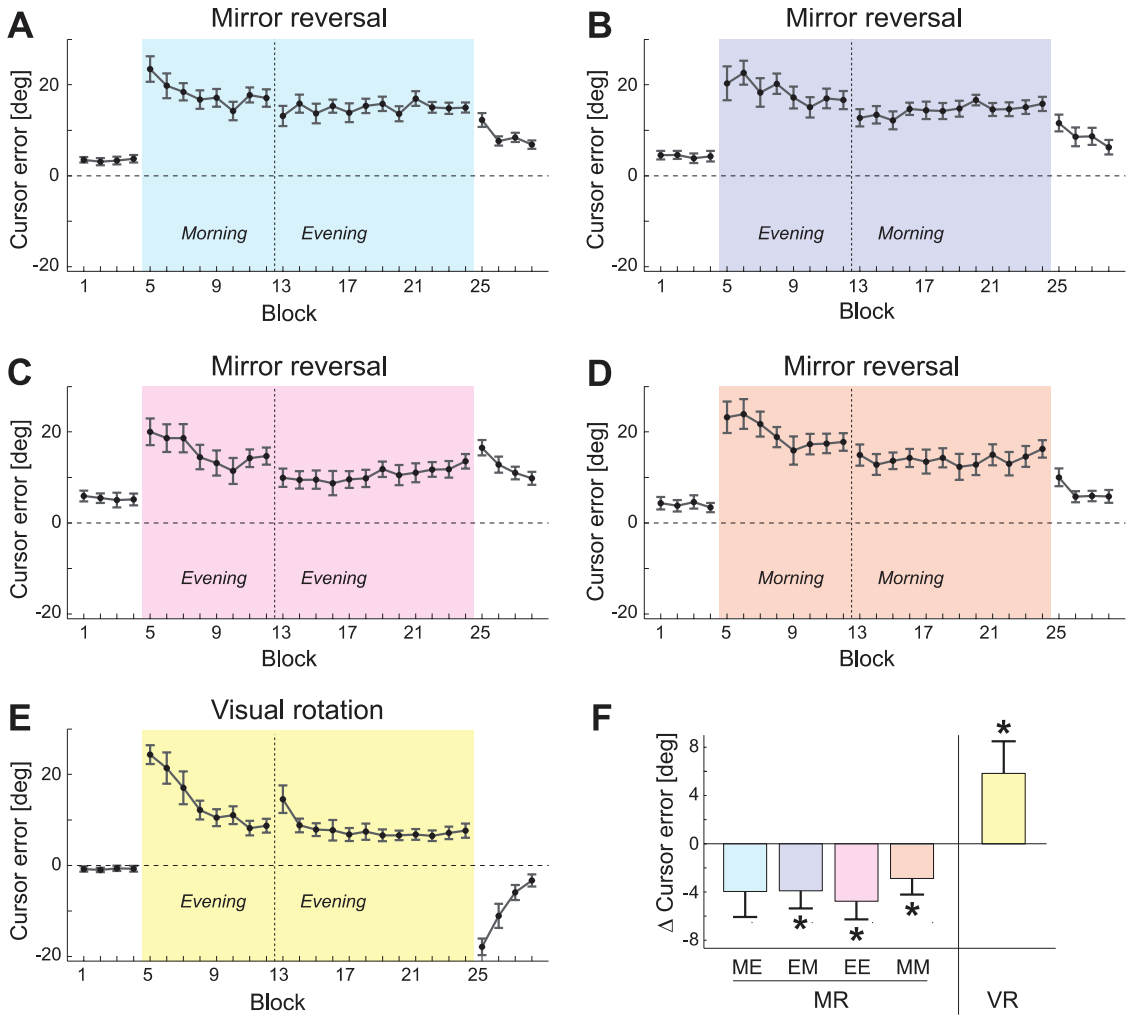

Figure 2.7: Consolidation of the feedforward command in Experiment 2 & 3.

Average angular errors 100 to 150ms after movement onset are plotted over the different blocks of the experiment (**A-D**) for the 4 mirror-reversal groups (Experiment 3) and (**E**) the visual rotation group (Experiment 2). The error is corrected for the influence of time-accuracy trade-off by calculating the average error at RT=250ms (see methods). Colored background indicates blocks with mirror-reversal or visual rotation. The vertical dashed line separates the two sessions. All mirror-reversal groups performed as well or better in the first block of the second session than in the last block of the first session. (**F**) Bar graph of the difference in error between the first block in the second session (block 13) and the last block in the first session (block 12) split up by the visual rotation and the four mirror-reversal (MR) groups ME=Morning Evening; EM=Evening Morning; EE=Evening Evening; MM=Morning Morning and the VR=visual rotation group. * indicates significant t-test against zero with p<.05.

Because error depended on RT, and because RT may differ from one session to the next, we quantified the skill level as the movement error that the participant would show for a fixed RT. The slightly non-linear relationship between error and RT was fitted using Gaussian Process Regression (see methods), and we then simply read off the movement error for an RT of 250ms. Errors from movements towards the +20° target were inverted, so that the RT-corrected directional error for both peripheral targets could be averaged.

We found that mirror-reversal learning did not show forgetting between sessions, but rather offline gains in performance (Fig. 2.7). Across all groups, there was a significant improvement in feedforward performance from the last block of the first session to the first block of the second session ($t_{(60)} = -4.72$, $p = 1.4*10^{-5}$). Tested individually, the EM group ($t_{(14)} = -2.678$, $p = .018$), the EE group ($t_{(12)} = -3.174$, $p = .008$) and the MM group ($t_{(16)} = -2.138$, $p = .048$) all significantly improved over night. The only group that did not show significant improvements was the ME group ($t_{(15)} = -1.872$, $p = .081$), which did not have a night of sleep between the two sessions. However, there was no significant direct difference between the group without sleep and the groups with a night of sleep between the two sessions in terms of their change in movement error from session 1 to session 2 ($t_{(59)} = -1.471$, $p = .147$).

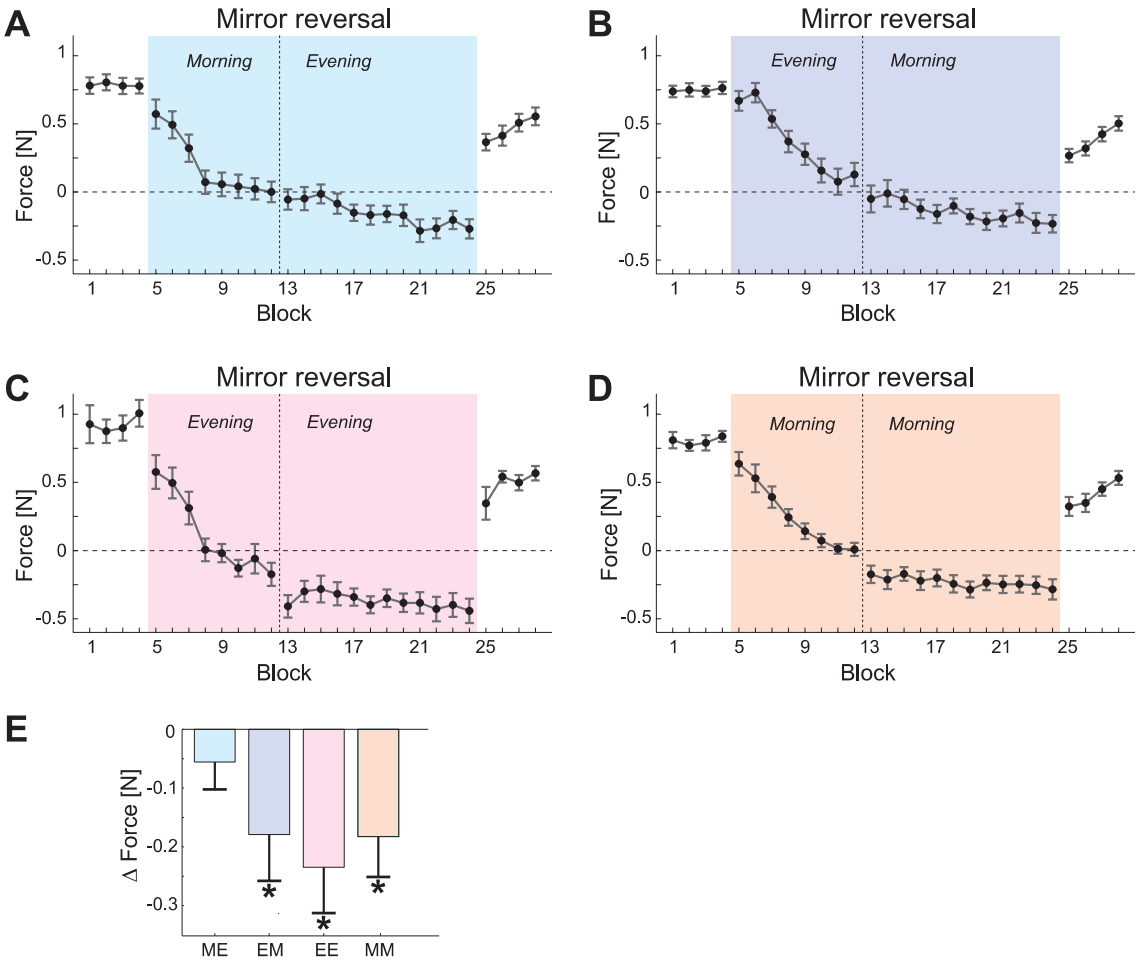

Figure 2.8: Consolidation of the feedback command in Experiment 3.

The average feedback command 250 to 350ms after the displacement is plotted over different blocks of the experiment. Colored background indicates mirror-reversal of the visual feedback. (**A - D**) Panels show the feedback commands of the four mirror-reversal groups. (**E**) Bar graph of the force differences between the first block in the second session (block 13) and the last block in the first session (block 12) split up by the four groups. ME=Morning Evening; EM=Evening Morning; EE=Evening Evening; MM=Morning Morning. * indicates significant t-test against zero with p<.05.

Offline gains were even more clearly visible in the feedback corrections (Fig. 2.8). For this analysis, we averaged the feedback response (Fig. 2.5) over the interval from 250 to 350ms after the displacement, as this time period showed the most profound learning-related changes. Again all participants combined showed

very strong offline gains ($t_{(60)}$ = -4.637, p = 1.9*10^{-5}). We also plotted this measure as a function of block for all 4 groups separately. The EM group ($t_{(14)}$ = 2.265, p = .04), the EE group ($t_{(12)}$ = 3.011, p = .011) as well as the MM group ($t_{(16)}$ = 2.656, p = .017) showed significant increases in performance from one session to the next. The only group that did not show improvements was the ME group ($t_{(15)}$ = 1.189, p = .253), i.e. the group that did not have a night of sleep between the two sessions. The groups with sleep had only marginally stronger offline gains than the group without sleep ($t_{(59)}$ = 1.837, p = .071), indicating that offline improvements may have been enhanced by sleep. There was no significant effect of time of day of the first ($t_{(59)}$ = 1.220, p = .227) or the second session ($t_{(59)}$ = .650, p = .518) nor an effect of the duration of the break between the sessions ($t_{(59)}$ = 1.314, p = .194). Taken together these results clearly demonstrate the existence of offline gains during mirror-reversal learning. In respect to the sleep dependency of this effect our results remain inconclusive. Even though there are some trends in the data that indicate that an intermitted night of sleep may amplify this effect, the direct comparison of the groups failed to reach significance.

In contrast to mirror-reversal, visual-rotation learning showed clear forgetting between sessions, in line with many other adaptation tasks (Tong et al., 2002; Klassen et al., 2005; Krakauer et al., 2005; Trempe and Proteau, 2010). Although we did not find a significant relationship between RT and angular error, we used, for the sake of consistency, the same method for RT correction as for the mirror-reversal data. Within the first day, the initial error reduced from 24.4° (±2.1°) to 8.7° (±1.5°) (Fig. 2.7e). When participants returned on the second day their error had increased again to 14.72° (±3°). Angular errors in the first block of the second session were significantly larger than angular errors in the last block

of the first session ($t_{(28)}$ = -2.192, p = .049, Fig. 2.7f). Thus, our results confirm previous literature showing that adaptation is forgotten between sessions, and provide evidence for a clear dissociation from mirror-reversal learning, for which offline gains are observed.

2.5 Discussion

We directly contrasted learning of two different visuomotor transformations. For mirror-reversal learning, we found a clear RT-dependency of initial movement error, with faster responses leading to larger errors than slow responses. We hypothesized that mirror-reversal learning involves the acquisition of a new sensorimotor mapping, which initially takes more time than the old mapping to perform the necessary computations. Therefore, under strict time constraints, the response was still dictated by the old mapping. With 2 days of training we found that the new mapping became increasingly automatic, achieving the same movement error at shorter RTs. It did not, however, achieve the same automaticity as the baseline mapping.

For visual-rotation learning, movement error did not decrease with increasing RT. We propose that this form of motor learning relies on the recalibration of an already existing mapping, and therefore can exploit the established automaticity of the underlying computational processes. Thus, in this view, the appearance of a time-accuracy trade-off at the beginning of learning with subsequent shifts of this relationship is a cardinal sign that the motor system acquires and automatizes a new mapping from goals to motor commands (Reis et al., 2009; Shmuelof et al., 2012).

Intriguingly, we found a parallel dissociation between mirror-reversal and visual-rotation learning during fast feedback responses to displacements of the visual cursor. For mirror-reversal learning, the corrective response was initially directed into the wrong direction, even after 2 days of training (Day and Lyon, 2000; Gritsenko and Kalaska, 2010) and reversed only in the late phases of the response. Thus feedforward and feedback control both require additional processing time in the beginning of learning and then are increasingly automatized.

In contrast, the feedback command during visual-rotation learning appeared to be fully adapted immediately. It has been suggested that feedback responses during large visual-rotation must adapt rapidly within a single trial, because the hand would otherwise circle around the target (Braun et al., 2009b). Another explanation might be that the feedback command does not need to adapt at all, because it always bases its reactions on the relative angle between the displacement and the visually observed trajectory. Whatever the exact mechanism, the presence of time-accuracy trade-offs in mirror-reversal, and their absence during visual rotation, provides clear evidence that the two visual transformations are learned via separate processes.

A previous study found a relationship between RT and how quickly participants learned a 60° visual rotation (Fernandez-Ruiz et al., 2011a). However, in this study RTs were unconstrained and on average 400-600ms. The authors argued that unconstrained RTs may have invited strategic re-planning of the endpoint (Mazzoni and Krakauer, 2006; Taylor et al., 2010; Taylor and Ivry, 2011), a process more related to an explicit mental rotation of the desired movement direction (Georgopoulos and Massey, 1987; Neely and Heath, 2009) than to visuomotor adaptation. Indeed, when RTs were constrained to below 350ms as in

our study, no evidence for a time-accuracy trade-off in visual-rotation learning was found. These results therefore argue that even visual rotations are not always learned purely through recalibration of an existing control policy: without speed constraints additional time-consuming processes (strategic remapping) can help to improve performance more quickly.

Why does the brain have to learn a new control policy for mirror-reversal, while it appears to recalibrate an existing control policy for visual rotations? At a computational level of description (Marr and Poggio, 1976), mirror-reversal and visual-rotation learning seem to be comparably difficult. Both can be described with a simple change in the function that transforms visual inputs into arm movements. However, what is difficult for the brain has to be viewed in the context of its prior experience. In ambiguous situations, the motor system appears to interpret visuomotor errors as being caused by visual rotations (Turnham et al., 2011), possibly reflecting inherent assumptions about the structure of the environment. These priors can be changed through repeated exposure to different environments, a process termed structural learning (Braun et al., 2009c). Viewed in this framework mirror-reversal learning would be slow, as it violates the learned structure of possible visuomotor transformations – requiring the slow acquisition of a new structure. A related explanation is based on the assumption that a visuomotor mapping is adapted by adding some part of the corrective response under the old mapping to the old motor command (Kawato and Gomi, 1992). Visual rotations up to 90° could be learned like this, whereas for mirror-reversal the initial corrective response would point in the wrong direction (Fig. 2.1), again requiring the establishment of a new control policy. This hypothesis would make the –yet to be tested - prediction that rotations larger than 90° should also show time-accuracy

trade-offs. Indeed, it has been suggested that such large rotations are learned by different mechanisms (Abeele and Bock, 2001a).

Rather than providing a clear computational-level explanation, the main empirical contribution of the paper is to show that mirror-reversal and visual-rotation learning clearly differ in their time-accuracy trade-off, both in feed-forward and feedback control. We hypothesize that these trade-offs are tightly related to the trade-off between *movement speed* and accuracy – as faster movements impose tighter time constraints on feedback processes. Consistent with our interpretation, shifts in such speed-accuracy trade-offs have been interpreted as a sign of the establishment of a new control policy (Haith and Krakauer, 2013). Following this definition, the learning of new trajectories (Shmuelof et al., 2012), finger sequences (Karni et al., 1995) or finger configurations (Waters-Metenier et al., 2014) should have some similarity to mirror-reversal learning.

Our second main finding is that the presence of a time-accuracy trade-off is associated with how the learned behavior consolidates between sessions. For visual-rotation learning for which no time-accuracy trade-off was found, forgetting occurred between sessions. This is in line with other studies of adaptation (Kassardjian et al., 2005; Krakauer et al., 2005; Galea et al., 2011). For mirror-reversal learning we found clear evidence for offline gains, both in the feedforward and the feedback command. So far, offline gains have mainly been reported for motor learning of sequential movements (Robertson et al., 2004a). Our study provides to our knowledge the first reported instance of offline improvement for learning of visuomotor transformations during reaching movements.

There has been an extensive debate on whether true offline gains in sequential finger movements depend on sleep (Stickgold, 2005; Wright et al., 2010;

Abe et al., 2011). Our results do not allow for a definite conclusion in the mirror-reversal learning task: For both feedback and feedforward commands we found trends indicating that offline gains are brought about by sleep - however, a direct comparison of the different mirror-reversal groups did not reach statistical significance. Thus, our failure to find evidence of sleep-dependency may be partly due to a lack of power - and the relationship between sleep and memory in this context may warrant further study.

The presence of a time-accuracy trade-off and offline gains suggests that the mechanisms that underlie learning of mirror-reversal and 40° visual-rotations have different physiological underpinnings. Specifically, one may speculate that the establishment of a new control policy relies on cortico-striatal circuits. Indeed, Gutierrez-Garralda et al. (Gutierrez-Garralda et al., 2013) showed that Basal Ganglia patients exhibit normal learning in a dart throwing task when the visual scene is horizontally displaced, but impaired performance when the visual scene is mirror reversed (Stebbins et al., 1997; Laforce Jr. and Doyon, 2001). The Basal Ganglia have been associated with action selection (Gerardin et al., 2004) and the acquisition of new control policies (Doya, 2000; Middleton and Strick, 2000; Hikosaka et al., 2002; Boyd et al., 2009; Doyon et al., 2009a). In addition, Parkinson's and Huntington's disease patients are impaired in learning sequential finger movements and learning of other novel tasks (Gerardin et al., 2004; Boyd et al., 2009; Penhune and Steele, 2012). In contrast, the adaptation of eye movements (Takagi et al., 1998, 2000), arm movements (Martin et al., 1996; Tseng et al., 2007) and gait (Reisman et al., 2007), heavily depends on the integrity of the cerebellum, while basal ganglia associated disorders affect adaptation to a lesser degree (Fernandez-Ruiz et al., 2003; Marinelli et al., 2009; Gutierrez-Garralda et al., 2013).

A strict dissociation between the cerebellum as the substrate for adaptation/ recalibration and the basal ganglia as the substrate for control policy acquisition has recently been called into question with increasing evidence that the cerebellum is involved in both adaptation and "skill learning" (Penhune and Steele, 2012). Cerebellar patients are impaired in dart throwing tasks with horizontally shifted as well as with mirror reversed visual feedback (Sanes et al., 1990; Vaca-Palomares et al., 2013).

It has to date been very difficult to determine whether any differences found between adaptation and skill-learning tasks can be truly attributed to the underlying learning mechanism or the differences between the tasks that are used to measure them. Here we demonstrate that the two mechanisms are differently engaged in the learning of two different visuomotor mappings during reaching movements. The current paradigm may therefore be ideally suited for studying the neural correlates of acquisition and recalibration of control policies using functional imaging or neurophysiologic recordings within a single task.

Chapter 3
The limits of recalibration

3.1 Abstract

This chapter addresses why time-accuracy trade-offs emerged for mirror-reversal learning but not for 40° rotations in chapter 2. Here we found that when participants made reaching movements under 180° rotations, they showed time-accuracy trade-offs that were similar to those elicited by mirror-reversal. Therefore it could be ruled out that mirror-reversal and visual rotations are learned differently per se. To find the point until which recalibration can be used, in a separate condition the size of the rotation was gradually increased until it reached 180°. Participants learned slower than the rotation size increased, such that they would increasingly fall behind and then abruptly switch to a different type of behavior that yielded more accurate movements. While prior to the switch longer reaction times did not result in higher accuracy, after the switch movements were characterized by a marked time-accuracy trade-off. The exact time point of the behavioral switch was

determined by fitting a model that consisted of two components: an error-based recalibration component, and a component that resembled more strategic movements, which always reached the target accurately. The error-based component learned by devaluing error signals depending on their magnitude. Crucially we found strong evidence that the switch was not caused by an upper limit on the amount of adaptation, but that instead the size of the error signal determined the time point at which participants would switch from recalibration to a more strategic type of reaching: weak learners, who experienced larger errors earlier, switched earlier than strong learners.

