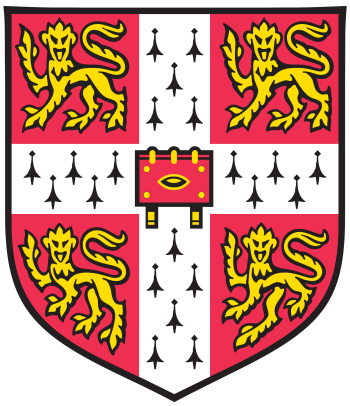
Sensorimotor learning under switching dynamics



James Barry Heald

Department of Engineering

University of Cambridge

This thesis is submitted for the degree of

Doctor of Philosophy

Selwyn College

September 2019

Declaration

I hereby declare that this thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as specified below. No substantial part of this thesis has already been submitted or is being concurrently submitted for any degree, diploma or other qualification at the University of Cambridge or any other university or similar institution. This thesis contains fewer than 65,000 words including appendices, bibliography, tables and equations and has fewer than 150 figures.

Chapter 2 has been published as

> Heald, J. B., Franklin, D. W., & Wolpert, D. M. (2018). Increasing muscle co-contraction speeds up internal model acquisition during dynamic motor learning. *Scientific Reports*, *8*(1).

Chapter 3 has been published as

> Heald, J. B., Ingram, J. N., Flanagan, J. R., & Wolpert, D. M. (2018). Multiple motor memories are learned to control different points on a tool. *Nat Hum Behav*, *2*(4), 300-311.

Chapters 4, 5 and 6 are a collaboration with Daniel Wolpert and Maté Lengyel.

Abstract

Sensorimotor learning under switching dynamics

James Barry Heald

Humans have a remarkable capacity to learn new motor behaviours without forgetting old ones. This capacity relies on the ability to acquire and express multiple motor memories without interference. Here we combine behavioural experiments and computational modelling to investigate how the sensorimotor system uses contextual information to create, update and recall motor memories. We first examine the role of muscle co-contraction in the learning of novel movement dynamics. We show that muscle co-contraction, as measured by surface electromyography, accelerates motor learning. We then explore the role of control points on objects in the formation of motor memories during object manipulation. We show that opposing dynamic perturbations, which interfere when controlling a single location on an object, can be learned when each is associated with a separate control point. To account for these results, we develop a parametric switching state-space model, in which the association between cues (control points) and contexts (dynamics) is learned from experience rather than fixed. We then extend this model to a Bayesian nonparametric switching state-space model, in which the number of contexts and cues are learned online rather than specified in advance. This model can instantiate new memories when novel perturbations are experienced and exhibits spontaneous recovery of a memory that has been ostensibly overwritten. To test the model, we perform an experiment in which we briefly present a previously experienced perturbation after behaviour has returned to baseline. As predicted, we observe a qualitatively distinct and more pronounced form of recovery, which we refer to as evoked recovery. Finally, we investigate Bayesian context estimation using single-trial learning. We show that people are able to learn novel associations between cues and contexts and that they use both contextual cues and state feedback to infer the current context and partition learning between memories. Taken together, these findings further the understanding of the behaviour and computational principles of sensorimotor learning under switching dynamics.

Acknowledgements

I would like to thank Daniel Wolpert first and foremost for giving me this opportunity and for his excellent supervision. James Ingram provided invaluable support throughout my PhD with the design and implementation of experiments. Goda Žalalytė and Alex Pantelides kindly ran many of my experiments. Daniel McNamee encouraged me to develop normative models of learning. Innumerable people in the Computational and Biological Learning Lab provided inspiration and friendship. The Engineering and Physical Sciences Research Council funded this PhD.

To my dad.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

Humans exhibit an enormous repertoire of goal-directed behaviors and can skillfully manipulate a wide range of objects. On a typical day, a person may pour a cup of coffee, tie their shoe laces and ride a bike, all with a level of dexterity and ease that belies the difficulty inherent in these tasks. Such complex behaviours are not innate but must be learned through interaction with the world. An important component of learning is the ability to compensate for the dynamics of the body (the mapping from muscle forces to motions of the joints). The dynamics of the body can change over time, both slowly (e.g., as the body develops) and quickly (e.g., as muscles fatigue). Therefore, to maintain skilled motor performance, a person must adapt their behaviour on multiple timescales throughout their lifetime. The dynamics of the body can also switch abruptly, such as when a person grasps an object, enters a swimming pool or steps on ice. To cope with switching dynamics, a person could learn multiple motor behaviours and switch between them as appropriate. The ability to learn new motor behaviours without forgetting old ones is central to this endeavour.

In the thesis, we will focus primarily on learning studies of goal-directed reaching movements. Our focus will be on adaptation to novel dynamic environments that alter the relation between actions and their consequences. We will examine the role that muscle co-contraction plays in dynamic motor learning (Chapter 2) as well as the role

of control points on objects in allowing multiple motor memories to be learned during object manipulation (Chapter 3). This will lead us to propose a new model of how multiple memories are instantiated and updated, which can explain existing data such as spontaneous recovery (Chapter 4). We then test predictions of the model related to spontaneous recovery (Chapter 5) and Bayesian context estimation (Chapter 6).

In this chapter we provide a review primarily of the behavior and computational underpinnings of sensorimotor learning. We start by reviewing two key components of sensorimotor processing: state estimation and control. We then review learning algorithms including error-based and reward-based learning. Finally, we focus on experiments, models and neural underpinnings of dynamic motor learning.

1.1 Internal models

A pre-requisite for control is to know the state of the body, for example, the body's joint angles and joint velocities. The rare person who loses their proprioceptive sense is severely incapacitated in their ability to perform skilled action (Rothwell et al., 1982). In humans and other animals, the state of the environment and body is not observed directly but must be inferred from sensory observations corrupted by noise at various sites along the neural pathway (e.g., transduction, action potential generation, synaptic transmission) (Faisal et al., 2008). Bayesian inference provides a principled way to estimate state so as to reduce uncertainty in a statistically optimal manner. Given a prior belief $p(x)$ about a latent state x (e.g. hand location) and a likelihood $p(y|x)$ describing the probability of an observation of sensory input y conditioned on the latent state, posterior inference can be performed via Bayes rule

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}, \tag{1.1}$$

where $p(y) = \int p(y|x)p(x)dx$ is a normalisation constant. Several lines of behavioural evidence suggest that humans approximate this inference procedure. First, when participants are instructed to report the value of a latent stimulus (e.g. the displacement of a cursor, the orientation of a line), their estimates are systematically biased away from the true value of the stimulus and towards the mean of the prior. Moreover, the magnitude of this bias increases with the level of noise in the stimulus (e.g. blur), consistent with optimal integration of the prior and the likelihood (Körding and Wolpert, 2004). Similar Bayesian processing has also been observed in a wide variety of studies including force estimation (Körding et al., 2004) and timing judgements (Jazayeri and Shadlen, 2010; Miyazaki et al., 2005). Although behavior tends to strive for Bayesian processing there appears to be approximation and limits to full Bayesian processing (Acerbi et al., 2014; Fiser et al., 2010). For example, while people are capable of internalising simple statistics of stimuli and noise (e.g., first and second moments of Gaussian distributions), higher-order statistical features (kurtosis, multimodality) of distributions seem much harder to acquire (Acerbi et al., 2012). Although human appear Bayesian there is still debate about the extent to which this behavior actually reflects Bayesian processing on individual trials rather than on average (Laquitaine and Gardner, 2018).

Rather than combine a prior with sensory evidence, statistical inference can also be used to infer a latent state of the environment (e.g., the height of an object) using multiple senses (e.g., vision and touch—a process known as cue combination). To perform optimal cue combination, each estimate should be weighted by the inverse variance (precision) of the noise in that sense; this yields the minimum-variance unbiased estimator. Evidence from psychophysics has demonstrated that humans do indeed perform cue combination in this way (Ernst and Banks, 2002; Jacobs, 1999; van Beers et al., 1999), which demonstrates that people possess well-calibrated estimates of the statistics of noise in different sensory modalities.

The state of the body and environment changes over time and to estimate the time-varying state of the motor plant, sequential Bayesian filtering can be performed such as by Kalman Filtering (Kalman, 1960; Kalman and Bucy, 1961). This involves predicting the next state of the motor plant and then updating the prediction once an observation has been made. The prediction step can use a probabilistic forward dynamics model, which maps the current motor command (i.e., efference copy) and state (e.g., position and velocity of the limb) to a distribution of next states (Haruno et al., 2001; Wolpert and Kawato, 1998). The predicted next state can in turn be mapped through an observation model to generate a corollary discharge (i.e., predicted sensory feedback, such as from joints or muscles). There are several advantages to having such an internal forward model. First, it reduces the variance of state estimates, increasing the robustness of control. Second, it mitigates delays inherent in the sensorimotor loop, which can be on the order of 80–150 ms for proprioceptive to visual feedback; these delays make feedback control too slow for rapid movements. With a forward model, a prediction of the current state can be used for rapid online control before sensory feedback is available.