3.2 Introduction

Learning to reach under different visuomotor transformations is supported by different memory systems. In chapter 2 we have shown that the type of learning that is commonly understood as adaptation is utilized when reaching under rotations of 40°, but not when learning left-right reversals. In this thesis, I refer to the former type of learning as recalibration of control policies, because it emphasizes the assumption that an existing mapping from internal target representation to motor output is being modified. However, for an existing mapping to be amenable to recalibration a mapping from inputs to outputs must have been acquired in the first place. It is intuitively attractive to think that skill learning tasks depend on the acquisition of new control policies, rather than the recalibration of existing ones. Mirror-reversal learning exhibits similar characteristics to skill learning tasks. In contrast to visual rotations of 40° mirror-reversal elicits offline improvements and time-accuracy trade-offs such that performance is higher at longer processing times (Chapter 2).

But why would humans require one memory system to learn visuomotor rotations of 40° and another memory system to learn mirror-reversal? I here test three alternative hypotheses that might explain the apparent dissociation. The first alternative hypothesis, proposes that the motor system has learned a structure of control policies that incorporates all possible visual rotations – i.e. where the visual rotation has become a free parameter in the control policy (Braun et al., 2009c, 2010; Turnham et al., 2011). Following this idea, irrespective of the size of the rotation, all rotations should be learned using parametric learning – or recalibration.

In contrast, mirror-reversal would not lie in the assumed structure and would therefore demand structural learning or the acquisition of a new control policy.

To test this idea we contrasted mirror-reversal learning with learning a 180° rotation. We found that 180° rotation and mirror-reversal learning showed marked time-accuracy trade-offs for those targets that required identical changes in the sensorimotor map for the two transformations (i.e. at -90° and +90°). Thus 40° rotations can be learned by recalibration, while 180° rotations and mirror-reversal are learned through the acquisition of a new control policy. As the structural hypothesis was falsified, we examined another condition in which we gradually increased the size of the visual rotation. We found that participants changed their learning behavior abruptly at rotation sizes of 148.5°-180°, showing discontinuous learning curves with instantaneous improvements in performance and emerging time-accuracy trade-offs. We proposed two alternative hypotheses to explain why learning switches from recalibration to acquisition in this condition.

One of our alternative hypotheses (the recalibration-limit hypothesis) proposed that there is an absolute upper bound on the amount that the system can recalibrate (Abeele and Bock, 2001a). If there was an absolute recalibration-limit, then strong learners should reach this limit earlier and thus change from recalibration to acquisition at an earlier time point to minimize the size of the error. In fact, given that recalibration elicits long-lasting after-effects (Kagerer et al., 1997), it might be desirable to learn separate policies for visuomotor transformations that lie very far apart, so that the nervous system can rapidly switch between different policies in a context-sensitive way.

Our third hypothesis follows from the role that we suggest for internal models in the conversion of prediction errors into state updates (the error-limit

hypothesis). It holds that not the absolute amount of recalibration has an upper bound. Instead there is an upper bound on the size of the error signals that can be utilized for recalibration (Körding and Wolpert, 2004; Wei and Körding, 2009). While this has previously been explained in a Bayesian framework through the attribution to internal and external sources of error, we here propose an alternative explanation for the case of visuomotor rotations. It has repeatedly been shown that performance is worst for visuomotor rotations of about 90° and that the deterioration in performance increases disproportionately strongly with the size of the rotation up to this point (Cunningham, 1989; Imamizu and Shimojo, 1995; Abeele and Bock, 2001b). A possible reason for this finding is that the error vector at 90° is orthogonal and thus uncorrelated to the old mapping in Euclidean space. We suggest that the system infers the required recalibration from the current inadequate mapping. Since the current mapping is wrong, the inferred correction will be inaccurate. Yet as long as the correlation - in Euclidean space - of the old mapping with the required mapping is bigger than zero, the inferred correction yields a recalibration that makes the following movement more accurate. Through iteration the system arrives at the correct solution. Figure 3.1 illustrates this idea at the hand of online corrections.

Therefore, contrary to the recalibration-limit hypothesis, the error-limit hypothesis predicts that fast learners experience the 90° upper bound on the prediction error later and thus they should switch from recalibration to acquisition at a later point in time than slow learners.

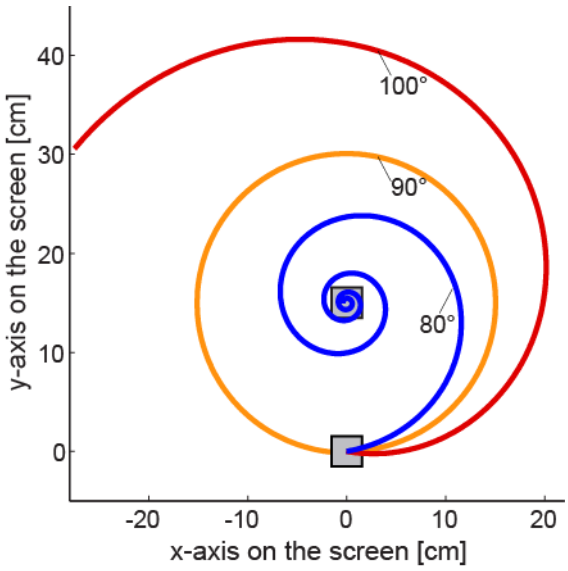

Figure 3.1: Simulation of high-frequency online corrections.

To illustrate the effects of relying on an outdated mapping to infer the updates for future motor commands we have simulated a system that starts a movement at [0,0] and aims at a target at [0,15], while updating its movement direction every .5mm based on the error vector. This simple system illustrates that when inferring online corrections from a mapping that is 80° off (blue), the correction will result in a reduction of the error. In fact, if we would allow the system to decrease the size of its updates based on the length of the error vector, it will converge at the target. In contrast, a mapping that is 90° off (orange) cannot result in corrections that reduce the error. Instead the hand would circle around the target at a constant distance. Finally if the mapping was off by more than 90° (red), the error would infinitely increase.

3.3 Materials and Methods

3.3.1 Participants

All participants (N=36, 15 male) were right-handed and aged 18-30. None had a history of neurological illness and or were taking medication. Participants were recruited through online advertising, and received 7£ per hour at the conclusion of the study. Before the study started the participants gave written informed consent. All procedures were approved by the UCL Ethics Committee.

3.3.2 General procedure

Participants made 15cm center-out reaching movements to targets displayed on a TFT LCD and viewed via a mirror, while holding a robotic handle with the right hand. The robotic device allowed unrestrained movement in the horizontal plane and was able to exert forces to the participant's hand. Movements were recorded at 200Hz. The delay of the visual display (50ms) was empirically measured using a photodiode and accounted for before data analysis. The horizontally placed surface mirror occluded the right arm and hand. Instead the position of the right hand was represented by a cursor (2mm diameter), which was displayed on the TFT LCD monitor (100Hz refresh rate) and viewed via the surface mirror.

At the beginning of each trial the cursor was invisible and the robot guided the participant's hand to the start location - a small rectangle, ~15cm in front of the participant's chest. Once the cursor was within a radius of 2cm from the center of the start location, the cursor became visible. After the hand remained inside the start rectangle for more than 400ms, a target (0.7x0.7cm^2 square) appeared on the

screen. Participants were instructed to execute a fast and accurate reach towards the target and to terminate the movement there. A movement was considered started when the tangential velocity exceeded 3.5cm/s and ended when it fell below 3.5cm/s. Visual feedback of the hand position (cursor) was removed once the movement was considered terminated. If the movement was successful, all items in the visual display turned red and a pleasant sound was played. After trials where the movement time - time from the start of the movement until termination - was too long or where the peak velocity was too low, all items turned blue; where the peak velocity was too high, yellow. Green feedback indicated that the peak velocity was in the correct range but movement termination was measured to be outside of the tolerance zone around the target rectangle.

To reveal the time-dependency of the feedforward command, we enforced lower and upper limits on reaction time, i.e. the delay between target appearance and movement onset. An unpleasant buzzing tone was played for slow reactions (RT>385ms), and an unpleasant high beep for anticipatory movements (RT<35ms). The specific thresholds were chosen consistent with earlier experiments (Chapter 2). In addition, feedback regarding movement time – the time from movement onset to termination – and spatial accuracy was delivered by changing the color of the cursor and target, and participants only received a point if all criteria were met. The target zone in which the movement had to end was initially set to 0.8cm, and the maximum movement time to 800ms during trials with normal visual feedback and 1.2cm and 1200ms respectively during trials with rotated or mirror reversed visual feedback. However, these criteria were not relevant for the inclusion of trials in offline data analysis.

3.3.3 **Conditions**

Participants were randomly assigned to 1 of 3 conditions. The three conditions differed only in the transformation applied to the cursor in blocks 33-128 (Table 3.1). In the mirror condition (7 participants) the cursor position was mirrored over the mid-sagittal axis. In the abrupt rotation condition (7 participants) the cursor feedback was instantly rotated by 180° relative to the instructed starting location of the hand. Finally, in the gradual rotation condition (14 participants), a visual rotation was introduced during blocks 33-72 starting from 4.5° and increasing by 4.5° in every block until it reached 180° in block 72 on day 1. It then remained at 180° for the following blocks on day 1 and until block 128 on day 2. For half of the participants in condition 3 the rotation was added clockwise and for the other half anticlockwise.

Day	1				2		
Block number	1-16	17-32	33-72	73-80	81-112	113-128	129-144
Block type	normal	CDB	normal	normal	normal	CDB	normal
Condition 1 (mirror)	0°		mirrored				0°
Condition 2 (abrupt)	0°		180° rotation				0°
Condition 3 (gradual)	0°		4.5°,…,180°	180° rotation			0°

Table 3.1: Structure of the experiment.
The 3rd row indicates the block type as either normal, i.e. without cursor displacements or CDB, i.e. cursor displacement blocks. The last three rows indicate the transformation of the visual feedback that the participants experienced in each of the conditions in the respective blocks, e.g. 0° indicates veridical feedback.

The experiment consisted of 2 sessions separated by a 24 hour break and taking place between 14.00 and 21.00. The first session lasted approximately 100 minutes and the second session approximately 70 minutes. Trials were arranged in blocks, which in turn were always presented in quadruples. Between any quadruple participants were free to take breaks and rest their arm. The first session consisted of 80 blocks (Table 3.1). Visual feedback in the first 32 blocks was normal. During the last 48 blocks in the first and the first 32 blocks in the second session the cursor position was either mirrored over the mid-sagittal axis between the participants' eyes (Cond. 1) or rotated around the starting point (Cond. 2 & 3). Visual feedback returned to normal again in the last 16 blocks of the 2nd session. Blocks 1-16, 33-112 and 129-144 were designed to assess the state of the feedforward command. Each of these blocks consisted of 6 targets - located at 150°, -90°, -30°, +30°, +90° and +150° - being presented three times and in random order within each block.

3.3.4 **Cursor displacement trials (not included in the analysis)**

Blocks 17-32 and 113-128 were designed to assess the feedback command to displacements of the cursor orthogonal to the required movement direction. Each of these cursor displacement blocks consisted of 20 trials, 2 of which were directed at each of the -150°, -30°, +30° and +150° targets, and 6 were directed at each of the lateral targets at -90° and +90° (Table 3.2). To obtain a sensitive measure of the feedback response, we clamped the hand to a straight-line trajectory towards the target using a force channel for three trials (downward, upward and no displacement) out of 6. These channels exerted a spring-like force of 6000N/m. When a cursor was displaced, participants pushed into the channel wall attempting

to correct for the displacement. On these trials the cursor was returned automatically in the end of the movement. The remaining three trials were channel-free and again consisted of one trial with an upward displacement, one with a downward displacement and one without displacement. During these movements the cursor was not realigned with the hand for the rest of the trial, and thus required an active feedback correction to counter the displacement. The channel-free cursor displacements were introduced to avoid possible attenuation of the feedback response which might be caused by the automatic realignment of cursor and hand during channel trials. However data collected in response to cursor displacements will not be presented in this thesis as they mainly replicate results from the feedback command in chapter 2 for 180° visual rotations.

Occurrences for each target location	2 for -150°, -30°, +30° & +150°; but only 1 for -90° & 90°	1				
Target location	-150°, -90°, -30°, +30°, +90° & +150°	-90°, +90°	-90°, +90°	-90, +90	-90°, +90°	-90°, +90°
Cursor displacement	×	←	→	←	×	→
Force channel	×			✓		

Table 3.2: The 20 trials of the cursor displacement blocks (33-48 and 113-128).

3.3.5 **Data analysis**

All data from the clockwise and anticlockwise gradual-rotation groups were combined, by inverting the reaching movements and rotations of the clockwise rotation group such that they overlaid with the data from the anticlockwise group. Circular statistics were used for calculating the moments of all circular distributions, such as reaching angles, reaching errors, reaching movement adaptation and model-prediction errors. For the gradual-rotation group, we hypothesized that participants would possibly abruptly change from recalibration to a different strategy. To detect such a change point, we created a model (Model 3.1), in which the reaches are either generated by a simple state space model, or by a different process, which always accurately hit the target. The moment for which this model switched from one behavior to the other was determined by a free parameter – the change-point – which was estimated along with the other parameters by minimizing the squared difference between the initial hand direction and the model predictions. We were exclusively interested in how learning feedforward control proceeds. Therefore we fitted the models to the average hand position 100-150ms into the movement and calculated the angular distance between this point and the straight line that connects start location and target. This early measure is highly variable, therefore to increase the reliability of the change point detection, we averaged the relative hand position across targets and repetitions within each block. Since A and B were highly correlated, we first estimated the initial state (constraints: z_0 [-2, 2], the retention factor A [0.6, 0.99] and the learning rate B [0, 0.6] freely. Thereafter we set A for all participants to the respective mean of A across participants for the cosine (0.895± 0.025) and the standard model (0.688± 0.037) separately and re-estimated B and z_0.

(I) $z_{n+1} = A * z_n - B * (u_n + z_n)$ if $n \in \{37,\ldots, CP-1\}$

(II) $z_{n+1} = -u_n$ if $n \in \{CP,\ldots, 112\}$

Model 3.1. Learning from angular errors. (**I**) describes a state space model that learns from angular cursor errors. (**II**) For all trials that occur after the change point the relative hand position is the inverse of cursor displacement and thus the cursor error equals zero. n = block, CP = change point, z = hand movement relative to the vector that connects start and cursor location, A = retention factor, B = learning rate, u = cursor displacement.

(I) $z_{n+1} = A * z_n - B * \cos(u_n + Z_n)$ if $n \in \{37,\ldots, CP-1\} \wedge (u_n + z_n) <= 90$

(II) $z_{n+1} = A * z_n$ if $n \in \{37,\ldots, CP-1\} \wedge (u_n + z_n) > 90$

(III) $z_{n+1} = -u_n$ if $n \in \{CP,\ldots, 112\}$

Model 3.2. Learning from cosine errors. (I) describes a state space model that learns from the cosine of the cursor error for cursor errors < 90°. This is equivalent to adding the inverse of the unit error vector in Euclidean space. (II) For errors larger 90° the error is set to 0. (III) Hand position is identical to the target location. n = block, CP = change point, z = hand location relative to the vector that connects start and cursor location, A = retention factor, B = learning rate, u = cursor displacement.

In general all movements were included in the analysis of the feedforward commands, except for procedures that directly involved reaction time or reaction time variance. In these analyses only trials with 150ms > RT < 1000ms were included to remove outliers which occasionally arose due to participants becoming inattentive or pausing within a block. These trials were removed because they might otherwise have obscured the learning related evolution of changes in RT and derived measures. For analyses that involved RT or related measures we excluded 0.8% of all trials for the mirror, 0.6% for the abrupt rotation, and 0.9% for the gradual rotation group.

3.4 Results

3.4.1 Reaction time changes under visuomotor transformations

We have previously shown that the reaction-time increase when learning a mirror-reversal was approximately 3 times larger compared to learning a 40° rotation (Chapter 2). Following the structural hypothesis, this is because the motor system is familiar with the structure of visual rotations but not of mirror-reversal. Thus, this hypothesis predicts that reaction time should increase more for mirror-reversal than for rotation learning. To test this idea, we forced participants to start their movements immediately after the target had appeared. To allow for a fair comparison between mirror-reversal and 180° rotation, we only compared trials in which both the stimulus and the required movement were identical under the two mappings, namely the -90° and the +90° targets. For these targets the required feedforward command was exactly opposite to the target direction, both for the mirror and the abrupt rotation condition. During the first block of mirror-reversal and abrupt 180° rotation-learning the reaction times increased on average by 189ms ±46ms ($t_{(6)}$ = -4.126, p = .006) and 190ms ±44ms ($t_{(6)}$ = -4.303, p = .005) respectively (fig 3.2). Thus, contrary to the prediction of the structural hypothesis, we found no significant difference in the change from the last block of normal to the first block of perturbed reaching between the two groups ($t_{(12)}$ = .017, p = .987).

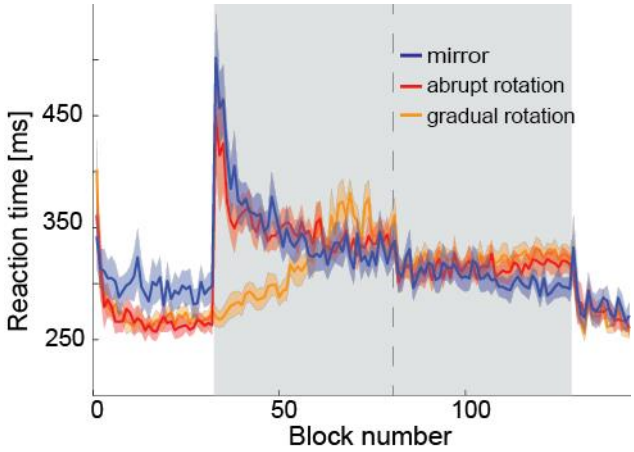

Figure 3.2: Reaction time for -90° and +90° targets.

White background indicates reaching under normal visual feedback, while the grey background indicates reaching during mirror reversed or rotated visual feedback. The dashed line indicates the break between sessions. Blue= mirror-reversal group mean (±SE), red= abrupt rotation, orange= gradual rotation.

3.4.2 Time-accuracy trade-offs in signed velocity

To test whether trade-offs between time and accuracy arise in the three conditions, we averaged the sign of the velocity for movements to the -90° and +90° targets and plotted it relative to target presentation. The sign of the velocity will henceforth also be referred to as the movement direction. Using the direction {+1, -1} sacrifices information about velocity magnitude but fully conserves information about the proportion of trials that moved in the target or anti-target direction, where a value of 0 means that movements in half of the trials are executed in the target direction, while the other half is directed in the anti-target direction. The movement direction started to differ from 0 at 125ms in both cases when reaching under normal visual feedback. For all further statistical analysis we used the average movement direction 150-200ms after target presentation. During the first day (blocks 33-80) of learning the movement direction was almost significantly positive (i.e. the hand moved towards and the cursor away from the target) in the abrupt

rotation group (fig 3.3a, mean .088 ±.037; $t_{(6)}$ = 2.407, p = .053) and significantly positive in the mirror-reversal group (fig 3.3b, mean .093 ±.035; $t_{(6)}$ = 2.653, p = .038). Thus, both groups started to move initially into the wrong direction, indicating that the early response was dictated by the old, original mapping. There was no significant difference between the two groups ($t_{(12)}$ = .097, p = .924) in the first day.

As practice continued during day 2, the movement stopped going significantly in the inappropriate direction (abrupt rotation: $t_{(6)}$ = 2.302, p = .061, mirror-reversal: $t_{(6)}$ = 1.858, p = .112). The gradual group, who had previously only been exposed to 180° rotations for the last 8 blocks of the first day, was still reaching into the wrong direction during the second day (fig 3.3c, $t_{(13)}$ = 4.099, p = .0013). Taken together mirror-reversal and abrupt 180° rotations elicited similar behaviors: Movements shortly after target presentation were still dominated by the old, no longer appropriate mapping, which led to movements in the wrong direction. Only after additional processing time could participants move in the correct direction in both conditions. We conclude that 180° rotations are learned in a very similar way as mirror-reversals. Both rely on the establishment of a new, time-intensive computation.

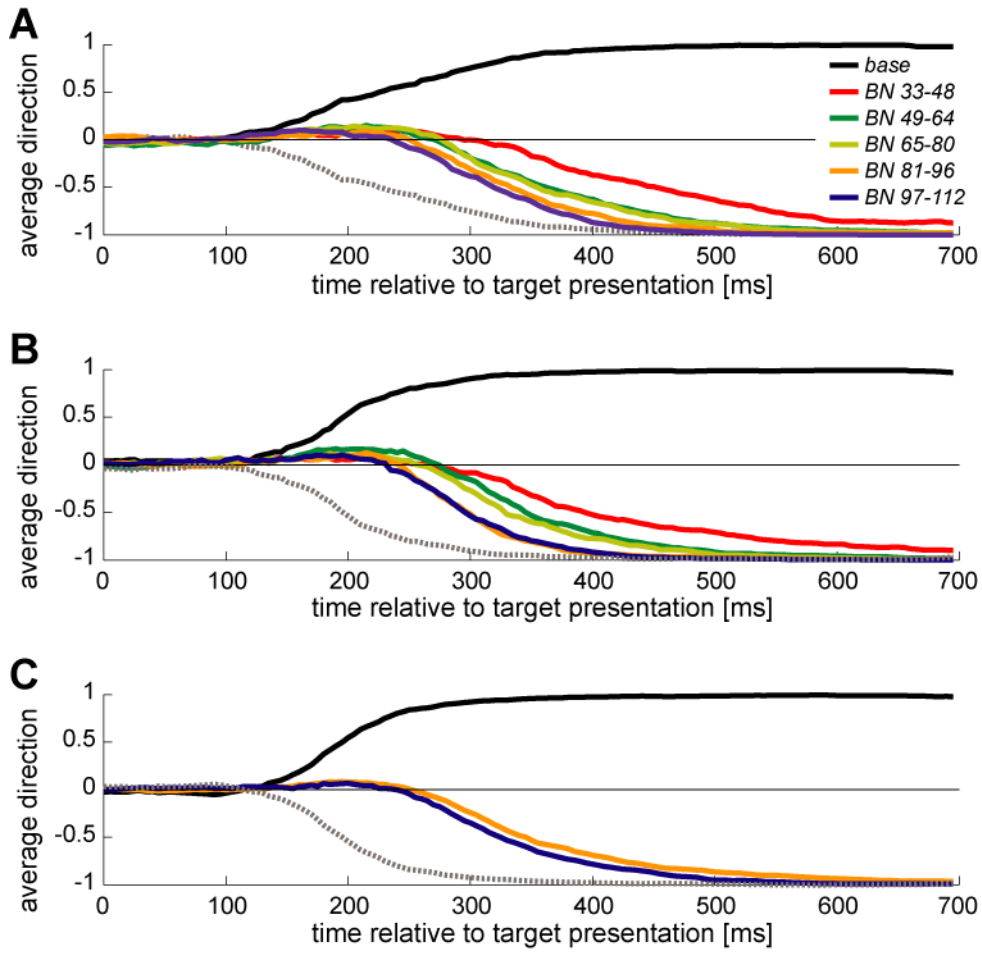

Figure 3.3: Average movement direction relative to -90° or 90° target presentation.

The group-average signed velocity in the X-direction is plotted relative to target presentation for different phases (each representing 36 blocks) of the experiment. Dotted line represents what a perfect reversal of the baseline velocity trace would look like. (**A**) Mirror-reversal (Cond. 1) (**B**) Abrupt rotation (Cond. 2) (**C**) Gradual rotation (Cond. 3). Note that for Condition 3 gradual rotations were still increasing throughout all phases of day 1, therefore only baseline and phases of day 2 are shown.

3.4.3 Time-Accuracy trade-offs in reaction times of mirror-reversal and 180° rotations

Time-accuracy trade-offs have previously been referred to as the key criterion to assess performance in finger sequence learning tasks (Reis et al., 2009) typically referred to under the umbrella term of skill learning. We argued that recalibration does not lead to time-accuracy trade-offs, because recalibration modifies an existing mapping and thus no additional computation is required. Since freshly formed control policies are not yet automatized, their computation requires additional processing time. Therefore, the tell-tale of de-novo control-policy acquisition is the emergence of time-accuracy trade-offs. If, as proposed in the second hypothesis, the absolute recalibration magnitude has an upper limit, then time-accuracy trade-offs should emerge at a certain rotation size. If however the size of the error signal is the limiting factor, as proposed in the third hypothesis, time-accuracy trade-offs should arise when visual errors approximate an upper limit. In Chapter 2, the required change during rotations was relatively small, i.e. 40° and hence we were able to study directional error. This is not possible in the current experiment, because the required change is 180° in the abrupt rotation and mirror-reversal for the -90° and +90° targets and at least in part 180° for the gradual rotation conditions. Thus we used absolute error, a measure that combines directional error and variance, to estimate potential trade-offs between time and accuracy. Briefly, for the mirror-reversal and the abrupt-rotation groups the trials for each of the lateral targets in each block are assigned to one of three bins depending on their reaction time. We then averaged across the -90° and +90° targets and across 8 consecutive blocks. For the gradual condition the procedure was identical except that movements towards all targets were included.

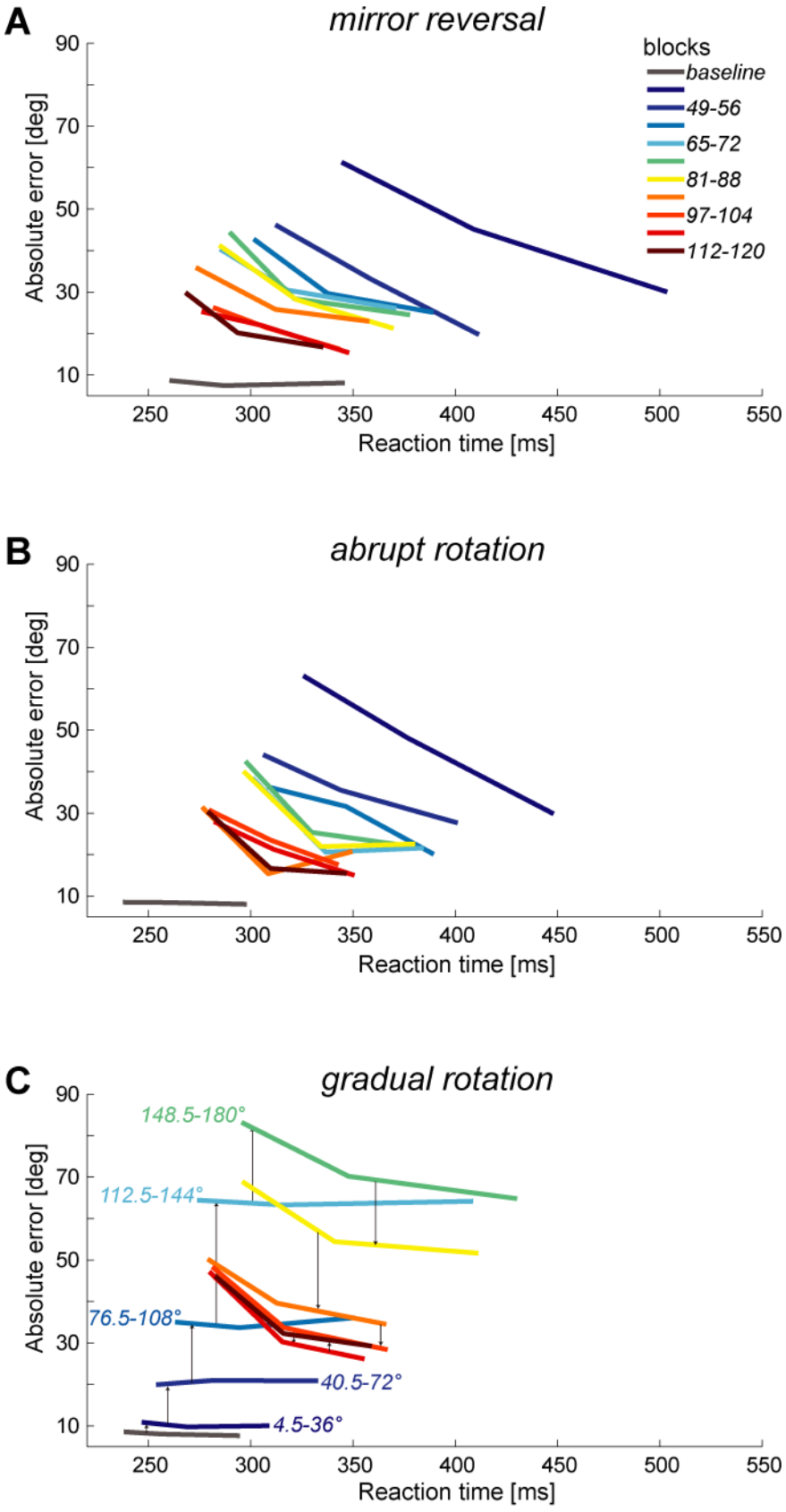
A
mirror reversal
blocks
baseline
49-56
65-72
81-88
97-104
112-120
Absolute error [deg]
Reaction time [ms]
B
abrupt rotation
Absolute error [deg]
Reaction time [ms]
C
gradual rotation
148.5-180°
112.5-144°
76.5-108°
40.5-72°
4.5-36°
Absolute error [deg]
Reaction time [ms]

Figure 3.4: Absolute error as a function of reaction time.

The trials were binned by reaction time (for the -90° and +90° targets in the mirror-reversal and the abrupt-rotation groups and all targets in the gradual-rotation group) for each participant and block. Each line is a group average of the lines calculated for 8 blocks and 14 participants. Visual feedback was veridical during baseline (grey). All other blocks were mirror reversed or rotated. The order of the line is indicated by a heat map from dark blue, through green to yellow, orange and red. (**A**) Mirror-reversal (Cond. 1). (**B**) Abrupt rotation (Cond. 2). (**C**) During gradual rotation (Cond. 3) absolute errors are initially small but increase over time and after a while decrease again. Therefore arrows are included to indicate the sequence of blocks. For phases where rotations still increased the range of rotations is explicitly stated in the figure.

Figure 3.4 shows that during baseline reaching there was no relationship between reaction time and reach error. However as soon as the manipulation of the visual feedback started, a strong trade-off between reaction time and adaptation evolved in both the mirror-reversal and in the abrupt rotation groups (Figure 3.4a, b): Longer reaction times led to smaller initial reaching errors in the first phase of the abrupt rotation ($t_{(6)} = 4.505$, $p = .004$). This effect was only marginally significant in the mirror-reversal group ($t_{(6)} = 1.998$, $p = .093$). Time-accuracy trade-off curves retained similar slopes throughout mirror-reversal learning. The trade-offs were visible in the mirror-reversal and the abrupt 180° rotation groups throughout both days (fig 3.4a, b). We had expected these trade-offs to arise during mirror-reversal learning based on the findings from Chapter 2. Thus in line with our previous analysis of movement velocity, we reconfirmed the existence of time-accuracy trade-offs in learning abrupt 180° rotations and mirror-reversal. The fact that large rotations show evidence for the establishment of a novel, time intensive computation refutes the structural hypothesis which states that learning rotations is different from learning mirror-reversal per se.

3.4.4 **Adaptation and variance during gradual rotation learning (Exp. 3)**

What causes the emergence of time-accuracy trade-offs? Is it the size of the rotation or the size of the visual error, i.e. the difference between adaptation and the rotation? Finally, what is the error or rotation size at which these time-accuracy trade-offs emerge?

To answer these questions we ran a group for which the visual rotation increased gradually over the course of blocks 9-72. We then tested the slopes of absolute errors over RT in each rotation block against baseline (Fig 3.4c). Note that until the rotation exceeded 144° there was no significant change in the trade-off between reaction time and visual error (table 3.3).

Blocks 9-16 vs. Blocks…	paired 2-tailed t-test	Rotation
33 – 40	t(13)=.321 p=.753	4.5 – 36°
41 – 48	t(13)=.444 p=.664	40.5 – 72°
49 – 56	t(13)=1.152 p=.270	76.5 – 108°
57 – 64	t(13)=1.129 p=.279	112.5 – 144°
65 – 72	t(13)=3.184 p=.007	148.5 – 180°
73 – 80	t(13)=5.142 p=.000	180°
81 – 88	t(13)=6.640 p=.000	180°
89 – 96	t(13)=6.115 p=.000	180°
97 – 104	t(13)=5.026 p=.000	180°
105 – 112	t(13)=5.863 p=.000	180°

Table 3.3: Slopes of absolute error over reaction time in the gradual condition.
The table shows t-tests of regression slopes between blocks with manipulated visual feedback and baseline (blocks 9-16) during different phases of the experiment.

In the gradual condition the visual rotation increased up to 180° at a pace that caused participants to increasingly fall behind. We expected participants to initially adapt to the visual rotation by recalibration of an existing internal model. However, we also predicted that at some point the visual error signal could not be used anymore for the purpose of recalibration. At this point, to compensate for the imposed rotation, there should be a behavioral switch from recalibration to control-policy acquisition. The establishment of a new internal model has been shown to be marked by initial increases in exploration (Roller et al., 2001; Cohen et al., 2005b). We thus tested whether variance in the movement direction and reaction time would peak around the same time as the switch from recalibration to control policy acquisition occurs. Figure 3.5a and b show data from the clockwise and the anticlockwise groups across all targets collapsed. On average the recalibration gain was too small for participants to catch up with the increasing size of the visual rotation in day 1. As adaptation continued to fall behind the ever increasing size of the rotation, there was a dramatic step-change in error reduction from block 65 to block 66. At this point adaptation increased from 77° to 107.5° (difference between block means 30.5° ±11.5). Subsequent blocks continued to show abnormally strong performance. When returning to normal reaching, participants showed an after-effect of initially 28° (±10.3) (t-test against zero $t_{(13)}$=-2.608, p= .011), which then decreased rapidly. Figure 3.5a shows the group average and Figure 3.5b the individual traces superimposed.

We examined the movement direction and reaction time variance within subjects and targets over the time course of gradual-rotation learning (Fig. 3.5c). At baseline the movement variance was 0.7 deg^2. During the introduction of the first 90° of rotation the average movement variance over these blocks increased by

1.07 ±0.2 deg^2 ($t_{(13)}$ = -5.281, p = .000). However, the increase between the change in movement variance from baseline to the first 90° of rotation was small compared to the more than 8-fold increase between 90° and 180° rotation ($t_{(13)}$ = 10.704, p = .000). Movement variance peaked at blocks 67 and 68, indicating that on average participants started to reach in a more exploratory fashion in these blocks or alternatively that the new control policy was still relatively unstable. Similarly reaction time variance did not significantly increase between baseline and the first 90° of rotation ($t_{(13)}$ = -1.096, p = .293), but showed a strong increase from 0°-90° to 90°-180° of rotation ($t_{(13)}$=3.695, p= .003). Reaction time variance peaked at blocks 65-66. Thus reaction time variance and movement variance increased roughly at similar points in time. Crucially, the increases in variation co-occurred with the onset of time-accuracy trade-offs as shown above (Table 3.3), indicating that at this point there was a qualitative change in behavior.

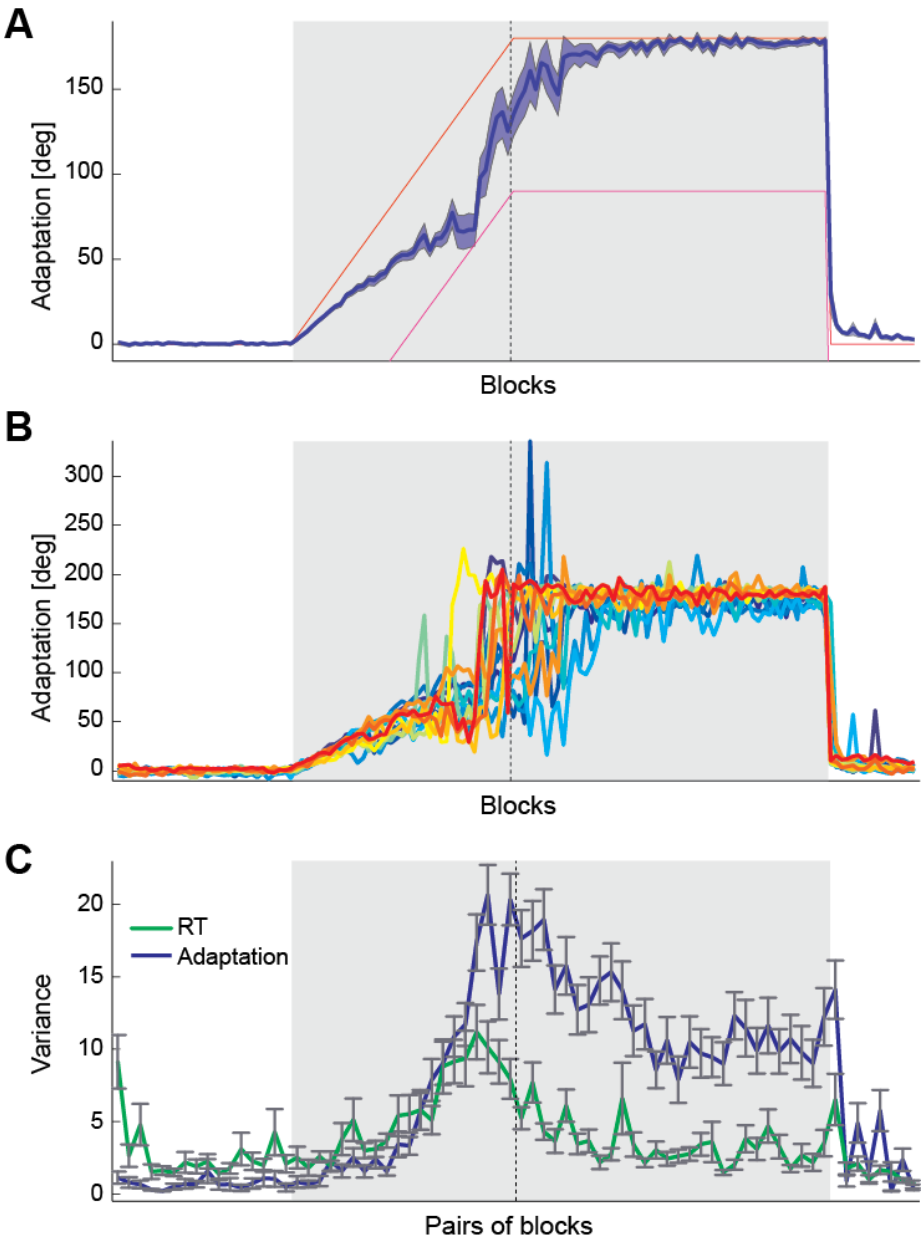

Figure 3.5: Time course of learning in the gradual rotation group.

Blocks of the experiment where visual feedback was rotated are shown against a grey shaded background. The vertical line indicates when the gradual rotation reached 180° and did not increase any further. (**A**) The group-average and SE of adaptation is plotted over the different blocks of the experiment. The red line indicates the absolute size of the rotation, whereas the pink line is the red curve shifted by 90°. It indicates what adaptation should look like to produce an absolute error of 90°. (**B**) Adaptation of individual participants over blocks. (**C**) Average within subject variance for reaction time (green) and adaptation (blue) estimated over pairs of blocks. Note that reaction time variance was scaled down by a factor of 1000 to allow for easier visualization.

3.4.5 Estimating the change point and learning rate during gradual rotations (Condition 3)

We have suggested two alternative explanations for the sudden behavioral change: The adaptation-limit hypothesis states that there is an absolute upper bound that limits how far the system can adapt. An upper bound on adaptation predicts that strong learners, who reach this bound after fewer trials than weak learners, should switch from recalibration to acquisition at an earlier time point. The alternative error-limit hypothesis states that there is an absolute upper limit on the size of the error signal that can be utilized by recalibration. It makes the opposite prediction: Weak learners should fall behind more quickly and reach the error limit after fewer trials than strong learners, who therefore should switch from recalibration to control policy acquisition only later in the experiment.

The accurate detection of this change point was therefore crucial to our further analysis. To better quantify the abrupt sudden change in behavior, we used a change-point model. Before the change point, the model recalibrated its output based on the last error, using a simple state space equation. After the change the model strategically switched to the correct answer, producing zero error on average. We then found for each participant the trial block that was the most likely candidate of the change point.

We ran two versions of this model, one of which (the standard model) learned from angular errors (Model 3.1, (Thoroughman and Shadmehr, 2000; Diedrichsen, 2007; Tseng et al., 2007) and the other model (Model 3.2) incorporated the notion that during recalibration the error signal is devalued as it approaches 90°. The latter model therefore did not learn from the angular visual error signal but instead for errors<=90° the cosine of the error was used as a teaching signal and for

errors>90° the teaching signal was set to 0, such that no learning could take place. In both models, the output behavior was allowed to instantly switch from recalibration (the state space model) to producing reaching errors of 0°. Initially the retention factor A, learning rate B, initial state z_0 and the change point were estimated freely for all participants over the blocks, with each block representing 1 datum, to reduce noise levels. We then fixed the retention factor A to the mean value over all participants (normal model .688 (±.037), cosine model .898 (±.026)) and re-estimated the other parameters. The mean variance explained was similar between the cosine (r^2= .866) and the standard model (r^2=.876) ($t_{(26)}$ = -1.113, p = 0.286). Yet the cosine model provided a better fit in 9 out of the 14 participants. Moreover it was less variant across participants with regard to its parameter estimates of B (SE=.024) and the error size at the change points (SE=5.291) compared to the standard model (SE of B=.037; SE of change point= 5.704). Curiously we noted that in some participants the learning curve would reach a plateau or even show decreased adaptation immediately before the switch between the two behaviors occurred (Fig. 3.6a, b). This plateau or forgetting behavior can be viewed as a consequence of learning from devalued error signals, and can be predicted by the cosine model. The more the visual error signal approaches 90°, the less the error signal can be utilized for adaptation, until the point is reached where the amount of adaptation that is forgotten from one time point to the next is bigger than the amount that is learned. At this point, net adaptation should decrease. We thus decided to use the cosine model for determining individual learning rates and change points for all further statistics. Note that in both models the estimated change points are identical in 11 out of the 14 subjects and the remaining 3 are

predicted to occur later in the normal state space model. Figure 3.6 shows the adaptation data and fits from 3 typical participants.