Signatures of a forward model can be appreciated in hand localisation tasks performed at the end of reaching movements in the absence of vision. For example, the bias and variance of estimates of the position of the hand (as a function of movement duration and external forces) are consistent with sequential Bayesian filtering using a probabilistic forward model (Wolpert et al., 1995). In addition, when participants perform hand localisation after first adapting to a visuomotor rotation, their estimates of the position of their hand are shifted in the same direction as the rotation (Izawa and Shadmehr, 2011), consistent with forward model recalibration.

A forward model also allows the sensory consequences of actions to be anticipated and cancelled from incoming sensory feedback. This process—known as sensory cancellation—increases the salience of events caused by the environment, which tend to be more behaviorally relevant (for a review see Crapse and Sommer, 2008). For example,

the projection of the world onto the retina shifts during eye movements. Despite this shift, visual perception remains stable. This occurs as the superior colliculus sends an efference copy of the eye movement command to the frontal eye field, causing the visual receptive fields of neurons to shift (Sommer and Wurtz, 2006). Sensory cancellation has also been demonstrated for touch. In somatosensory cortex, the response to a self-produced tactile stimulus is attenuated compared to the response to an externally-produced tactile stimulus (Blakemore et al., 1998). Furthermore, participants perceive a self-produced tactile stimulus to be less ticklish than an externally-produced tactile stimulus (Blakemore et al., 2000). Sensory cancellation is also believed to underlie force attenuation to self-generated action (Bays et al., 2006, 2005; Shergill et al., 2003).

Finally, a forward model can be used to simulate state-action-reward trajectories for use in planning, physical reasoning and offline reinforcement learning. For example, knowledge of state transitions can be used to update a value function via backups, which transfer value information back to a state from its successor states (Sutton and Barto, 2018). A forward model can also be used to directly optimise the parameters of a policy by performing gradient ascent on the expected return (Deisenroth and Rasmussen, 2011), thus reducing sample complexity and increasing data efficiency.

1.1.1 Control

We now examine the type of control that the sensorimotor system might use including inverse models and optimal feedback control as well as stiffness.

Inverse models and optimal feedback control

Classic motor control theory suggests that forward models operate in parallel to a separate class of internal model known as inverse models, which compute the motor command required to achieve a desired state. To execute a movement using an inverse

model, the movement must first be specified as a desired state trajectory (i.e., a set of desired states indexed by time) (Haruno et al., 2001; Kawato et al., 1987; Wolpert and Kawato, 1998). Early work on unconstrained point-to-point reaching movements suggested that the desired state trajectory could be obtained by optimising a kinematic cost function, such as the squared jerk of the hand (Flash and Hogan, 1985), although it is not clear how this method generalises to more complex movements, such as grasping the handle of a cup. During execution of a movement, deviations from the desired state (e.g., due to noise or external perturbations) can be corrected either by passing the current state as input to the inverse model (e.g., Haruno et al., 2001) or by incorporating a separate feedback controller (e.g., a PID controller) (e.g., Kawato, 1990; Kawato et al., 1987; Wolpert and Kawato, 1998). The goal of corrections is to restore the state to its desired path. Crucially, however, the notion of a desired state trajectory is not supported by recent behavioural studies of goal-direct reaching (Nashed et al., 2012, 2014). For example, when reaching to either a small circular target or an elongated rectangular bar, participants tend to move in a straight line to the nearest point on the target/bar. However, if the movement is physically perturbed, participants reaching to the circle make a correction, whereas participants reaching to the bar allow themselves to be redirected to other locations along the bar (Nashed et al., 2012). Goal-directed corrections of this nature can be explained by alternative frameworks such as optimal feedback control and reinforcement learning, in which motor commands are generated to optimise a cost/reward function, rather than produce a pre-specified desired trajectory.