With this change point detection model it was possible to dissociate between alternative hypotheses 2 (the type of learning depends on the size of the rotation) and 3 (the type of learning depends on the size of the visual error). In line with the error limit prediction there was a strong correlation between the learning rate B and the change point across participants (fig 3.6d) (r^2 = .565, p=.002), such that strong learners switched only later in the experiment compared to weak learners. This result suggests that the behavioral switch does not occur due to an absolute upper bound on adaptation size. To the contrary, the error limit hypothesis correctly predicted that participants who learned quicker had later change points.

Interestingly, the average change point was found to occur at around 84.22° (±5.29°) of visual error. Thus once the visual error was close to 90°, participants showed an abrupt change in reaching behavior. As outlined in the introduction, an error limit of slightly less than 90° is to be expected from a system that represents the mapping from visual errors to motor updates in a Euclidean coordinate frame.

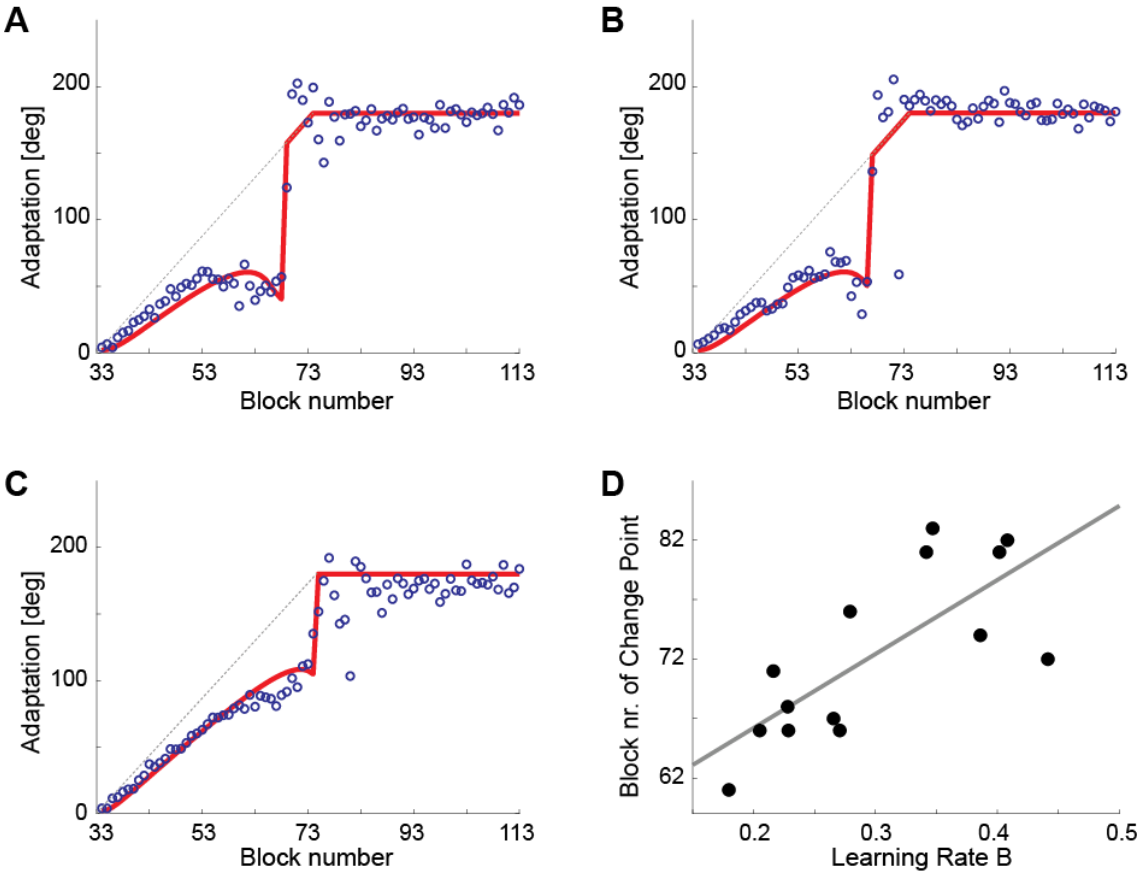

Figure 3.6: Learning rates and change points during gradual rotation learning.

Panels **(A-C)** show adaptation (blue circles) and model fits (red line) for three different participants. Grey stippled lines show the amount of adaptation that would be required to fully cancel out the rotation for any given block. (**A-B)** comparably weak learners. **(C)** Strong learner. **(D)** The estimated block numbers at which the change points occur are plotted over the learning rate B for all 14 participants. The lower a participant's learning rate B, the earlier the change point occurs on average.

3.4.6 **Consolidation**

We have previously observed that learning visual rotations of 40° results in forgetting between sessions, whereas learning mirror-reversals benefits from offline gains (Chapter 2). We have argued that recalibration leads to forgetting, while control policy acquisition benefits from offline gains. To obtain an unbiased performance measure despite time-accuracy trade-offs and reaction times that

decreased with learning, we fixed reaction times at 350ms and inferred the absolute error using linear regression. Neither learning mirror-reversal ($t_{(6)}$ = -.946, p = .380) nor abrupt 180° rotations ($t_{(6)}$ = 1.195, p = .277) resulted in significant performance increases from blocks 77-80 to blocks 81-84 and there was no significant difference between the consolidation between these groups ($t_{(12)}$ = 1.493, p = .161). The absence of offline gains in these two groups might be explained by ceiling effects. Indeed, the gradual rotation condition however benefitted from strong offline gains ($t_{(13)}$ = 3.118, p = .008), equaling a reduction of 20.3° (±6.5°) in absolute error (Fig. 3.7). Importantly all participants in this condition had switched from recalibration to acquisition already during the first day. Therefore taken together with the finding that there was no forgetting in the other two groups, our results are consistent with the claim that control policy acquisition benefits from offline gains.

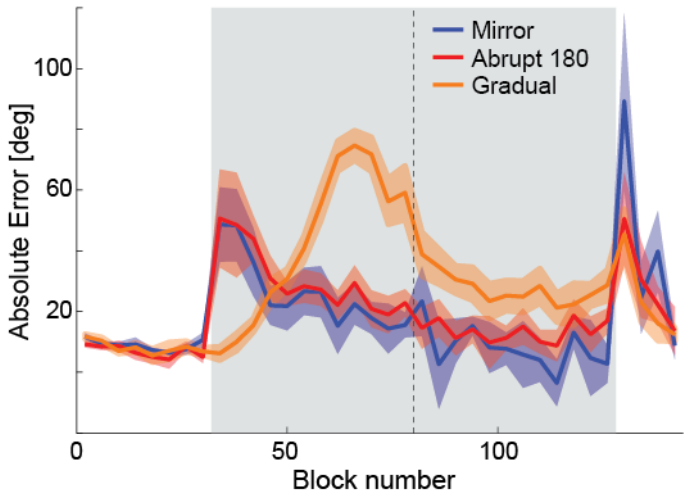

Figure 3.7: Absolute error decreases over the time course of training.

Grey background indicates training blocks under manipulated visual feedback. Dotted vertical line= break between sessions. Blue= mean ±SE of Mirror-reversal condition, red= abrupt 180° rotation condition and orange= gradual Rotation.

3.5 Discussion

In chapter 2 we argued that time-accuracy trade-offs during mirror-reversal learning were a signature of control policy acquisition and their absence during the learning of small visuomotor transformations was a sign of recalibration. Here we tested three alternative hypotheses that could have explained why different learning mechanisms are used. The structural hypothesis states that the motor system has a representation of the structure of visual rotations (Braun et al., 2009c, 2010), but less so of mirror-reversal. Conditions 1 & 2 have been designed to address this idea by direct comparison of learning of 180° rotations with mirror-reversal for those targets where the required visuomotor transformation was equivalent. Based on the presence of comparable time-accuracy trade-offs in both conditions, the structural hypothesis must be rejected and we conclude that 180° rotations and mirror-reversal are both learned by establishing a new time-intensive mapping between internal targets and motor outputs.

In chapter 2 we have also demonstrated offline gains for mirror-reversal learning, but offline forgetting for learning 40° visual rotations and attributed these to the underlying learning mechanisms. The literature is largely consistent with the finding that the gradual process that we term recalibration exhibits forgetting (Tong et al., 2002; Krakauer et al., 2005; Trempe and Proteau, 2010). In the current study we did not find offline gains in mirror-reversal or abrupt 180° rotation learning. Possible explanations for the absence of offline gains in the mirror-reversal and the abrupt rotation groups might be that the sample sizes of the two groups were relatively small. Furthermore performance seemed already to be close to the optimum in the end of day 1, which might have been caused by

having to rely on absolute error as the performance measure, a rather crude compound of reaching error and variance. However, the gradual rotation group which was still performing far from optimal in the end of day 1 expressed strong offline gains, suggesting that the absence of offline gains in the other two conditions might indeed have been caused by ceiling effects (Kuriyama et al., 2004; Wilhelm et al., 2012) and small sample sizes. Thus we conclude that learning of large rotations (Doyon et al., 2009b), just as mirror-reversal (Chapter 2) and other types of skills (Maquet et al., 2003; Stickgold, 2005; Wright et al., 2010; Abe et al., 2011) can benefit from offline gains.

Why does the type of learning used for 40° rotations differ from the type of learning used for 180° rotations and mirror-reversal? In line with previous suggestions, an upper boundary to the overall size of the imposed rotation that the motor system can adapt to using a "gradual" process was addressed in the recalibration limit hypothesis (Abeele and Bock, 2001a). In the third hypothesis we proposed that the process which is typically modelled using a state space model devalues the error signal with increasing error sizes. We implemented the error devaluation by applying the cosine to the visual error. The specific limit of 90° is predicted by a system that represents the mapping between visual inputs and motor outputs in a Euclidean coordinate frame (Fig. 3.1). Although the experiments presented here were not designed to test if the mapping is represented in Euclidean coordinates, as this would have required the study of movement magnitude, the cosine devaluation would be predicted by a Euclidean sensorimotor map. We have tested the general idea of a devalued error signal in Condition 3, where the error gradually increased at a pace that caused adaptation to increasingly fall behind. In line with the cosine model prediction, the estimated

change points from recalibration to strategic reaching occurred just before the average error would have reached 90°. While until this point there was no time-accuracy trade-off, the sudden change in behavior was accompanied by marked time-accuracy trade-offs that remained throughout the rest of the experiment. This indicates that indeed there was a qualitative behavioral change from recalibration to a mechanism that involves additional time-intensive computations.

Based on these results alone it would still have been impossible to attribute the behavioral switch to either the recalibration-limit or the error-limit hypothesis. However, our results clearly confirmed the prediction of the error-limit hypothesis, which states that the type of learning depends on the angular size of the error signal: We found that participants who learned slower and thus experienced larger error signals earlier on in the experiment also had earlier change points than strong learners. Based on these data we could reject the recalibration-limit hypothesis which would have predicted exactly the opposite relationship. In conclusion, our data strongly argue that the size of the error signal is devalued, and that therefore the angular size of the visual error signal determines the type of learning used. This model also relates to a number of findings in the literature. For example there is a long standing debate about whether learning gradually-increasing rotations differs from learning abrupt rotations (Kagerer et al., 1997; Criscimagna-Hemminger et al., 2010). It has been shown that when a 90° rotation was learned abruptly, participants performed worse by the end of training compared to another condition in which the rotation was introduced gradually (Kagerer et al., 1997). In line with these findings our framework predicts that sufficiently large gradual and abrupt rotations should be learned by different mechanisms. In the same study it was also shown that the abrupt condition

exhibited smaller after-effects after training, which is consistent with our idea that a new control policy is established that can be activated and deactivated in a context sensitive manner. One might even speculate that with sufficient training, the nervous system could be able to switch between the old and the new control policy in a similar fashion as a pianist switches from one song to the next.

Abeele and Bock found that when rotation size increased the movement error peaked at 120°, whereas during decreases from 180° to 0°, the movement error peaked at 70° of rotation (Abeele and Bock, 2001a). They interpreted these results as evidence for different change points during rotation increase and rotation decrease adaptation. Based on the data in the current experiment, an alternative explanation might be that participants acquire or already partly possess an internal model when abruptly being exposed to visual rotations of 180°. It seems plausible that (as we will examine in Chapter 4 in more detail) a newly learned internal model should itself be amenable to recalibration. If as in our third condition participants were not adapting as fast as the rotation size in- or decreased, one might speculate that the perceived size of the error signal was close to 90° in the decreasing and the increasing conditions.

In conclusion we have presented strong evidence that not the amount of adaptation, but instead the size of the angular error dictates the type of learning being used. While errors <90° can be learned by recalibration, errors> 90° elicit the establishment of a new time intensive computation.

Chapter 4
Can error-based learning be relearned?

4.1 Abstract

During mirror-reversal, the correction vector points in the wrong direction when using trial-by-trial recalibration. Here we address whether this mapping is fixed or whether it can be altered through experience. Recalibration with mirror reversed error signals requires a mirror reversed internal model. Participants made reaching movements under mirror reversed visual feedback over 4 days. To test how participants corrected for perturbations from one trial to the next, during half of the trials, we added small random rotations to the visual hand representation (cursor). Corrections derived from an unreversed model should increase the error. Once the error becomes too big, some form of strategic re-aiming is to be expected. We fitted a mixture of Gaussians consisting of an error-based (state space) component and a component for strategic changes. Without a good model for the strategic movements, a stationary distribution with a large variance was used.

Initially the learning gradient pointed in the wrong direction, leading to increasing deviation from the optimal movement direction and movements predominantly originated from a process that could not be adequately described by a simple state space model. Subsequently, the learning gradient decreased and showed weak indications of reversal. Simultaneously, the amount of trials that could be characterized by the state space model increased. Crucially, participants who produced more outlier movements during early mirror-reversal learning also had higher reaction time increases in a separate set of mirror reversed trials, suggesting that the outlier distribution in trial-by-trial corrections might be related to control-policy acquisition in feedforward control.

4.2 Introduction

To achieve the flexibility and precision that is required to finely manipulate its ever changing environment, the human motor system constantly recalibrates the mapping from sensory inputs to motor outputs (Tseng et al., 2007). Recalibration or adaptation is a form of error-based learning. Whenever the system perceives a prediction error - i.e. a mismatch between the actual outcome and the predicted outcome for a motor command – it updates its sensorimotor map (Tseng et al., 2007; Wei and Körding, 2009; Shadmehr et al., 2010). Note that the error signal is not just used as a reward prediction error but instead the directional information contained in the error signal is utilized (Tseng et al., 2007; Srimal et al., 2008). Recalibration has been demonstrated for many different movement types such as eye movements (McLaughlin, 1967), gate (Reisman et al., 2007) and reaching movements (Thoroughman and Shadmehr, 2000; Mazzoni and Krakauer, 2006; Tseng et al., 2007).

But how is a visual error signal used to adjust subsequent motor commands? One possible solution would be to infer the required update from the current inverse model (Kawato and Gomi, 1992). For example when applying a translation to the visual hand representation during a reaching movement task, the inverse model can generate an adequate update in the same way as it generates feedforward or feedback motor commands. In Chapters 2 and 3 we have shown that mirror-reversal and rotations in excess of 90° are learned differently from smaller visual rotations. We have attributed the switch from control policy recalibration to control policy acquisition to the nervous system's inability to generate adequate updates in response to large rotations and mirror-reversal. Under normal visual feedback, a

rightward error should be compensated by moving a bit further to the left in the following trial. In contrast, in a mirror reversed environment the update needs to be inverted, such that the hand moves even further to the left when experiencing a leftward error. Therefore in chapter 4 we tested explicitly if the nervous system can invert the way it learns from errors when confronted with mirror-reversed visual feedback over 4 training sessions. We also tested if trial-by-trial learning in a mirror-reversed environment is related to feedforward control, as would be predicted the updates were inferred from the same inverse model.

State space models have – despite their apparent simplicity - been used very successfully to quantitatively simulate trial-by-trial recalibration (Thoroughman and Shadmehr, 2000). In a nutshell, the motor output can be described as an internal state (z), combined with output noise ε. If the hand travels too far to the right of the target, a rightward error signal (error) is perceived. The mapping is updated by a fraction of the error signal, such that in the following movement, it will reach further to the left. From a computational standpoint the mapping from visual error signals to the necessary change in the motor command could be calculated by the current inverse model. In control theory the inverse model computes the feedforward motor command, but also needs to know how to change the motor command in response to visual error signals. Thus, the inverse model possesses knowledge of $\frac{\partial motor\ command}{\partial visual\ error}$, i.e. the Jacobian matrix $\begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_j}{\partial y_1} & \cdots & \frac{\partial x_j}{\partial y_k} \end{pmatrix}$, that tells the system how to change the motor command (x) to minimize the visual error ($\hat{y} - y$).

Correspondingly state updates during parametric recalibration require knowledge about the differential $\frac{\partial motor\ update}{\partial visual\ error}$ to choose a motor update that decreases the size of the visual error in the following trial.

If we assume that the system reacts to visual errors first and foremost as if they were rotations (Turnham, 2011), the Jacobian of interest now becomes the derivative of the reach direction in respect to the angular visual error – that is simply a scalar. That is, the learning rate *B* (equation 4.1) is simply the learning gradient.

In the current experiment participants made reaching movements throughout 4 sessions with mirror reversed cursor feedback. Unbeknownst to the participants we added small random rotations onto the cursor on top of the mirror-reversal. Crucially when visual errors are corrected for in a mirror-reversed environment, a non-reversed learning gradient will amplify the error in the next trial. We therefore expected that the way in which the motor system learns should change during mirror-reversal.

By modelling how the system corrected for random visual errors with a state space model, we were able to infer whether the nervous system relearned to learn from errors. One theoretical possibility was that there would be no reversal of the learning gradient (*B*). In this case the hand should over the course of the experiment either have drifted out further and further until it crossed the mirror-reversal axis again or until some other equilibrium was reached as a cosine update rule would suggest (Chapter 3). Alternatively the learning gradient *B* could have been inverted, leading to error-based learning in a mirror-reversed fashion. Another possibility was that error-based learning would become suppressed, such that trial-by-trial learning stopped altogether.

4.3 Methods

4.3.1 Participants

16 healthy participants (6 males; mean age 23.1 ± 2.7 years) participated in the experiment, all of which were right handed according to the Edinburgh handedness questionnaire (Oldfield, 1971); median handedness score = 0.84; IQR = 0.34). Participants were instructed not to drink alcohol during the 4 days of training and to have a minimum of 6 hours of sleep per night. The study was approved by the UCL ethics committee. All participants gave written informed consent and were remunerated with £25.

4.3.2 Apparatus

Participants controlled a robotic manipulandum (Fig. 4.1) that restricted movements along a two-dimensional horizontal plane and sampled their hand position at 200Hz. Torque motors that actuated the robotic arm were used to guide the participant's hand to the start location. Reaching movements were made underneath a mirror, which prevented participants from viewing their hand and arm. The mirror reflected the image of the monitor placed above, resulting in a superposition of the screen below the mirror. Hand position was represented by a 3mm diameter circle and the screen had a lag time of 68 ± 5ms. Participants sat on a chair and rested their foreheads on a pad placed 10-15cm above the mirror to keep the viewing position constant.

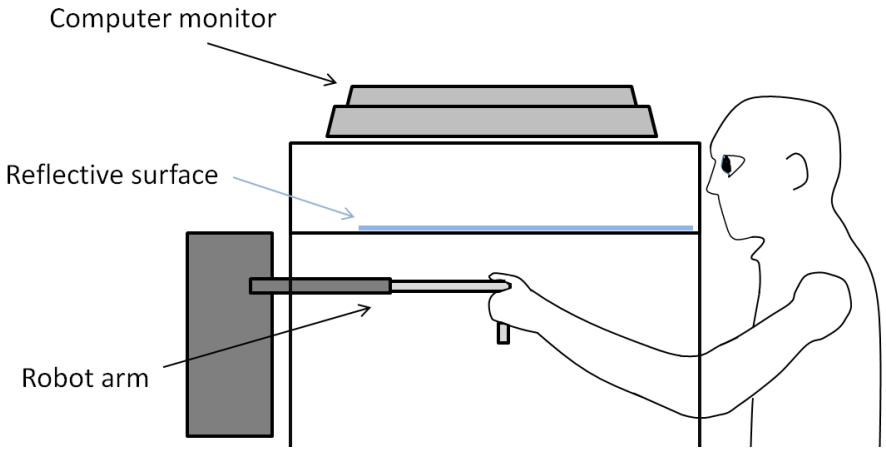

Figure 4.1: Virtual reality environment setup

Participants sat in front of a virtual reality setup and performed reaching movements while holding on to a robotic arm. Vision of the hand was occluded by a reflective surface (mirror) that presented the visual environment.