Optimal control theory (OCT) aims to find an sequence of motor commands that optimise a cost function under boundary conditions that define a movement (e.g., start and end state, movement duration). Originally, in the field of motor control, OCT was applied in an open-loop manner so as to generate a feedforward sequence of states or commands to minimise a cost. Cost functions such as integrated jerk (Flash and Hogan, 1985), integrated change in torque (Uno et al., 1989) or integrated endpoint variance (Harris and Wolpert, 1998) can successfully reproduce invariant features of

movements, such as bell-shaped speed profiles, Fitts's law and the 2/3 power law. Crucially, however, OCT cannot predict how a person will respond to perturbations. This shortcoming is addressed by optimal feedback control (OFC), which extends OCT by incorporating sensory feedback into the control process. Rather than finding an optimal sequence of feedforward motor commands, OFC finds an optimal policy (i.e., a mapping from the current state to a motor command, or time-varying feedback controller) that optimises ecologically relevant costs such as task accuracy and effort. A common cost function used in OFC is a quadratic function of states and control signals.

In OFC, the motor system interacts with the environment in a closed loop by generating motor commands and observing sensory feedback (Fig. 1.1). The motor command at each time step depends on the current state of the body and the environment, which must be estimated from time-delayed and noisy sensory feedback. To reduce the variance of the state estimate, the next state is predicted via an internal forward model that takes an efference copy of the motor command as input. When sensory feedback is observed, the predicted state is updated. The updated state estimate is used to generate a motor command from the optimal control policy. The state of the body and the environment then undergo a transition due to dynamics that depend on the motor command, and so the cycle continues.

OFC was first proposed as a solution to the problem of coordinating multiple effectors (e.g., muscles, joints, limbs) to achieve a common goal (Todorov and Jordan, 2002), which exists as the human body possesses more kinematic degrees-of-freedom than are required to perform any particular task. A key emergent property of the OFC framework is the *minimal intervention principle*, which states that deviations from the average trajectory (e.g., due to noise or external perturbations) should only be corrected if they degrade task performance. This conservative and goal-directed approach to error correction avoids unnecessary effort expenditure and reduces the chance of inadvertently worsening performance via motor noise. Due to the minimal intervention principle,