4.3.3 **Paradigm overview**

4.3.3.1 Training trials

Training trials were designed to train the mapping between motor commands (hidden hand movements) and the cursor on the screen. The task required participants to terminate their movement inside the target rectangle while being given full visual feedback. Targets were randomly presented in one of 6 locations, which were located 15 cm from the start location at 200°, -20°, 0°, 20°, 160° and 180° relative to the midline (Fig. 4.2). Targets were symmetrically distributed around the start location to prevent anticipatory movements.

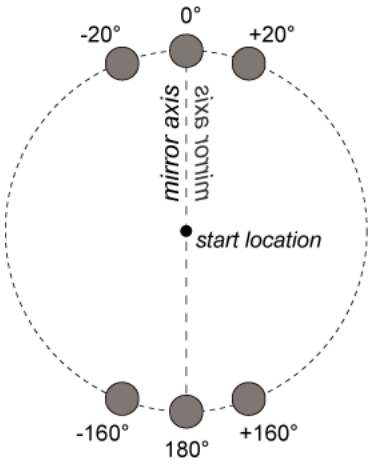

Figure 4.2: Target arrangement in training trials.

The dashed vertical line indicates the mirror-reversal axis. During testing trials only the 0° target was presented.

4.3.3.2 Testing trials

Testing trials were used to estimate the learning gradient (*B*), reflecting how participants updated reaching movements based on previous errors. In these trials the target was always presented at 0°. During these trials, we introduced random cursor rotations (without the participants' awareness) so that the state space model could be estimated on a trial-by-trial basis. Cursor rotations were drawn from a pseudo-random Gaussian distribution with a standard deviation of 7°. In order to encourage the production of straight reaching movements without any feedback corrections, we instructed participants to reach through the target and only presented end point feedback where the distance between the instructed start location and cursor exceeded 15cm.

4.3.4 Performance feedback

Symbolic feedback was given to participants to indicate whether their movement velocity or reaction time (RT) were within the thresholds. Movement onset and termination were defined respectively as when the movement velocity exceeded 2.5cm/s and when it remained below 2.5cm/s for 40ms. RT was defined

as the time between target appearance and movement onset. For RTs longer than 382ms or shorter than 32ms (indicating anticipatory movements), an unpleasant (punishing) low or high pitch tone respectively, was presented. Movement time was defined as the time between movement onset and termination. The maximum allowed movement time was 1200ms. The lower and upper thresholds for peak velocity for training and testing trials, respectively, were 50 and 80cm/s and 85 and 125cm/s. The faster velocities for testing trials were designed to encourage participants to make straight path reaches, which are more suitable for modelling with a state space model. Peak movement velocity feedback was provided in the form of color changes of the target and cursor: yellow signalled movement speeds that were too fast. Blue signalled that the movement speed was too slow or that the movement time –the interval from movement onset to termination – was too long. When the movement parameters were within the required bounds and the cursor accurately reached the target, the target turned red and expanded, a pleasant sound was produced, and 3 points were awarded. When participants were within all thresholds but missed the target, the target turned green and they received 1 point. The total number of points for each block was displayed in all 4 corners of the screen and presented at the end of each block.

4.3.5 **Experimental Design**

Each block consisted of 30 training trials followed by 30 testing trials. The testing trials enabled us to estimate changes in the learning gradient, while training trials, where the targets were displayed to the left or the right of the mirror-reversal axis, enabled us to assess feedforward control in the same block. The transition from training to testing trials was indicated by a message on the screen. At the beginning

of each trial, the manipulandum guided the participant's hand to the start location located 25-30cm in front of the participant's chests. A variable delay (uniformly distributed between 0-500ms) was added at the start of each trial to prevent participants from moving early in anticipation of the visual presentation of the target. During mirror-reversed reaching, visual feedback was flipped across the central midline.

On the first day, participants were given 5 baseline blocks with normal visual feedback (blocks 1-5), followed by 10 blocks of mirrored visual feedback (blocks 6-15) (Fig 4.3). On the second and third days participants were given 10 blocks of mirrored feedback (block 16-25, 26-35). On the fourth day participants were given 10 blocks of mirrored feedback (blocks 36-45) followed by 5 washout blocks with normal feedback.

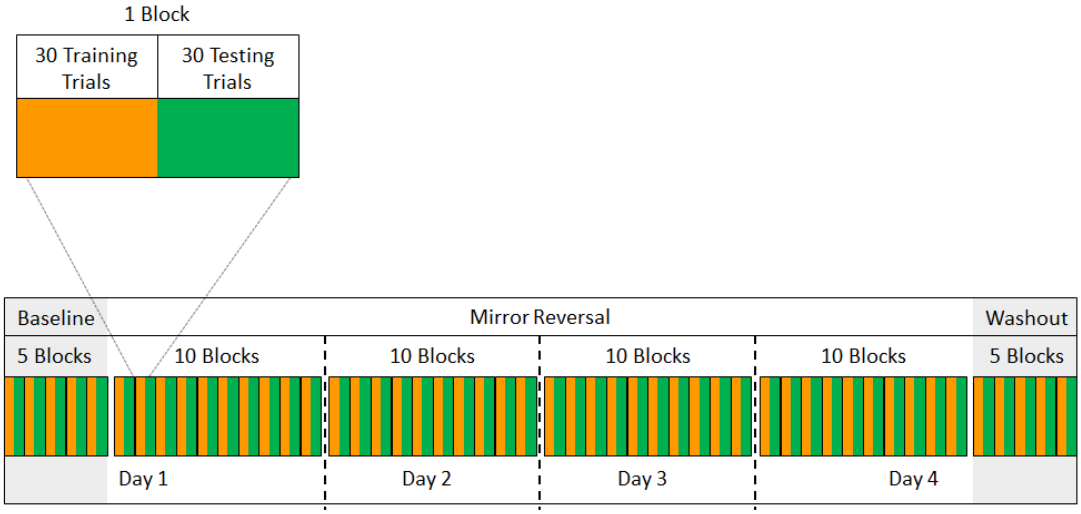

Figure 4.3: The structure of the experiment.

Each block contained 30 training trials followed by 30 testing trials. The experiment began and ended with 5 blocks with normal feedback. For four consecutive days participants performed 10 blocks of mirror-reversed reaching.

4.3.6 **Modelling**

All modelling and analysis was performed using custom written Matlab™ code.

4.3.6.1 The State Space Model

In order to estimate the learning gradient (B), we fitted a state space model (Thoroughman and Shadmehr, 2000) to each participant's hand position data for testing trials. At the core of the state space model is its adaptive state (z_n) that represents the system's current estimate of the mapping between movement goals and motor commands. This adaptive state is used to determine what motor commands are sent to the muscles to reach the target. This term is constantly updated on a trial-by-trial basis according to the visual error (e_n) of the previous movement. The amount by which the state is updated by error feedback is given by the learning gradient (B_v) in block v. The retention factor (A) determines how much of the current state will be retained in the next state (z_{n+1}). This behavior can be mathematically described as follows:

Equation 4.1

$$z_{n+1} = A * z_n - B_v * e_n$$

Although hand movements will naturally produce errors due to motor noise, it is computationally impossible to estimate trial-by-trial learning rates based on these errors alone as the model cannot differentiate between actual updates of the adaptive state and motor noise. Therefore, in order to estimate the learning rate, we introduced artificial errors in the form of cursor rotations (u_n). The cursor position (y_n) is a combination of the adaptive state (z_n) plus a noise term (ε_n) which is assumed to be normally distributed with a standard deviation of σ. An additional factor (m_v) specified whether the visual feedback was normal (m_v=1) or whether it

was mirror reversed (m_v=-1). The cursor position ($\hat{Y}_n$) that was predicted by the state space model was calculated as:

Equation 4.2

$$\hat{Y}_n = (z_n + \varepsilon_n) * m_v + u_n$$

The error term (e_n) that is used to update the internal state (z_{n+1}) was described as follows:

Equation 4.3

$$e_n = target - \hat{Y}_n$$

The state space model estimated parameters A, B and z_0 (the initial state estimate) by minimising the total sum of squares of the difference between the output predicted by our model ($\hat{Y}_n$) and the actual cursor angles that were produced by the participants (y_n):

Equation 4.4

$$Model\ prediction\ error = \sum (\hat{Y}_n - y_n)^2$$

4.3.6.2 Preprocessing

Previous studies have shown that motor adaptation can be modelled by two simultaneous 'slow' and 'fast' adaptive processes (Smith et al., 2006; Turnham et al., 2012). Since our study was designed to exclusively examine fast trial-by-trial adaptation, we subtracted out the mean slope from the raw hand positions of each block before fitting our model, assuming that in this way we would be able to isolate the fast learning system.

4.3.6.3 Mixture Model

When correcting for errors under mirror-reversed feedback, there are two possible ways to succeed: changing the way of automatically adapting to errors, or using certain strategies (e.g. to mentally shift the target location to compensate for the adaptation, or simply trying to move in a direction that feels proprioceptively straight). Lacking a precise model for such behaviors, we simply assumed that movements that were not well fit by the state space model (i.e. outliers) were more likely generated by such a strategic mechanism. We therefore modelled all movements as originating from one of two Gaussian distributions: a model-based distribution (equation 4.5) or an outlier distribution (equation 4.6) (Fig. 4.4).

Equation 4.5

$$Likelihood(y_n \mid Modelbased) = \frac{1}{\sigma_{Modelbased}\sqrt{2\pi}} e^{-\frac{(y_n - \hat{Y}_n)^2}{2(\sigma_{Modelbased})^2}}$$

Equation 4.6

$$Likelihood(y_n \mid Outlier) = \frac{1}{\sigma_{Outlier}\sqrt{2\pi}} e^{-\frac{{y_n}^2}{2(\sigma_{Outlier})^2}}$$

We define model-based movements as those which more likely originated from the state space model and the outlier movements as those which more likely originated from the outlier distribution. The variance of the model-based distribution ($\sigma^2_{Modelbased}$) was estimated on an individual participant basis from the motor variance during the baseline blocks (based on the assumption that all movements made in baseline blocks were from the model-based distribution). The outlier distribution was designed to account for movements in which subjects may have used a different process (e.g. aiming towards an imaginary target or making exploratory movements). For this distribution, we freely estimated the variance

($\sigma^2_{Outlier}$) for each block, assuming that the distribution was centred on the target. As a result of modelling the data as originating from two distributions, we obtained the parameter $P_{Modelbased}$, the overall probability of making a model-based movement. Therefore, for each given trial we could calculate the posterior probability that it originated from either the model-based or outlier distribution. Finally, to estimate the free parameters we needed to maximize the marginal likelihood (ML):

Equation 4.7

$$ML = \sum_{n=1}^{N} P_{Modelbased} * Likelihood_{Modelbased(n)} + \left((1 - P_{Modelbased}) * Likelihood_{Outlier(n)}\right)$$

The model minimized the negative summed log likelihood of the two distributions, giving rise to the data observed in each block and each subject by simultaneously estimating the free parameters.

Equation 4.8

$$Negative\ Marginal\ LogLikelihood = -\ln(ML)$$

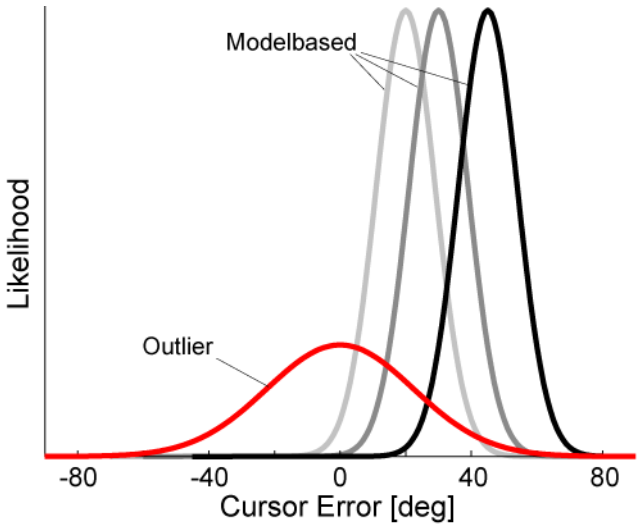

Figure 4.4: The 2 distributions in the Gaussian mixture model.

A simulated example of the model-based distribution for 3 different trials and the outlier distribution. The variance of the model-based distribution was estimated from baseline blocks and its center was identical with the state z of the state space model. Thus, the distribution was allowed to change its center as the state space model changed its state from one trial to the next. In contrast, the variance of the outlier distribution was estimated freely in each block, while its center was always stationary at zero.

4.3.6.4 Gaussian mixture model fitting

The model was only fitted on data from the testing trials. On the data from the baseline blocks, we freely estimated A, B, z_0 and $\sigma^2_{Modelbased}$ using the state space model (equation 4.1). For each subject, we used their average A and $\sigma^2_{Modelbased}$ as a constant for the entire experiment. This was done because the parameters A and B were not independent from each other, and thereby assuming a constant A for each participant across all phases allowed us to estimate changes in B. Thus, after the initial fit, the free parameters B, z_0, $\sigma^2_{Outlier}$ and $P_{Modelbased}$, were estimated block-by-block throughout the whole experiment. The lower and upper bounds on their estimates were B -1 to 1; z_0 -20 to 20; $P_{Modelbased}$ 0 to 1; and $\sigma^2_{Outlier}$ 1 to 100.

4.3.7 **Reaction time versus error analysis in training trials**

In contrast to the testing trials where all trials were included in the data analysis, in the training trials movements with RT<50ms and RT>730ms were excluded from the analysis, which led to 99.1% of all training trials to be analysed. In order to evaluate the relationship of RT versus accuracy for lateral targets during training trials, we defined the initial reach angle as the angle of the vector between start location and the cursor position 150ms after movement onset. This allowed us to measure the feedforward motor command with relatively little influence from feedback corrections (Franklin and Wolpert, 2008). To average the relationships across the four lateral targets, a positive error was defined as being in the un-adapted (un-mirrored) direction. For each block, we assigned the trials to one of 6 bins depending on their RT. We then averaged the initial reach error for each bin across each day (10 blocks) and plotted the results over the average RT for each bin. For the training trials, we only analysed movements made towards lateral targets, as the targets located at 0° and 180° require the same movement for baseline and mirrored reaching.

4.4 Results

Participants made mirror reversed reaching movements over 4 sessions which took place over 4 consecutive days. Each block consisted of 30 training trials, during which reaching movements were aimed at 6 different targets and a subsequent testing phase consisting of 30 trials, during which movements had to be aimed at the 0° target. In the testing trials we applied small random rotations to the cursor feedback. Participants were forced at all times to keep their reaction times below 335ms.

4.4.1 Training trials

4.4.1.1 Reaction time

In line with previous studies (Chapter 2), there was an increase in reaction times during mirror-reversal learning (Fig 4.5). During training and testing, RTs significantly increased by 80ms ($t_{(15)}$ = 12.379, p=3*10^{-9}) and 16ms ($t_{(15)}$ = 3.730, p = .002) respectively when mirror-reversal was first introduced (blocks 1-5 vs. blocks 6-10). By the end of the first day (blocks 11-15), RTs were still significantly elevated ($t_{(15)}$ = -3.770, p = .002) in the training trials and only in the beginning of the second day did this difference disappear. RTs for testing trials did not significantly differ from baseline during the end of the first day ($t_{(15)}$ = -0.467, p = .647). Over the course of the experiment RTs further decreased for training and testing trials, while the RT of training trials stayed consistently higher than the RT of testing trials. The RT difference between training and testing trials is likely due to the presence of 6 target locations during training as opposed to 1 target location during testing. Furthermore the 0° target during testing did not require a change in

the feedforward command, because the straight-line trajectory from the start location to the target was identical with the mirror-reversal axis.

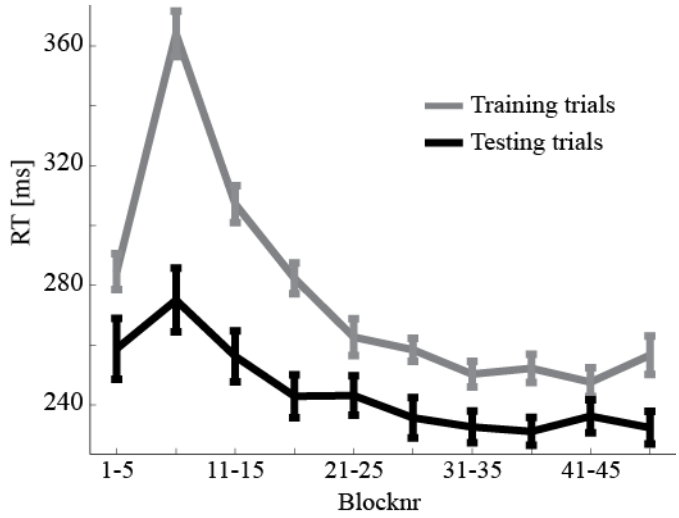

Figure 4.5: Reaction times increased when mirror-reversal started.

Reaction times for test (black) and training trials (grey) are plotted over groups of 5 blocks. Blocks 1-5 were normal baseline reaching and blocks 46-50 were washout. Blocks 6-45 were with mirror reversed visual feedback.

4.4.1.2 Time-accuracy trade-off

We analyzed the relationship between RT and error for movements towards the 4 lateral targets during training. Errors from the 20° and -160° targets were inverted and thereafter the trials within any block were assigned to 1 of 6 bins for each target separately according to their reaction time (see methods). Next these bins were averaged over all targets and 10 blocks (Fig. 4.6). During mirror-reversal on day 1, movements in the fastest bin were still directed in the incorrect, un-mirrored direction resulting in errors of approximately 25°. On average movements in the fastest and in the slowest bins under mirror-reversal on day 1 differed by up to 20°, depending on RT. Upon visual inspection the relationship between the first 5

bins of mirror-reversal during day 1 looks approximately linear, while the last bin is at baseline level of 5° and the slope flattens out. It is to be expected that errors can even at the longest RTs never be smaller than at baseline levels. Therefore, to compare the strength of the relationship between RT and error, all trials with RT>400ms were excluded from this and only this analysis: The slopes of error over RT were computed individually for each subject and phase. During baseline reaching, the slopes did not significantly differ from zero and thus there was no indication that error depended on RT ($t_{(15)}$ = 0.171, p = .867). However, during mirror reversed training on day 1 a strong RT-error relationship emerged compared to baseline ($t_{(15)}$ = -7.734, p = $6*10^{-7}$). In other words, trials with longer RTs had lower errors. Even on day 4 this relationship remained significantly different from baseline ($t_{(15)}$ = -4.320, p = .001). This is in line with findings from a previous experiment and has been interpreted as a tell-tale for the emergence of a novel, computationally expensive control policy (Chapter 2). Interestingly we also observed a reduction in the time-accuracy trade-off over the training days (repeated measures ANOVA over slopes ($F_{(3, 45)}$ =4.224, p=.01) leading to slightly smaller slopes on day 2 ($t_{(15)}$ = -1.964, p = .068) and day 3 ($t_{(15)}$ = 1.920, p = .074) and significantly smaller slopes on day 4 ($t_{(15)}$ = -4.320, p = .001) compared to day 1. This lends further support to the idea that during mirror-reversal learning a new control policy is first established and subsequently automatized, requiring increasingly less time for the same computation.

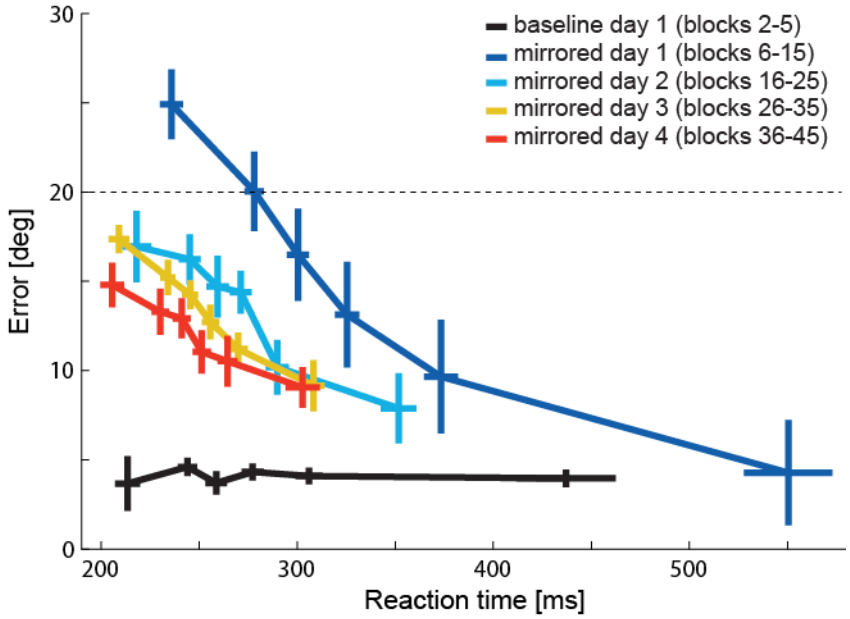

Figure 4.6: Speed-accuracy trade-offs in feedforward commands.

Averaged angular error recorded 150ms after movement onset is plotted over reaction time. Within each phase the trials were assigned to one of six bins, depending on their reaction time. Error bars represent SE of each bin for each of the two dimensions. The horizontal dashed line shows the mirror-reversal axis.