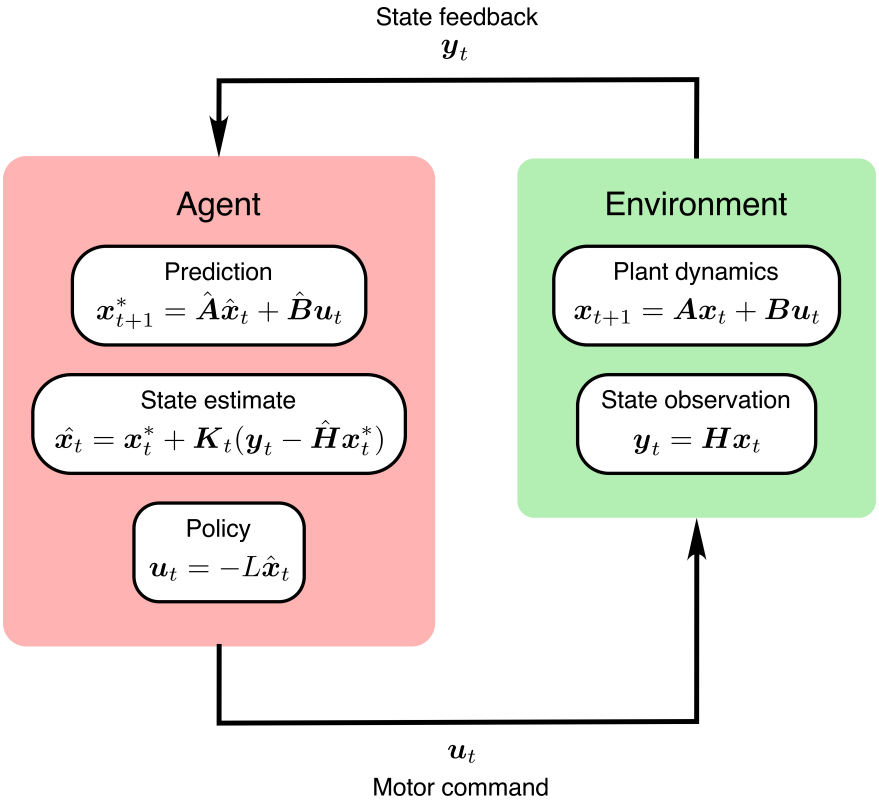


Figure 1.1. Schematic of the optimal feedback control loop. The agent interacts with the environment by generating motor commands ($\boldsymbol{u}$) and observing sensory feedback ($\boldsymbol{y}$). To estimate the state ($\boldsymbol{x}$) of the body and the environment, the agent first predicts the next state ($\boldsymbol{x}^*$) via an internal forward model and then updates this prediction ($\hat{\boldsymbol{x}}$) once state feedback has been observed. The size of the update depends on the Kalman gain ($\boldsymbol{K}$) and the prediction error ($\boldsymbol{y} - \hat{\boldsymbol{H}}\boldsymbol{x}^*$). The estimate of the state is used to generate a motor command via a control policy (L). The optimal control policy minimises a task-relevant cost function, $J[q(\boldsymbol{x}) + r(\boldsymbol{u})]$, where $q(\boldsymbol{x})$ encodes the goal state and $r(\boldsymbol{u})$ penalises effort. At the next time step, the body and the environment transition to a new state according to dynamics parameterised by $\boldsymbol{A}$ and $\boldsymbol{B}$.

motor variability accumulates in task-irrelevant dimensions while being reduced in task-relevant dimensions. This induces structured variability in goal-directed tasks, which has been well documented (Diedrichsen, 2007; Domkin et al., 2002; Latash et al., 2002; Scholz and Schöner, 1999; Valero-Cuevas et al., 2009). As an example, when participants control a single cursor at the midpoint of both hands, the variability in the endpoint location of each hand is greater than the variability in the endpoint location of the cursor (Diedrichsen, 2007). Indeed, the endpoint locations of the left and the

right hand are negatively correlated as deviations in one hand are corrected (in part) by the other hand. Such correlations between effectors, which emerge naturally in OFC as a consequence of optimising state and effort costs, have previously been interpreted as evidence for motor synergies (Santello and Soechting, 2000). In addition to these findings, OFC can also explain the distribution of work across effectors (Diedrichsen et al., 2010a) and the kinematics of movements in velocity-dependent curl fields (Izawa et al., 2008).

Stiffness

Although learning an inverse model or applying OFC can specifically compensate for the dynamics of the body and environment so as to be optimal, other less specific control strategies are also possible. For example, to compensate for an unpredictable perturbation muscles can co-contract to stiffen the limb so as to minimize deviations of the arm independent of the direction of a force perturbation. Such as strategy can reduce movement variability, by increasing the gain of the response to destabilising or unexpected perturbations (Hogan, 1985; Mussa-Ivaldi et al., 1985). In addition, using stiffness as a form of control can reduce time delays as the response is driven by the properties of the muscles.

Many studies have examined stiffness of the arm measured at the hand both during static posture as well as during movements in stable and unstable force fields. Stiffness at the hand can be quantified by an ellipse, which describes the relation between displacement and restoring force vectors in different directions in the horizontal plane. The magnitude, shape and orientation of this ellipse are primarily modulated through co-contraction of antagonistic muscles (Burdet et al., 2001; Franklin et al., 2007), though static limb posture also plays a role (Flash and Mussa-Ivaldi, 1990; Mussa-Ivaldi et al., 1985).

Stiffness control offers a non-specific strategy to compensate for perturbations until a more specific and efficient pattern of muscle activation can be learned. For example, when participants are first exposed to a velocity-dependent force field, rapid co-contraction of arm and chest muscles causes a global increase in stiffness (Franklin et al., 2003) and such increases in muscle activity scale up responses to mechanical perturbations (Pruszynski et al., 2009). As adaptation proceeds, muscle co-contraction decreases on a time course that correlates with the reduction in kinematic error, suggesting that the error signals themselves could drive co-contraction (Franklin et al., 2008). Increasing stiffness during dynamic learning may have a number of effects. As increased stiffness reduces kinematic errors it might reduce learning, but as stiffness keep the arm close to the final adapted paths it may also increase learning. In Chapter 2 we will ask how artificially increasing or decreasing stiffness changes the rater of internal model formation during dynamic learning.

Not only can overall stiffness be modulated but the orientation of the stiffness ellipse can also be controlled to some extent. For example, when moving in a force field that is unstable in one direction (an inverse spring that acts in one direction), the stiffness ellipse has been shown to expand only in the task-relevant direction (Burdet et al., 2001; Franklin et al., 2007). This strategy can limit the increase metabolic cost compared to a global increase in stiffness while maintaining task performance.

Muscle co-contraction also reduces movement variability in the absence of perturbations. Such an observation may seem counterintuitive given that muscle force variability scales with the mean force level (Harris and Wolpert, 1998; Jones et al., 2002). However, force variability increases at a slower rate than joint impedance as a function of muscle co-contraction (Selen et al., 2005). Hence, muscle co-contraction has a net stabilising effect on endpoint kinematics (Selen et al., 2009). This stabilising effect provides a normative explanation for why muscle co-contraction increases i) towards the end of movements, where noise has a greater effect on endpoint error and task success (Gomi and Kawato, 1996); ii) when spatial accuracy requirements are high (Gribble et al.,

2003; Lametti et al., 2007); and iii) during high velocity movements that require large noisy joint torques (Bennett, 1993; Suzuki et al., 2001).

1.2 Learning

Motor learning refers to improvement in performance of a sensorimotor behaviour, typically as a result of practice. Broadly speaking, sensorimotor learning can be divided into three main categories: adaptation, skill learning and sequence learning (Krakauer and Mazzoni, 2011).

Adaptation is the process of internal model recalibration so that errors are reduced. This process is essential to maintain accurate motor performance in the face of a changing body and environment. For example, during development, bone length and muscle mass change; during exercise, muscles fatigue; and during disease, muscles may weaken. Many different types of adaptation have been studied in the laboratory including reaching under perturbations such as visuomotor displacements (including prisms) (Krakauer et al., 1999, 2000) and robotic-induced force fields (Shadmehr et al., 2010), saccadic adaptation of eye movements to target jumps (Pélisson et al., 2010), and adaptation of walking patterns on a split-belt treadmill with different belt speeds (Roemmich and Bastian, 2018). In these studies, a perturbation is applied either to the dynamics or kinematics of the body or environment. This initially results in worsening of performance relative to baseline. However, over repeated trials, performance tends to improve and gradually returns towards baseline. If the perturbation is suddenly removed following adaptation, erroneous movements are made in the opposite direction to the initial perturbation—a phenomenon known as an after-effect (Lackner and Dizio, 1994; Lackner and DiZio, 1996).

Adaptation ensures the performance of a learned behaviour (e.g., tennis) is maintained when the environment changes (e.g., altered playing conditions, new tennis racquet,

physical exhaustion). However, adaptation is only one component of motor behavior. Knowing the dynamics of a tennis racquet is insufficient to be a skilled player, which require among others thing decisions on how to swing the racquet and what strategies to use. This is where skill learning comes in. Everyday skills such as playing tennis, tying your shoelaces or riding a bike involve complex spatiotemporal coordination of effectors (e.g., arm, eye and leg muscles) guided by multimodal sensory feedback (e.g., vision, proprioception and touch). Skill learning proceeds by trial-and-error exploration of the solution space under feedback provided by the environment. This process can be accelerated by observing an expert or demonstrations from a teacher, who can also provide explicit instructions and feedback. The initial phases of skill learning often require conscious cognitive input, but with practice, the skill becomes automatic and independent of conscious cognitive input. Classic assays of skill learning include the pursuit rotor task, the serial reaction time task and the mirror tracing task, all of which require complex hand-eye coordination Krakauer and Mazzoni (2011). More recently, researchers have explored more naturalistic tasks such as carrying a cup of coffee (Nasseroleslami et al., 2014), rhythmically bouncing a ball (Ronsse et al., 2010) and playing the throwing game of skittles (Zhang et al., 2018). Skill learning has been operationally defined as a shift in the speed-accuracy trade-off curve and reduction in variability. Moreover, variability reduction has been further decomposed into i) the discovery of error-tolerant solutions; ii) the covariance of execution variables to minimise the effect of variability on the task; and iii) reduction in the variance of noise Müller and Sternad (2004).

Skilled movements are often composed of discrete sub-movements ordered in a particular sequence. Consider, for example, playing the piano. In the motor domain, sequence learning has been investigated using a variety of tasks, most notably the serial reaction time task (Nissen and Bullemer, 1987; Willingham et al., 1989) and the finger-tapping task (Kami et al., 1995; Willingham et al., 1989). These two tasks emphasise different components of sequence learning. The serial reaction time task examines the acquisition of the order of elements in a sequence, and learning is demonstrated as a reduction

in the interval between the presentation of a stimulus and the selection of the correct response. The finger-tapping task examines the ability to execute an entire sequence quickly and accurately. An increase in the number of correct repetitions of the sequence in a fixed duration provides a signature of learning.