4.4.2 Testing trials

4.4.2.1 Average drift in reaching movements

For the testing trials towards the 0° target the straight line trajectory between start position and target is aligned with the mirror-reversal axis, such that when moving along this axis the mirror-reversal should not have any effect on the cursor position. However, during baseline, participants moved on average slightly to the left of the target: mean= -2.25°, SE=±0.31° (Fig. 4.7). Throughout the blocks 5-10 of mirror-reversal the reaching angles became increasingly more negative, i.e. they drifted to the left by a further -5.58° (±0.92°): $t_{(15)}$ = -6.038, p = 1.5*10^{-5}.

Although the peak was reached within blocks 6-10, during blocks 11-15 ($t_{(15)}$ = -4.149, p = .001) on day 1 and blocks 16-20 ($t_{(15)}$ = -2.691, p = .017 and blocks 21-25 ($t_{(15)}$ = -2.482, p = .025) on day 2 participants still produced significantly more negative reaching angles than during baseline blocks 1-5.

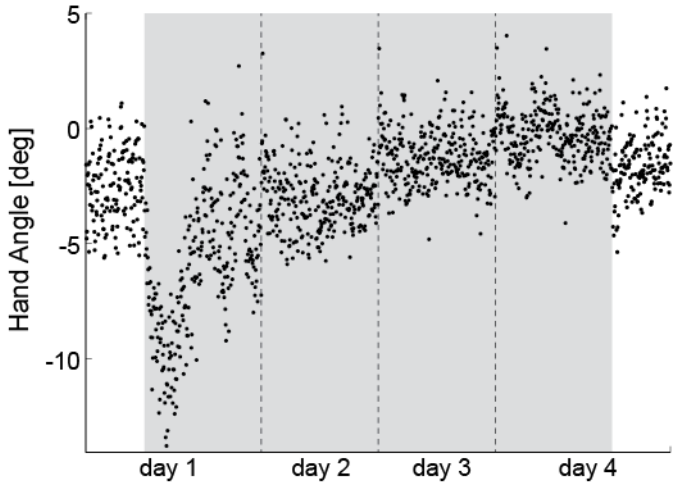

Figure 4.7: Drifts in average reach angle 15cm from the start during testing trials.

Grey background indicates mirror reversed visual feedback. Dotted vertical lines separate data collected on different days. Each dot represents a single reaching movement averaged across all participants.

Why did participants drift away while reaching towards a target for which the minimum distance trajectory during baseline and mirror-reversal was identical? Human reaching movements travel along curved trajectories due to biomechanical constraints and to optimize energetic expenditure (Todorov and Jordan, 2002). Furthermore, certain movement directions might encounter slightly more inertia by the robotic manipulandum than others. This might explain why with endpoint feedback reaching movements were biased to the left. If the motor system would not constantly recalibrate its reaching movements (Tseng et al., 2007; Srimal et al., 2008) trial-by-trial, participants would have moved even further to the left during

baseline. However, the usually beneficial recalibration becomes detrimental in a mirror-reversed environment. Under these circumstances any bias in the mean reach angle becomes amplified and even increases from one trial to the next, thus explaining the initial drift in the mean reaching angle. Interestingly, on the 3rd day the mean reach angle did not differ significantly from baseline. There are two alternative explanations: Either the nervous system had correctly inverted the way it learned from errors or, alternatively, error-based learning had stopped altogether and participants used other learning mechanisms to reduce the error.

4.4.2.2 Movements are generated by two different processes

To study recalibration in the mirror-reversal task, we injected small random rotations on top of the mirrored cursor during the testing trials. In this way we could observe how the motor system reacts to small perturbations. As corrections derived from an outdated internal model are detrimental, the motor system should change how it reacts to visual error signals under mirror reversed visual feedback. Figure 4.8 shows reaching movements during a single block of testing trials from an individual participant during mirror-reversal. While the majority of movements were relatively close to one another, testing trial 17 deviated strongly from all other trials.

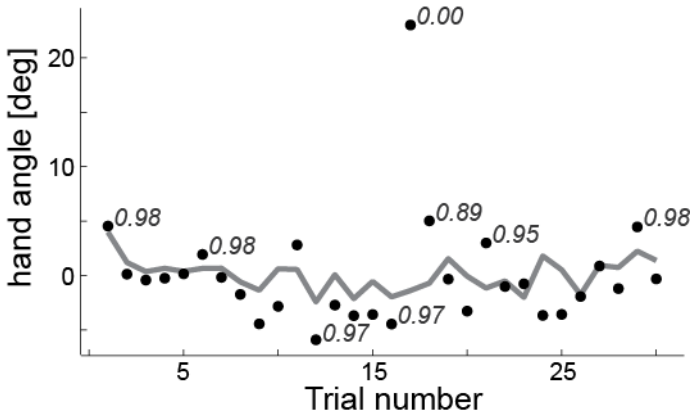

Figure 4.8: The probability of a trial to belong to the model-based distribution.

Data from the testing trials in block 21 in participant 3 are shown. The grey line shows the state space model predictions based on the estimated parameters. Each dot represents a single trial and the numbers indicate the probability for a given trial to have arisen from the state space model distribution.

Our intuition was that the data were produced by two different processes. One process is an error-based recalibration process, while a separate process generates movements that might be strategic or exploratory. We estimated the likelihood of each movement to stem from either the model or the outlier distribution. During baseline the average probability ($P_{Modelbased}$) for a given trial to originate from the model-based distribution was 0.948 (±0.043) (Fig. 4.9). However as soon as the cursor position was mirror-reversed the average probability for a trial to belong to the model-based distribution dropped to 0.416 (±0.273). Over the course of the experiment $P_{Modelbased}$ increased again so that it was significantly higher by the first half of day 3 (paired t-test blocks 6-10 vs. blocks 16-20: $t_{(15)} = -2.243$, $p = .04$). Despite the recovery, even after 4 days of mirror-reversed reaching $P_{Modelbased}$ was only 0.764 (±0.215) and still significantly smaller than at baseline ($t_{(15)} = 3.197$,

p = .006). During washout (blocks 46-50) $P_{Modelbased}$ was at 0.924(±.029) and did not differ significantly from baseline ($t_{(15)}$ = .825, p = .422). Taken together, the exposure to mirror-reversal caused a large proportion of movements to be generated by a process other than recalibration. However with further training the average probability of trials to originate from the model-based recalibration process increased.

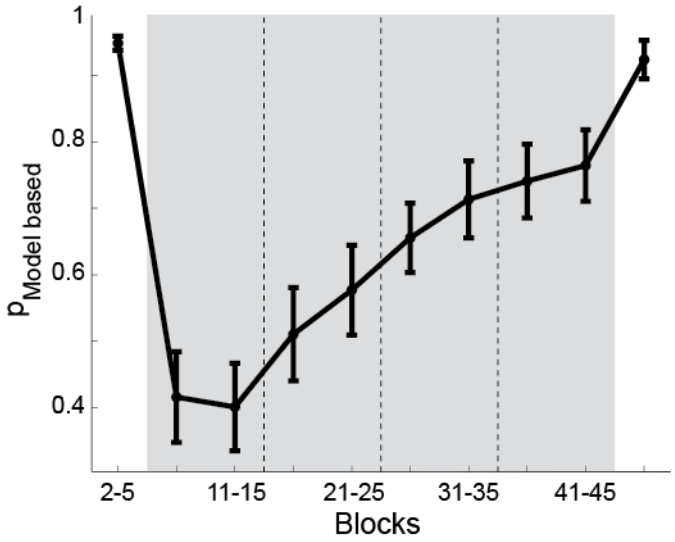

Figure 4.9: $P_{Modelbased}$ plotted over blocks.

Grey background indicates reaching under mirror reversed visual feedback. Shown is the mean ±SE of $P_{Modelbased}$ over groups of 5 blocks across participants. Dotted vertical lines separate data collected on different days.

If such a large proportion of movements cannot be explained by parametric recalibration which other process might have generated them instead? Decreases in $P_{Modelbased}$ likely reflect not just an increase in variance but instead a shift towards a different process. If a novel process would underlie the outlier distribution, it might require additional processing time, similar to a novel control policy in feedforward control. We thus correlated across participants the change in reaction time at the

onset of mirror-reversal (blocks 6-10) during training trials with $P_{Modelbased}$ during testing trials (Fig. 4.10). We found that participants whose movements were more likely to belong to the outlier distribution also had higher increases in reaction time during training trials (r^2=0.298, p= .0288). This correlation links the model estimates in testing trials with feedforward control in training trials and thereby supports the credibility of the method used for estimating these probabilities and the learning gradient B. Furthermore this link supports the idea that feedforward control and recalibration rely on shared internal models.

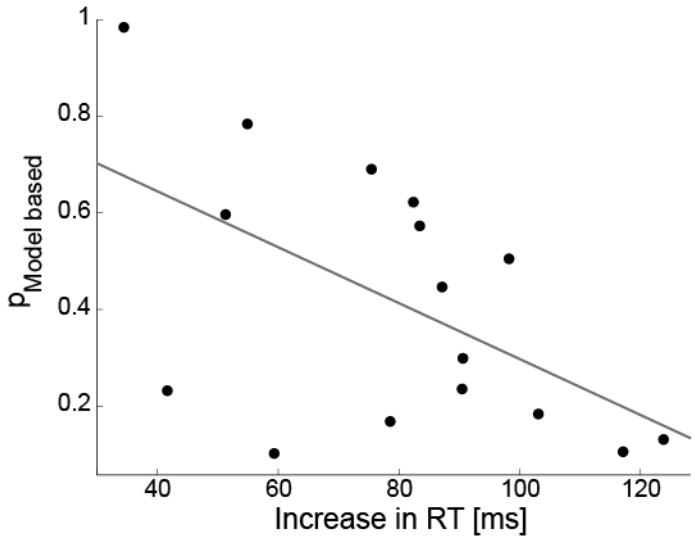

Figure 4.10: $P_{Modelbased}$ correlates with reaction time increase.

Each point represents one participant's average $P_{Modelbased}$ during the testing trials in blocks 6-10 plotted over the average increase in reaction time during the training trials of the same blocks.

4.4.2.3 Outlier movements correct differently for errors than model-based movements

Within any block the quality of the estimate of B was expected to depend on $P_{Modelbased}$. In blocks in which few trials could be explained by the state space model, fewer trials would be available for the estimation of the learning gradient, thus leading to less reliable estimates. We also found that the mean of the learning gradient B depended on $P_{Modelbased}$, such that the more likely it was that movements had originated from the outlier distribution, the more negative the learning gradient B of the state space model was estimated to be (Fig. 4.11). Indeed across participants there was a positive correlation between $P_{Modelbased}$ and B (r^2= 0.298, p= .029).

4.4.2.4 Changes in the learning gradient *B*

Since B depended on $P_{Modelbased}$, which changed over the course of the experiment, it was necessary to remove the effects of changes in $P_{Modelbased}$ to obtain independent estimates of the learning gradient B. Therefore we performed linear regression of B over $P_{Modelbased}$. Next we fixed $P_{Modelbased}$ at its baseline value of 0.948 (Fig. 4.11) and read off the estimated B value and corresponding standard errors. These values are shown in figure 4.12 for groups of 5 blocks.

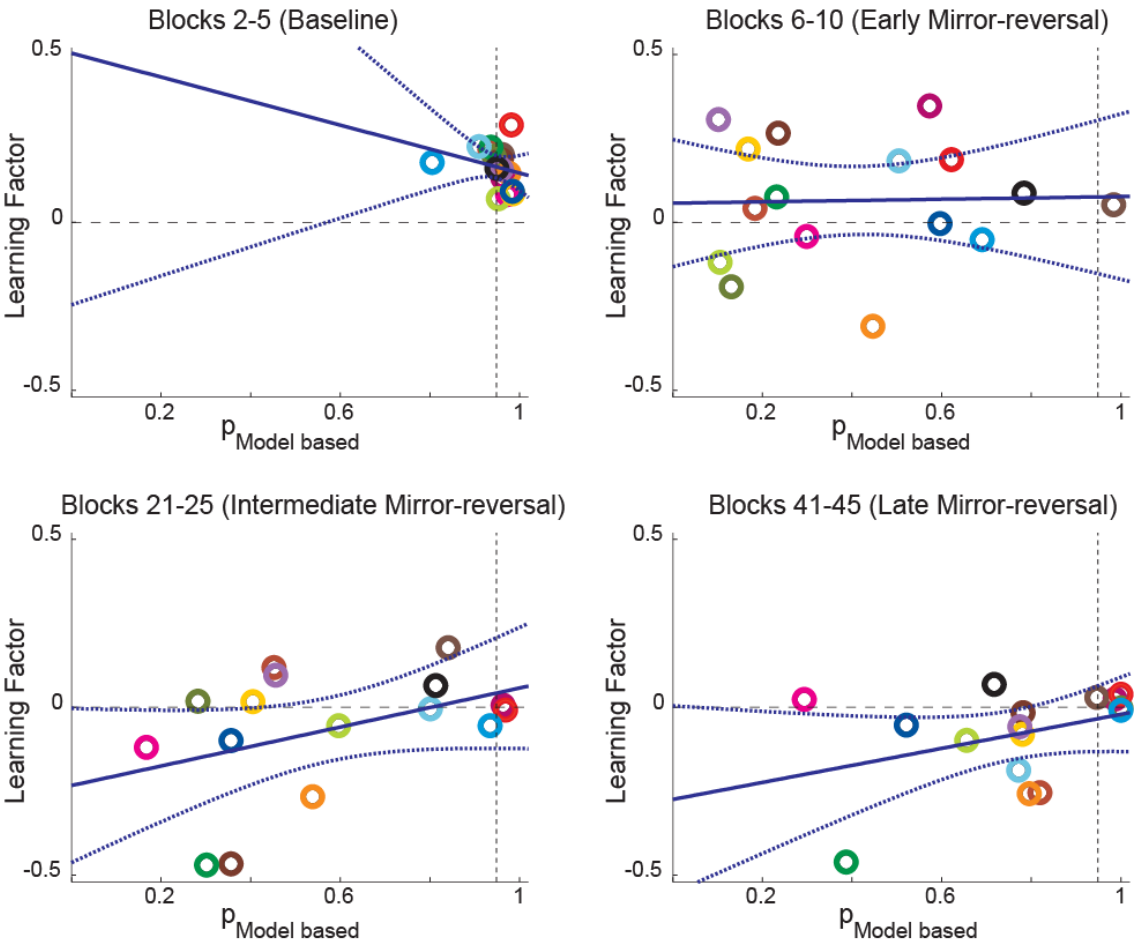

Figure 4.11: During mirror-reversal $P_{Modelbased}$ correlates with the B-values.

Each color represents one participant. Linear regression (solid blue line) and 95% confidence intervals are shown across the range of $P_{Modelbased}$. During baseline (upper left panel) all participants produced movements likely to have originated from the model and all estimates of *B* were positive. During mirror-reversal (upper right and both lower panels) $P_{Modelbased}$ decreased for most participants. Lower $P_{Modelbased}$ in turn led to lower estimates of the learning gradient *B*. To obtain estimates of *B* that were independent of $P_{Modelbased}$, we read off the value of *B* at $P_{Modelbased}$ = 0.945 (corresponding to the mean during baseline), which is indicated by the vertical dotted line. Henceforth this value will be referred to as B_{fixed}. The corresponding standard errors required for statistical tests were inferred from the size of the confidence interval at B_{fixed}.

At baseline B was 0.165 (±.169) and significantly different from zero ($t_{(15)}$ = 9.759, p = $3*10^{-8}$). At the onset of mirror-reversal B_{fixed} did not significantly differ from zero ($t_{(15)}$ =0.648, p =.263), but it also did not significantly differ from baseline ($t_{(15)}$ =-0.966, p =.349). Given the increase in variance of the estimated learning rates during early mirror-reversal, this is hardly surprising. Over the course of mirror-reversal learning B_{fixed} (Fig. 4.12) significantly decreased (r^2= .519, p=.044). By the end of the first session B_{fixed} was close to zero (0.004± .142), and marginally differed from baseline ($t_{(15)}$ =2.026, p = .061). Blocks 21-25 ($t_{(15)}$ = 2.435, p = .028), blocks 26-30 ($t_{(15)}$ = 3.205, p =.006), blocks 31-35 ($t_{(15)}$ = 4.415, p =$5*10^{-4}$), blocks 36-40($t_{(15)}$ = 5.062, p =10^{-4}) and blocks 41-45 ($t_{(15)}$ = 5.971, p =$3*10^{-5}$) all significantly differed from baseline, however B never became significantly negative.

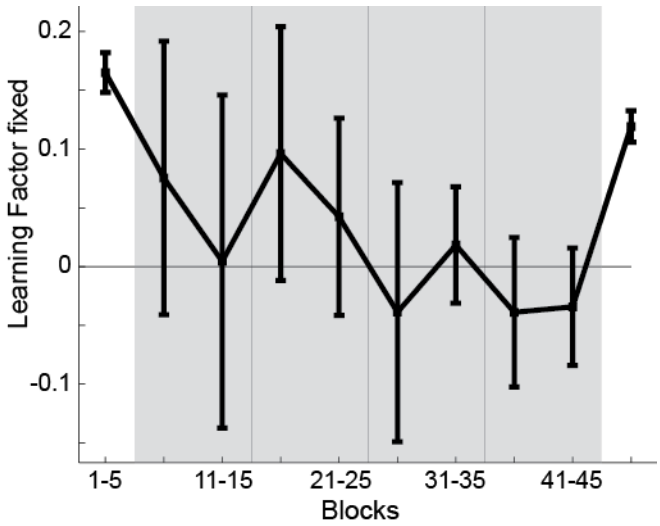

Figure 4.12: Learning decreases over the course of mirror-reversal learning.

B_{fixed} (±SE) is plotted over groups of 5 blocks. Grey background indicates mirror reversed visual feedback. Vertical dotted lines separate the different days of the experiment.

Note that when reaching under random visual rotations of the cursor feedback, an ideal learner should set the learning gradient to zero. If the learning rate would have been reduced only because the motor system stopped to correct for the imposed visual rotations, then in the washout blocks B_{fixed} should be close to zero. However, when visual feedback was switched back to normal again (blocks 46-50), the learning gradient B_{fixed} instantly assumed positive values again ($t_{(15)} = 8.822$, $p = 3*10^{-7}$), which were significantly higher than in the blocks 41-45 of mirror reversed reaching ($t_{(15)} = -4.844$, $p = 2*10^{-4}$). Taken together with the trend towards negative B_{fixed} in the last day of training it seems unlikely that the participants ignored the visual feedback altogether. We interpret the reduction in the learning gradient over the time course of learning as a true change in the way the system learns from errors under mirror reversed visual feedback. Although the negative values in the last day might suggest that B ultimately can be reversed, our data do not provide statistical evidence for a learning gradient reversal. The reduction in B might thus have been caused either by a general gain decrease of the process that underlies parametric trial-by-trial recalibration or alternatively a gradual reversal of the learning gradient. We can however conclude with certainty that mirror-reversed reaching changes the way the motor system learns from errors.

4.5 Discussion

In the current chapter we asked whether the motor system can relearn how to learn from errors during continued exposure to mirror-reversed visual feedback. To address this question, 16 participants made reaching movements in a mirror-reversed environment over 4 consecutive days. To test how participants learned from visual error signals, we applied small rotations – varying randomly in sign and magnitude from trial to trial - to the cursor position. We then fitted a model which explains the behavior of the subjects as a mixture of a recalibration process and an outlier distribution that may contain strategic movements. In this way we estimated how the learning gradient B of the state space model and the probability of movements to be generated by either distribution changed over the course of the experiment.

If participants had continued to correct for errors using the internal model they had used before mirror-reversal started, they would have corrected for errors in the wrong direction and thus with every unreversed state update drifted further away from the target. Indeed we observed such drifts during the first blocks of mirror reversed reaching, indicating that initially participants continued to use an unreversed learning rule. However, the average reach direction drifted back towards the target location as a result of further training. There were two possible explanations that could have explained why participants stopped to drift out and instead reached towards the target again. They had either reversed the learning gradient B or stopped error-based learning.

From a purely computational perspective reversing the sign of the learning gradient is trivial. However, computational difficulty is not always a good predictor

for how difficult the actual implementation in the brain is (Marr and Poggio, 1976). It has previously been claimed that a complete reversal of the learning gradient is possible (Abdelghani et al., 2008; Abdelghani and Tweed, 2010). In these experiments the learning gradient was never directly assessed during error-based learning, but instead only the control Jacobian that underlies feedback corrections was tested using a joystick. However, it is unclear whether the control Jacobian that underlies feedback control is equivalent with the learning gradient used for error-based recalibration.

Recalibration is supported by two simultaneous processes that operate at different time scales (Smith et al., 2006). The slow process learns slowly, but retains well, whereas the fast process learns faster but retains little. It has been shown that learning in the fast state can be altered through experience. It has also been shown that after participants had been exposed to random rotations their learning rate in the fast model was subsequently increased compared to previous exposure to no or consistent rotations (Turnham et al., 2012). Similarly we found facilitation of the learning gradient B, which after regressing out slow trends in the data, should mainly reflect the fast process. However in Turnham's study as in our study it is impossible to tell whether the change in the learning gradient represented a true change in the underlying mapping, or whether the mapping remained unchanged and the outputs derived from that mapping were inhibited instead by some mediating process. A significant sign reversal of the learning gradient would have been strong evidence for plastic changes in the mapping underlying trial-by-trial recalibration. Alternatively the entire fast learning process might have been scaled up or down in response to the randomness of the external perturbations (Wei and Körding, 2009; Turnham et al., 2012).

Structural learning describes learning to learn phenomena in motor control (Braun et al., 2010; Turnham et al., 2011). In this framework, the motor system could be characterized as being unfamiliar with the structure of the mirror-reversed environment. Based on the system's Bayesian priors over the mapping between motor commands and sensory consequences, visual error signals were interpreted to stem from a rotation, scaling or lateral translation. The current results could equally well have been described in a Bayesian framework, where the prior exposure to an environment predicts the control policies chosen for future movements (Tenenbaum et al., 2011; Kobak and Mehring, 2012; Yousif and Diedrichsen, 2012). However, given that more than 2000 trials of mirror reversed reaching were not sufficient for reversing the learning gradient, it seems likely that the required mapping was not available to the system. Thus from a structural learning perspective, the structure of learning from mirror-reversal has not been fully established by the motor system within the time frame under study.

Similar to the structural learning perspective, from a control theoretical perspective, if the learning gradient had at some point reversed, this would have been sufficient for the assertion that during mirror-reversal not only a new control policy but also a new internal model was established (Lalazar and Vaadia, 2008; Haith and Krakauer, 2013). Note that from a reinforcement learning perspective, control policies can be simple point to point mappings, whereas an internal model would possess knowledge about the nonlinear function that relates sensory inputs to motor outputs (Sutton and Barto, 1998). In this sense an established internal model could have been used to correctly interpret error signals during mirror-reversal and thus have driven parametric recalibration.

Over the course of training the learning gradient was reduced until it approximated zero by the 3rd day. This can be interpreted to be due to gradual changes in the internal model from which the corrections were derived or as silencing of error-based learning. The optimal learning rule in the face of uncorrelated, random errors is not to learn at all. Thus any reduction in the learning gradient might have been explained simply by the system learning to not learn from the imposed rotations. One might argue that the learning gradient on the 4th day – although not significantly negative – suggests that error-based learning was not shut down but instead started to show first signs of reversal. Yet, on statistical grounds we have to conclude that error-based learning did not reverse even after 4 days of training. One implication of the unreversed learning gradient is that as explained in chapters 2 and 3 mirror-reversal cannot be learned by parametric recalibration, not even after 4 days of training. A new control policy has to be established, but cannot be derived from the internal model(s) previously used for feedforward and feedback control and recalibration.

At the onset of mirror-reversal in the current experiment, movements were more likely to be generated by the outlier distribution than the state space model. Thus, with an unreversed recalibration process participants relied very strongly on a different process.

Likewise reaction times increase during mirror-reversed reaching movements, reflecting the use of a new computationally expensive process. Consequently we here found that participants with stronger increases in reaction times during training trials were more likely to produce movements belonging to the outlier distribution in testing trials. This finding establishes a link between feedforward control and the way the system responds to visual error signals,

suggesting that feedforward control and parametric recalibration rely on similar mappings (Wagner and Smith, 2008). While the underlying processes giving rise to the outlier distribution are only captured in a non-specific way in the mixture model employed here, one might speculate that these movements depend on more cognitive mechanisms (Mazzoni and Krakauer, 2006; Taylor and Ivry, 2012). Participants might for example start to actively ignore visual feedback, or alternatively draw on explicit knowledge about the mirror-reversal to counter the drift. Even though, reaction times were restricted, there was only a single target presented during testing trials. Thus, it was possible to mentally prepare the movement long before the target was even visually presented, with the visual target presentation being used as a go-signal similar to anti-pointing (Neely and Heath, 2010) or anti-saccade tasks (Zhang and Barash, 2000; Munoz and Everling, 2004), which might explain why reaction times during testing trials did not significantly correlate with the likelihood of belonging to the model-based or the outlier distribution.

In conclusion we confirmed that speed-accuracy trade-offs emerge in feedforward control of mirror-reversal learning. In addition we have observed that these trade-offs did not only shift, but also became weaker as a result of continued training. We have found no evidence for a reversal of the internal model that underlies parametric recalibration. However, we found that participants stop to engage in trial-by-trial error-based recalibration, when reaching under mirror reversed feedback. It is possible that further training might have led to a reversal of recalibration. However 2400 trials under mirror reversed visual feedback were insufficient for the development of a reversed mapping in trial-by-trial recalibration. Analogous to the emergence of new control policies in feedforward and feedback

control, a new process dominated error-based learning to ensure normal functioning in the face of detrimental internal models.

Chapter 5
Discussion

5.1 Summary

It is the main goal of this thesis to contribute to a more principled understanding of the types of learning that are commonly described by the term adaptation. In the following I first provide a short summary of the key experimental findings. Thereafter I discuss the results and implications presented here in the light of existing neuroscientific concepts. I conclude with a bullet point summary and state the contributions to the field of neuroscience.

In chapter 2, in an experiment where the required magnitude of the change in the sensorimotor map was equivalent in both conditions, I found that different processes governed the behavior and consolidation in learning a 40° visuomotor rotation and mirror-reversal. During mirror-reversal learning trade-offs between processing time and accuracy emerged in feedforward and feedback control,

whereas accuracy was independent of reaction times during rotation learning. In addition, within a 24-hour period, learning mirror-reversal resulted in offline gains whereas visual rotations resulted in forgetting. I have argued that these dissociations reveal the presence of two different learning mechanisms: recalibration when learning small visuomotor rotations and control-policy acquisition when learning mirror-reversals. In chapter 3 I found that a visual rotation of 180° elicits time-accuracy trade-offs similar to those found in mirror-reversal. Therefore the presence of these trade-offs cannot be explained by structural-learning differences between mirror-reversal and rotations per se. Subsequently I asked why 40° rotations but not 180° rotations can be learned through recalibration. Crucially when the rotation size was gradually increased, strong learners, who did not fall behind as quickly, switched from recalibration to acquisition later than weak learners. Therefore the magnitude of the angular error rather than the amount of absolute adaptation determined the behavioral switch. The switching point was found to be close to 90° of error, which is consistent with the interpretation that recalibration can only be used when the correction issued under the old sensorimotor map is correlated with the required change in the map. Finally in chapter 4 I asked whether the mapping that underlies trial-by-trial recalibration can be mirror reversed. The results suggest that trial-by-trial recalibration itself cannot be significantly mirror reversed after 4 days of training. However, since reaction times during training trials and the learning gradient during testing trials were correlated across participants, the results suggest that recalibration and feedforward control rely on shared internal models.

5.2 Either recalibration or acquisition?

The experiments were designed to dissociate between recalibration and control policy acquisition as cleanly as possible. However I do not mean to suggest that the learning processes described here can only be used in isolation. Rather in many real-world situations both might be active concurrently. When reaction times are not restricted, time-accuracy trade-offs emerge for small rotations as well (Fernandez-Ruiz et al., 2011b), meaning that in addition to the recalibration process, performance is in part also improved by additional time-intensive computations such as mentally rotating the target (Georgopoulos and Massey, 1987; Mazzoni and Krakauer, 2006; Taylor et al., 2010; Taylor and Ivry, 2011). Therefore even small visual rotations are not always learned purely through recalibration of an existing control policy: without speed constraints additional time-consuming processes (strategic remapping) can help to improve performance more quickly. Besides, additional types of motor learning such as use-dependent learning (Diedrichsen et al., 2010) are presumably involved as well.

5.3 Devaluation of the error signal

In chapter 4 a model was introduced that discounted the visual error signal by the cosine of its magnitude for recalibration learning. Indeed we found that this model was able to fit the behavior of the participants relatively accurately. The switching points from recalibration to a more strategic mechanism were estimated to be close to 90°, which is exactly where they should be given a learning mechanism that devalues angular errors by their cosine. The error devaluation assumption follows from the idea that the brain represents visuomotor errors in a Cartesian

reference frame. However, error devaluation is equally predicted by feedback-error learning (Kawato, 1995): The bigger the prediction error is, the more a motor update should influence not only the movement components that are perpendicular to the feedforward command, but also those that are parallel. For a 45° error, an error-based update should then result in equal amounts of perpendicular as in parallel velocity changes in the following movement. It should be possible to test this hypothesis by either measuring movement velocity or by measuring movement endpoints. In chapter 3, I could not study changes in movement velocity because the range of instructed movement speed was rather wide. In addition the cursor was visible at all times, such that online corrections would have affected velocity and endpoint measurements. Most importantly movement speeds tend to naturally vary considerably over time. However, in the gradual learning experiment, error sizes were relatively consistent at any point in time. An experiment specifically designed to test this hypothesis, should only include endpoint feedback and apply rotations that vary at random in size and magnitude on a trial-by-trial basis. Taken together while our results hint at the existence of a Cartesian reference frame for updating motor commands, the strongest prediction of this hypothesis was not addressed here.

5.4 The level of the argument

One might argue that when the old control policy is recalibrated it becomes a new control policy itself. Therefore from a computational perspective one might ask how the *recalibrated new* control is policy different from the *acquired new* control policy. However, the current thesis does not aim to provide a computational framework for recalibration and skill learning.

Motor learning can be understood as a change in the function (or control policy) that generates both the feedforward and the feedback command. So what does "recalibrating an old function" or "learning a new function" mean then? On a computational level of description (Marr and Poggio, 1976) it is unclear how such a distinction could be made in a principled fashion. However, when asking how the motor system performs these computations on the level of algorithms and representations, the distinction can be made whether the same or different structures are used to generate the motor commands. From the perspective of internal models it has been argued that multiple paired forward and inverse models exist and can be activated in a context-dependent fashion (Wolpert and Kawato, 1998). Empirically this idea has been addressed by investigating whether the motor system can switch between two motor behaviors. When participants were alternatingly presented with two opposing force fields during reaching movements over four consecutive days, they were unable to learn to switch from the representation of one field to another (Karniel and Mussa-Ivaldi, 2002). Interestingly though, when additional contextual cues were presented with rapidly alternating force-fields, participants were able to rapidly switch between the force fields (Osu et al., 2004), which has been interpreted as evidence for the acquisition

of separate internal models. Another way to probe the existence of separate representations is to study the extent to which the memory for an external perturbation generalizes to other similar movements (Nozaki et al., 2006; Kluzik et al., 2008; White and Diedrichsen, 2013). The ability to hold different adaptation states for different behaviors is an indication for partially separate control policies. One might even speculate whether with several weeks of training in a mirror-reversal task, it is possible to learn visual rotations in a mirror-reversed environment, similar to the experiment described in Chapter 4 of this thesis. If participants could be trained to rapidly switch between mirror-reversals and normal visual feedback, one could test how much the memory for the visual rotation generalizes to reaching under normal visual feedback during identical movements. The ability to switch between the two control policies should be highly correlated with the ability to hold separate calibrations. The more independent the underlying control policies are, the less recalibration should generalize between the two.

This thesis provides a different insight into this problem by studying the speed with which the underlying computations can be performed, by examining feedforward and feedback motor commands after short processing times. I assume that reaching is controlled by a set of context-dependent control policies, which can support very fast and automatic computations. When learning is achieved by recalibrating existing control policies, their automaticity can be inherited, as the same neural circuits are used to perform the computation. However, when a new control policy is established, the new output is first achieved by slower processes, which then as a result of practice become increasingly automatic (Fu and Anderson, 2006). That is, like generalization and switching, the speed-accuracy trade-off curve can serve as one possible characterization of the underlying control policy or

computational process, which indicates whether two behaviors are generated by the same or different control policies. I here make the distinction between recalibration and learning de novo purely on the basis of the underlying speed-accuracy trade-off curve – i.e. on the basis of how the brain implements these processes. The finding that mirror-reversal and visual-rotation learning results in offline gains and forgetting respectively further supports the notion that different memory systems are involved in learning of the two transformations.

5.5 Offline gains

There is an on-going debate about the existence of offline gains per se. Part of this debate stems from the problem that like adaptation the concept of offline gains is ill-defined. For example in many studies in which offline gains have been reported, hundreds of trials from the end of the pre-sleep session were averaged and compared to hundreds of trials from the beginning of the post-sleep session (Maquet et al., 2003; Stickgold, 2005; Landsness et al., 2009). It has been argued that this procedure is problematic because if performance continues to increase throughout the entire first session, the averaged performance in the first session might underestimate the true level of performance expressed just before the break (Rickard et al., 2008). Likewise performance improvements throughout the second session can lead to overestimation of performance levels present immediately after the break. Thus by averaging over too many trials, trial-by-trial improvements that happened in the second session can mask the effect of forgetting.

The critical insight to be gained from the experiments presented in this thesis is not so much whether there is a true performance improvement from one session

to the next or whether instead relearning is just quicker in the second session. Instead, the results gain their strength from the dissociation between forgetting and offline gains in a simple reaching movement task, with an identical performance measure for mirror-reversal and visual-rotation learning.

The contrast of these different consolidation signatures might improve the understanding of the findings from previous studies. Offline improvements in reaching movement tasks have been demonstrated in two experiments where participants made reaching movements under visual rotations (Landsness et al., 2009; Määttä et al., 2010). In the first session participants had a 10-minute break before continuing for a few more trials (post-break). The performance during the post-break movements was worse than during the pre-break period. When the experiment continued on the following day, participants showed spontaneous improvements relative to the post-break period of the previous day. However compared to the pre-break trials of the previous day they still performed significantly worse. In the light of the findings presented in chapter 2, one could speculate whether two concurrently active learning mechanisms gave rise to initial forgetting during the rest break in day 1 and subsequent offline gains overnight (Mazzoni and Krakauer, 2006; Taylor et al., 2010; Huang et al., 2011): The recalibration mechanism forgets quickly during the 10-minute break on the first day and leads to decreased performance in the post-break period. In contrast, the mechanism that acquires the new control policy benefits from the overnight break, resulting in offline improvements on the following day.

5.6 Reinforcement learning

The concept of recalibration as defined in this thesis is equivalent to first-order error-based learning. The proof that recalibration depends on the sign of the error signal - and not just on reward signals - comes from experiments on trial-by-trial adaptation (Tseng et al., 2007; Srimal et al., 2008; Turnham et al., 2012). Although, during these experiments trial-by-trial learning leads to worse performance than not learning at all, participants still adapt. Furthermore, when reaching in a mirror-reversed environment, where the sign of the error signal has to be inverted, error-based learning is impaired (Chapter 4).

In contrast, during control policy acquisition the cursor position might primarily serve as a reward signal. This would also explain why patients with lesions in the Basal Ganglia show abnormal learning in a range of motor skills (Graybiel, 2005; Boyd et al., 2009) and mirror-reversal learning, but relatively unimpaired learning of small visual rotations (Fernandez-Ruiz et al., 2003; Gutierrez-Garralda et al., 2013). Parkinson patients, who have repeatedly been shown to be impaired in operant reinforcement learning, exhibit normal adaptation, but decreased savings after learning a 30°-visuomotor rotation (Leow et al., 2012). However they are impaired in adapting to 90°-rotations (Contreras-Vidal and Buch, 2003). A possible explanation might be that recalibration can compensate for impairments in reinforcement learning when errors are small, but only as long as visuomotor errors are smaller than 90°.

In sum it seems likely that control policy acquisition is implemented through reinforcement learning. This idea is in line with the promising reinforcement-

learning framework of model-based and model-free learning, which has recently been introduced to the field of motor learning (Huang et al., 2011).

5.7 Model-based and model-free learning

Is recalibration model-based and acquisition model-free learning? Clearly recalibration is first-order model-based learning. However for control policy acquisition the answer depends on the level of models that we consider. The terms model-based and model-free originate from the field of reinforcement learning (Sutton and Barto, 1998; Fu and Anderson, 2006), where the term model (of the environment) is used in the most general sense to refer to an internal representation of the environment. A model can simulate the environment and in this way is invaluable for the purpose of planning future actions. When participants acquired new control policies they did not explore the entire search space. Instead it is reasonable to assume that more cognitive and strategic components guided their behavior. Being able to re-plan a movement by mentally rotating the target position in a strategic way requires a higher order cognitive model of the environment. Following this original reinforcement-learning definition control policy acquisition is probably model-based.

However, if instead we define models exclusively as the control Jacobian that is implemented by the forward and inverse model, then recalibration is model-based, whereas control policy acquisition as described in this thesis is model-free. Note that when recalibration and acquisition are active simultaneously a useful internal model must be present for recalibration to proceed. It is unclear whether the acquisition of a new control policy can be guided by an existing internal model,

if possible. Therefore in the debate about model-based and model-free learning in the context of motor control (Huang et al., 2011; Haith and Krakauer, 2013) it is important to be precise about the model definition. With these limitations in mind, the findings presented here do support the general notion of a first-order-model-based recalibration mechanism and an additional first-order-model-free mechanism in learning to compensate visuomotor perturbations, the latter of which might proceed by reinforcement learning and might be similar to the type of learning that is often observed in skill learning tasks such as finger sequence learning and chording (Boyd et al., 2009; Wright et al., 2010; Waters-Metenier et al., 2014).

5.8 Key findings and interpretation

1) Recalibration relies on an inverse model that can generate beneficial state updates from the directional information contained in the prediction error (Chapters 2, 3 & 4).
2) Adaptation can be achieved by recalibration and/or the establishment of a novel time intensive computation (Chapter 2).
3) When adaptation involves the establishment of a new time intensive computation/ control policy, offline gains can occur (Chapter 2).
4) The original non-adapted control policy remains active, such that it dominates motor outputs that are generated after relatively short processing times (Chapters 2, 3 & 4).
5) The new control policy becomes increasingly automatic and can be computed at shorter latencies as a result of training (Chapters 2, 3 & 4).

6) I did not find evidence for an upper boundary on recalibration. Instead there was an upper boundary on the size of error signals that can drive recalibration (Chapter 3).
7) The mapping that underlies recalibration itself could not be reversed within 4 days of training (Chapter 4).
8) Recalibration and feedforward control might rely on shared internal models (Chapter 4).

5.9 Contributions to the field of motor learning

One of the major contributions of this thesis to the field of motor learning is the establishment of solid and testable behavioral criteria that can be used as indicators for the type of motor learning being used. Furthermore the current thesis demonstrates that characteristics like time-accuracy trade-offs and offline gains that are typical for complex skill learning tasks can be elicited in arm movement adaptation tasks as well. Thereby it provides a bridge between skill learning and adaptation, via the concept of de novo control policy acquisition in contrast to recalibration. Finally this thesis offers the possibility to directly compare recalibration and control policy acquisition in the same well controlled task (reaching movements under visual rotations of the cursor feedback) with the same movements, largely identical visual feedback and the same performance measure, which is ideal for contrasting the underlying neural substrates with neurophysiological tools.

References

Abdelghani MN, Lillicrap TP, Tweed DB (2008) Sensitivity Derivatives for Flexible Sensorimotor Learning. Neural Comput 20:2085–2111.

Abdelghani MN, Tweed DB (2010) Learning course adjustments during arm movements with reversed sensitivity derivatives. BMC Neurosci 11:150-161.

Abeele S, Bock O (2001a). Sensorimotor adaptation to rotated visual input: different mechanisms for small versus large rotations. Exp Brain Res 140:407–410.

Abeele S, Bock O (2001b) Mechanisms for sensorimotor adaptation to rotated visual input. Exp Brain Res 139:248–253.

Abe M, Schambra H, Wassermann EM, Luckenbaugh D, Schweighofer N, Cohen LG (2011) Reward Improves Long-Term Retention of a Motor Memory through Induction of Offline Memory Gains. Curr Biol 21:557–562.

Beilock SL, Bertenthal BI, Hoerger M, Carr TH (2008) When does haste make waste? Speed-accuracy tradeoff, skill level, and the tools of the trade. J Exp Psychol Appl 14:340–352.

Bell CC (1989) Sensory coding and corollary discharge effects in mormyrid electric fish. J Exp Biol 146:229–253.

Blakemore S-J, Wolpert DM, Frith CD (1998) Central cancellation of self-produced tickle sensation. Nat Neurosci 1:635–640.

Boyd LA, Edwards JD, Siengsukon CS, Vidoni ED, Wessel BD, Linsdell MA (2009) Motor sequence chunking is impaired by basal ganglia stroke. Neurobiol Learn Mem 92:35–44.

Braun DA, Aertsen A, Wolpert DM, Mehring C (2009a) Motor Task Variation Induces Structural Learning. Curr Biol 19:352–357.

Braun DA, Aertsen A, Wolpert DM, Mehring C (2009b) Learning Optimal Adaptation Strategies in Unpredictable Motor Tasks. J Neurosci 29:6472–6478.

Braun DA, Aertsen A, Wolpert DM, Mehring C (2009c) Motor Task Variation Induces Structural Learning. Curr Biol 19:352–357.

Braun DA, Mehring C, Wolpert DM (2010) Structure learning in action. Behav Brain Res 206:157–165.

Brawn TP, Fenn KM, Nusbaum HC, Margoliash D (2010) Consolidating the Effects of Waking and Sleep on Motor-Sequence Learning. J Neurosci 30:13977–13982.

Calvo-Merino B, Glaser DE, Grèzes J, Passingham RE, Haggard P (2005) Action Observation and Acquired Motor Skills: An fMRI Study with Expert Dancers. Cereb Cortex 15:1243–1249.

Cohen DA, Pascual-Leone A, Press DZ, Robertson EM (2005a) Off-line learning of motor skill memory: A double dissociation of goal and movement. Proc Natl Acad Sci 102:18237–18241.

Cohen HS, Bloomberg JJ, Mulavara AP (2005b) Obstacle avoidance in novel visual environments improved by variable practice training. Percept Mot Skills 101:853–861.

Contreras-Vidal JL, Buch ER (2003) Effects of Parkinson's disease on visuomotor adaptation. Exp Brain Res 150:25–32.

Costa RM (2011) A selectionist account of de novo action learning. Curr Opin Neurobiol 21:579–586.

Criscimagna-Hemminger SE, Bastian AJ, Shadmehr R (2010) Size of Error Affects Cerebellar Contributions to Motor Learning. J Neurophysiol 103:2275–2284.

Cunningham HA (1989) Aiming error under transformed spatial mappings suggests a structure for visual-motor maps. J Exp Psychol Hum Percept Perform 15:493–506.

Day BL, Lyon IN (2000) Voluntary modification of automatic arm movements evoked by motion of a visual target. Exp Brain Res 130:159–168.

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

Derégnaucourt S, Mitra PP, Fehér O, Pytte C, Tchernichovski O (2005) How sleep affects the developmental learning of bird song. Nature 433:710–716.

Desmurget M, Grafton S (2000) Forward modeling allows feedback control for fast reaching movements. Trends Cogn Sci 4:423–431.

Diedrichsen J (2007) Optimal Task-Dependent Changes of Bimanual Feedback Control and Adaptation. Curr Biol 17:1675–1679.

Diedrichsen J, White O, Newman D, Lally N (2010) Use-Dependent and Error-Based Learning of Motor Behaviors. J Neurosci 30:5159–5166.

Diekelmann S, Born J (2010) The memory function of sleep. Nat Rev Neurosci 11:114–126.

Dietterich TG, Michalski RS (1981) Inductive learning of structural descriptions: Evaluation criteria and comparative review of selected methods. Artif Intell 16:257–294.

Donders FC (1969) On the speed of mental processes. Acta Psychol (Amst) 30:412–431.

Doya K (2000) Complementary roles of basal ganglia and cerebellum in learning and motor control. Curr Opin Neurobiol 10:732–739.

Doyon J, Bellec P, Amsel R, Penhune V, Monchi O, Carrier J, Lehéricy S, Benali H (2009a) Contributions of the basal ganglia and functionally related brain structures to motor learning. Behav Brain Res 199:61–75.

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

Fernandez-Ruiz J, Diaz R, Hall-Haro C, Vergara P, Mischner J, Nuñez L, Drucker-Colin R, Ochoa A, Alonso ME (2003) Normal prism adaptation but reduced after-effect in basal ganglia disorders using a throwing task. Eur J Neurosci 18:689–694.

Fernandez-Ruiz J, Wong W, Armstrong IT, Flanagan JR (2011a) Relation between reaction time and reach errors during visuomotor adaptation. Behav Brain Res 219:8–14.

Fernandez-Ruiz J, Wong W, Armstrong IT, Flanagan JR (2011b) Relation between reaction time and reach errors during visuomotor adaptation. Behav Brain Res 219:8–14.

Fischer S, Hallschmid M, Elsner AL, Born J (2002) Sleep forms memory for finger skills. Proc Natl Acad Sci 99:11987–11991.

Fitts PM (1954) The information capacity of the human motor system in controlling the amplitude of movement. J Exp Psychol 47:381–391.

Franklin DW, Wolpert DM (2008) Specificity of Reflex Adaptation for Task-Relevant Variability. J Neurosci 28:14165–14175.

Fu W-T, Anderson JR (2006) From recurrent choice to skill learning: A reinforcement-learning model. J Exp Psychol Gen 135:184–206.

Galea JM, Vazquez A, Pasricha N, Xivry J-JO de, Celnik P (2011) Dissociating the Roles of the Cerebellum and Motor Cortex during Adaptive Learning: The Motor Cortex Retains What the Cerebellum Learns. Cereb Cortex 21:1761–1770.

Gentilucci M, Daprati E, Toni I, Chieffi S, Saetti MC (1995) Unconscious updating of grasp motor program. Exp Brain Res 105:291–303.

Georgopoulos AP, Massey JT (1987) Cognitive spatial-motor processes: 1. The making of movements at various angles from a stimulus direction. Exp Brain Res 65:361-370.

Gerardin E, Pochon J-B, Poline J-B, Tremblay L, Van de Moortele P-F, Levy R, Dubois B, Le Bihan D, Lehéricy S (2004) Distinct striatal regions support movement selection, preparation and execution. Neuroreport 15:2327–2331.

Graybiel AM (2005) The basal ganglia: learning new tricks and loving it. Curr Opin Neurobiol 15:638–644.

Gritsenko V, Kalaska JF (2010) Rapid Online Correction Is Selectively Suppressed During Movement With a Visuomotor Transformation. J Neurophysiol 104:3084–3104.

Gutierrez-Garralda JM, Moreno-Briseño P, Boll M-C, Morgado-Valle C, Campos-Romo A, Diaz R, Fernandez-Ruiz J (2013) The effect of Parkinson's disease and Huntington's disease on human visuomotor learning. Eur J Neurosci 38:2933–2940.

Haith AM, Krakauer JW (2013) Model-Based and Model-Free Mechanisms of Human Motor Learning. In: Progress in Motor Control (Richardson MJ, Riley MA, Shockley K, eds), pp 1–21. Springer New York.

Hay L, Redon C (1999) Feedforward versus feedback control in children and adults subjected to a postural disturbance. Exp Brain Res 125:153–162.

Helmholtz H von (1866) Handbuch der physiologischen Optik. Voss.

Hikosaka O, Nakamura K, Sakai K, Nakahara H (2002) Central mechanisms of motor skill learning. Curr Opin Neurobiol 12:217–222.

Holst E von, Mittelstaedt H (1950) Das Reafferenzprinzip. Naturwissenschaften 37:464–476.

Huang VS, Haith A, Mazzoni P, Krakauer JW (2011) Rethinking Motor Learning and Savings in Adaptation Paradigms: Model-Free Memory for Successful Actions Combines with Internal Models. Neuron 70:787–801.

Illert M, Trauner M, Weller E, Wiedemann E (1986) Forearm muscles of man can reverse their function after tendon transfers: An electromyographic study. Neurosci Lett 67:129–134.

Imamizu H, Shimojo S (1995) The locus of visual-motor learning at the task or manipulator level: Implications from intermanual transfer. J Exp Psychol Hum Percept Perform 21:719–733.

Jordan MI, Rumelhart DE (1992) Forward models: Supervised learning with a distal teacher. Cogn Sci 16:307–354.

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

Karni A, Meyer G, Jezzard P, Adams MM, Turner R, Ungerleider LG (1995) Functional MRI evidence for adult motor cortex plasticity during motor skill learning. Nature 377:155–158.

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

Kassardjian CD, Tan Y-F, Chung J-YJ, Heskin R, Peterson MJ, Broussard DM (2005) The Site of a Motor Memory Shifts with Consolidation. J Neurosci 25:7979–7985.

Kawato M (1995) Neural Networks for Control (Miller WT, Werbos PJ, Sutton RS, eds). pp 197-228. MIT Press Cambridge, MA, USA

Kawato M, Gomi H (1992) The cerebellum and VOR/OKR learning models. Trends Neurosci 15:445–453.

Klassen J, Tong C, Flanagan JR (2005) Learning and recall of incremental kinematic and dynamic sensorimotor transformations. Exp Brain Res 164:250–259.

Kluzik J, Diedrichsen J, Shadmehr R, Bastian AJ (2008) Reach Adaptation: What Determines Whether We Learn an Internal Model of the Tool or Adapt the Model of Our Arm? J Neurophysiol 100:1455–1464.

Kobak D, Mehring C (2012) Adaptation Paths to Novel Motor Tasks Are Shaped by Prior Structure Learning. J Neurosci 32:9898–9908.

Körding KP, Wolpert DM (2004) The loss function of sensorimotor learning. Proc Natl Acad Sci 101:9839–9842.

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

Krakauer JW, Ghez C, Ghilardi MF (2005) Adaptation to Visuomotor Transformations: Consolidation, Interference, and Forgetting. J Neurosci 25:473–478.

Krakauer JW, Mazzoni P (2011) Human sensorimotor learning: adaptation, skill, and beyond. Curr Opin Neurobiol 21:636–644.

Krakauer JW, Pine ZM, Ghilardi M-F, Ghez C (2000) Learning of Visuomotor Transformations for Vectorial Planning of Reaching Trajectories. J Neurosci 20:8916–8924.

Kuriyama K, Stickgold R, Walker MP (2004) Sleep-dependent learning and motor-skill complexity. Learn Mem 11:705–713.

Laforce Jr. R, Doyon J (2001) Distinct Contribution of the Striatum and Cerebellum to Motor Learning. Brain Cogn 45:189–211.

Lalazar H, Vaadia E (2008) Neural basis of sensorimotor learning: modifying internal models. Curr Opin Neurobiol 18:573–581.

Landsness EC, Crupi D, Hulse BK, Peterson MJ, Huber R, Ansari H, Coen M, Cirelli C, Benca RM, Ghilardi MF, Tononi G (2009) Sleep-Dependent Improvement in Visuomotor Learning: A Causal Role for Slow Waves. Sleep 32:1273–1284.

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

Lesage E, Morgan BE, Olson AC, Meyer AS, Miall RC (2012) Cerebellar rTMS disrupts predictive language processing. Curr Biol 22:R794–R795.

Lotze M, Scheler G, Tan H-R., Braun C, Birbaumer N (2003) The musician's brain: functional imaging of amateurs and professionals during performance and imagery. Neuroimage 20:1817–1829.

Määttä S, Landsness E, Sarasso S, Ferrarelli F, Ferreri F, Ghilardi MF, Tononi G (2010) The effects of morning training on night sleep: A behavioral and EEG study. Brain Res Bull 82:118–123.

Maquet P, Schwartz S, Passingham R, Frith C (2003) Sleep-Related Consolidation of a Visuomotor Skill: Brain Mechanisms as Assessed by Functional Magnetic Resonance Imaging. J Neurosci 23:1432–1440.

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

Marr D, Poggio T (1976) From Understanding Computation to Understanding Neural Circuitry. Available at: http://dspace.mit.edu/handle/1721.1/5782 [Accessed March 8, 2014].

Martin TA, Keating JG, Goodkin HP, Bastian AJ, Thach WT (1996) Throwing while looking through prisms I. Focal olivocerebellar lesions impair adaptation. Brain 119:1183–1198.

Mazzoni P, Krakauer JW (2006) An Implicit Plan Overrides an Explicit Strategy during Visuomotor Adaptation. J Neurosci 26:3642–3645.

McLaughlin SC (1967) Parametric adjustment in saccadic eye movements. Percept Psychophys 2:359–362.

Meister I, Krings T, Foltys H, Boroojerdi B, Müller M, Töpper R, Thron A (2005) Effects of long-term practice and task complexity in musicians and nonmusicians performing simple and complex motor tasks: Implications for cortical motor organization. 25:345–352.

Miall RC, Wolpert DM (1996) Forward Models for Physiological Motor Control. Neural Netw 9:1265–1279.

Middleton FA, Strick PL (2000) Basal ganglia and cerebellar loops: motor and cognitive circuits. Brain Res Rev 31:236–250.

Miller I, W.T. (1987) Sensor-based control of robotic manipulators using a general learning algorithm. IEEE J Robot Autom 3:157–165.

Munoz DP, Everling S (2004) Look away: the anti-saccade task and the voluntary control of eye movement. Nat Rev Neurosci 5:218–228.

Neely KA, Heath M (2009) Visuomotor mental rotation: Reaction time is not a function of the angle of rotation. Neurosci Lett 463:194–198.

Neely KA, Heath M (2010) Visuomotor mental rotation: Reaction time is determined by the complexity of the sensorimotor transformations mediating the response. Brain Res 1366:129–140.

Noto CT, Robinson FR (2001) Visual error is the stimulus for saccade gain adaptation. Cogn Brain Res 12:301–305.

Nozaki D, Kurtzer I, Scott SH (2006) Limited transfer of learning between unimanual and bimanual skills within the same limb. Nat Neurosci 9:1364–1366.

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113.

Osu R, Hirai S, Yoshioka T, Kawato M (2004) Random presentation enables subjects to adapt to two opposing forces on the hand. Nat Neurosci 7:111–112.

Oudiette D, Paller KA (2013) Upgrading the sleeping brain with targeted memory reactivation. Trends Cogn Sci 17:142–149.

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

Penhune VB, Steele CJ (2012) Parallel contributions of cerebellar, striatal and M1 mechanisms to motor sequence learning. Behav Brain Res 226:579–591.

Peyrache A, Khamassi M, Benchenane K, Wiener SI, Battaglia FP (2009) Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. Nat Neurosci 12:919–926.

Prokop T, Berger W, Zijlstra W, Dietz V (1995) Adaptational and learning processes during human split-belt locomotion: interaction between central mechanisms and afferent input. Exp Brain Res 106:449–456.

Rasmussen CE, Williams CKI (2006) Gaussian processes for machine learning. MIT Press Cambridge, MA, USA.

Reichenbach A, Franklin DW, Zatka-Haas P, Diedrichsen J (2014) A Dedicated Binding Mechanism for the Visual Control of Movement. Curr Biol 24:780–785.

Reis J, Schambra HM, Cohen LG, Buch ER, Fritsch B, Zarahn E, Celnik PA, Krakauer JW (2009a) Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. Proc Natl Acad Sci 106:1590–1595.

Reisman DS, Wityk R, Silver K, Bastian AJ (2007) Locomotor adaptation on a split-belt treadmill can improve walking symmetry post-stroke. Brain 130:1861–1872.

Rickard TC, Cai DJ, Rieth CA, Jones J, Ard MC (2008) Sleep does not enhance motor sequence learning. J Exp Psychol Learn Mem Cogn 34:834–842.

Robertson EM, Pascual-Leone A, Miall RC (2004a) Current concepts in procedural consolidation. Nat Rev Neurosci 5:576–582.

Robertson EM, Pascual-Leone A, Press DZ (2004b) Awareness Modifies the Skill-Learning Benefits of Sleep. Curr Biol 14:208–212.

Robertson EM, Press DZ, Pascual-Leone A (2005) Off-Line Learning and the Primary Motor Cortex. J Neurosci 25:6372–6378.

Roller CA, Cohen HS, Kimball KT, Bloomberg JJ (2001) Variable practice with lenses improves visuo-motor plasticity. Cogn Brain Res 12:341–352.

Sanes JN, Dimitrov B, Hallett M (1990) Motor Learning in Patients with Cerebellar Dysfunction. Brain 113:103–120.

Sarlegna F, Blouin J, Bresciani J-P, Bourdin C, Vercher J-L, Gauthier GM (2003) Target and hand position information in the online control of goal-directed arm movements. Exp Brain Res 151:524–535.

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

Shadmehr R, Smith MA, Krakauer JW (2010) Error Correction, Sensory Prediction, and Adaptation in Motor Control. Annu Rev Neurosci 33:89–108.

Shelhamer M, Tiliket C, Roberts D, Kramer PD, Zee DS (1994) Short-term vestibulo-ocular reflex adaptation in humans. Exp Brain Res 100:328–336.

Shmuelof L, Krakauer JW, Mazzoni P (2012) How is a motor skill learned? Change and invariance at the levels of task success and trajectory control. J Neurophysiol 108:578–594.

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

Sperry RW (1947) Effect of crossing nerves to antagonistic limb muscles in the monkey. Arch Neurol Psychiatry 58:452–473.

Sperry RW (1950) Neural basis of the spontaneous optokinetic response produced by visual inversion. J Comp Physiol Psychol 43:482–489.

Srimal R, Diedrichsen J, Ryklin EB, Curtis CE (2008) Obligatory Adaptation of Saccade Gains. J Neurophysiol 99:1554–1558.

Stebbins ED, Glenn T, Singh J, Willingham DB, Goetz CG (1997) Intact mirror-tracing and impaired rotary-pursuit skill learning in patients with Huntington's disease: Evidence for dissociable memory systems in skill learning. Neuropsychology 11:272–281.

Stickgold R (2005) Sleep-dependent memory consolidation. Nature 437:1272–1278.

Sutton RS, Barto AG (1998) Reinforcement Learning: An Introduction. A Bradford Book. MIT Press Cambridge, MA, USA.

Takagi M, Zee DS, Tamargo RJ (1998) Effects of Lesions of the Oculomotor Vermis on Eye Movements in Primate: Saccades. J Neurophysiol 80:1911–1931.

Takagi M, Zee DS, Tamargo RJ (2000) Effects of Lesions of the Oculomotor Cerebellar Vermis on Eye Movements in Primate: Smooth Pursuit. J Neurophysiol 83:2047–2062.

Taylor JA, Ivry RB (2011) Flexible Cognitive Strategies during Motor Learning. PLoS Comput Biol 7:e1001096.

Taylor JA, Ivry RB (2012) The role of strategies in motor learning. Ann N Y Acad Sci 1251:1–12.

Taylor JA, Klemfuss NM, Ivry RB (2010) An Explicit Strategy Prevails When the Cerebellum Fails to Compute Movement Errors. The Cerebellum 9:580–586.

Tenenbaum JB, Kemp C, Griffiths TL, Goodman ND (2011) How to Grow a Mind: Statistics, Structure, and Abstraction. Science 331:1279–1285.

Thoroughman KA, Shadmehr R (2000) Learning of Action Through Adaptive Combination of Motor Primitives. Nature 407:742–747.

Todorov E, Jordan MI (2002) Optimal feedback control as a theory of motor coordination. Nat Neurosci 5:1226–1235.

Tong C, Wolpert DM, Flanagan JR (2002) Kinematics and Dynamics Are Not Represented Independently in Motor Working Memory: Evidence from an Interference Study. J Neurosci 22:1108–1113.

Trempe M, Proteau L (2010) Distinct consolidation outcomes in a visuomotor adaptation task: Off-line leaning and persistent after-effect. Brain Cogn 73:135–145.

Tseng Y, Diedrichsen J, Krakauer JW, Shadmehr R, Bastian AJ (2007) Sensory Prediction Errors Drive Cerebellum-Dependent Adaptation of Reaching. J Neurophysiol 98:54–62.

Turnham EJA, Braun DA, Wolpert DM (2011) Inferring Visuomotor Priors for Sensorimotor Learning. PLoS Comput Biol 7:e1001112.

Turnham EJA, Braun DA, Wolpert DM (2012) Facilitation of learning induced by both random and gradual visuomotor task variation. J Neurophysiol 107:1111–1122.

Vaca-Palomares I, Díaz R, Rodríguez-Labrada R, Medrano-Montero J, Vázquez-Mojena Y, Velázquez-Pérez L, Fernandez-Ruiz J (2013) Spinocerebellar Ataxia Type 2 Neurodegeneration Differentially Affects Error-Based and Strategic-Based Visuomotor Learning. The Cerebellum 12:848–855.

Wagner MJ, Smith MA (2008) Shared Internal Models for Feedforward and Feedback Control. J Neurosci 28:10663–10673.

Wagner U, Gais S, Haider H, Verleger R, Born J (2004) Sleep inspires insight. Nature 427:352–355.

Walker MP, Brakefield T, Morgan A, Hobson JA, Stickgold R (2002) Practice with Sleep Makes Perfect: Sleep-Dependent Motor Skill Learning. Neuron 35:205–211.

Waters-Metenier S, Husain M, Wiestler T, Diedrichsen J (2014) Bihemispheric Transcranial Direct Current Stimulation Enhances Effector-Independent Representations of Motor Synergy and Sequence Learning. J Neurosci 34:1037–1050.

Wei K, Körding K (2009) Relevance of Error: What Drives Motor Adaptation? J Neurophysiol 101:655–664.

Werner S, Bock O (2010) Mechanisms for visuomotor adaptation to left–right reversed vision. Hum Mov Sci 29:172–178.

White O, Diedrichsen J (2013) Flexible Switching of Feedback Control Mechanisms Allows for Learning of Different Task Dynamics. PLoS ONE 8:e54771.

Wilhelm I, Metzkow-Mészàros M, Knapp S, Born J (2012) Sleep-dependent consolidation of procedural motor memories in children and adults: the pre-sleep level of performance matters. Dev Sci 15:506–515.

Willingham DB (1998) A Neuropsychological Theory of Motor Skill Learning. Psychol Rev 105:558–584.

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

Wolpert DM, Miall RC, Kawato M (1998) Internal models in the cerebellum. Trends Cogn Sci 2:338–347.

Wright DL, Rhee J-H, Vaculin A (2010) Offline Improvement during Motor Sequence Learning Is Not Restricted to Developing Motor Chunks. J Mot Behav 42:317–324.

Yousif N, Diedrichsen J (2012) Structural learning in feedforward and feedback control. J Neurophysiol 108:2373–2382.

Zhang M, Barash S (2000) Neuronal switching of sensorimotor transformations for antisaccades. Nature 408:971–975.