1.2.1 Computation

In this thesis, we focus primarily on motor adaptation. We now review the computations thought to be associated with sensorimotor adaptation.

Sensorimotor-error-based learning

Motor adaptation is believed to depend on error-based recalibration of an internal forward model (Tseng et al., 2007). Recalibration is computationally straightforward, as the target values of the forward model—the sensory consequences of an action—are directly sensed. Therefore, supervised learning can be performed: the forward model predicts the sensory consequences of an action (Blakemore et al., 2000; Mehta and Schaal, 2002; Shergill et al., 2003), and any discrepancy between the actual and predicted sensory consequences is used to compute a sensory prediction error and train the forward model (e.g., using standard gradient descent methods). Forward model recalibration alters the perceived location of the hand at the end of a reaching movement (Izawa and Shadmehr, 2011; Modchalingam et al., 2019).

How forward model recalibration leads to adaptation of motor behaviour is still an open question (Shadmehr et al., 2010). In the framework of optimal feedback control, the optimal control policy could be recomputed every time the forward dynamics model is updated (Izawa et al., 2008; Liu and Todorov, 2007). This would produce incremental changes in behaviour that accompany the gradual recalibration of the forward model (Mitrovic et al., 2010). Similarly, in model-based reinforcement learning, the updated

forward dynamics model could be used to calculate the change in policy that would improve the expected reward, where reward can be defined, for example, in terms of task accuracy and effort (Deisenroth and Rasmussen, 2011). Alternatively, in classical sensorimotor control theory, it has been suggested that an inverse controller could be updated using standard supervised learning techniques via the forward model to minimise the difference between the desired and actual sensory consequences (Jordan and Rumelhart, 1992).

Reward-based learning

Motor learning can also be driven by rewards and punishments in the absence of sensory prediction errors. Unlike error-based learning, reward-based learning does not lead to forward model recalibration and hence does not alter the perceived location of the hand (Izawa and Shadmehr, 2011). Rather, reward-based learning reinforces successful actions (Huang et al., 2011). Reward feedback can be used to learn gradual (Izawa and Shadmehr, 2011) and abrupt (Nikooyan and Ahmed, 2014) rotations, as well as more difficult tasks. For example, scalar reward feedback can be used to learn a curved or sinusoidal reaching movement without visual feedback about the hand trajectory or the rewarded shape (Wu et al., 2014).

Because scalar rewards do not provide information about the behavioural change required to increase reward, the learner must explore the action space. One way to explore the action space is through movement variability. In tasks that rely heavily on exploration, movement variability—far from being a hindrance—can actually accelerate learning (Herzfeld and Shadmehr, 2014). Indeed, in the curve tracing task mentioned above (Wu et al., 2014), participants with higher baseline variability exhibited faster learning. Movement variability can be modulated by reward incentives. For example, in reward-based learning tasks, movement variability can be upregulated to promote exploration (Wu et al., 2014), whereas outside of learning, movement variability can

be suppressed to improve task accuracy to obtain rewards (Manohar et al., 2015). Taken together, these results suggest that humans can activity regulate their movement variability to navigate the trade-off between exploration and exploitation and maximise long-term reward.

Reward-based learning and error-based learning can interact to produce greater learning than either alone. For example, the addition of reward feedback to vision accelerates learning (Nikooyan and Ahmed, 2014), leads to greater retention (Galea et al., 2015) and can even overcome forgetting of an adapted movement (Shmuelof et al., 2012). Punishment feedback can accelerate learning even more than reward feedback, although there are no benefits to retention (Galea et al., 2015). Omitting a reward—a type of punishment—can also induce retrieval of a previously rewarded behaviour (Pekny et al., 2011). These results suggests that different mechanisms underpin learning and retention in motor adaptation, and that motor adaptation can be influenced by motivational feedback.

The notion of reward-based learning poses the question: what constitutes reward? In a typical reward-based learning paradigm, the experimenter provides explicit rewards or punishments (e.g., points, money, aversive stimuli). However, hitting a target and expending little energy (i.e., successfully completing a task) may also provide an implicit reward. Indeed, in an adaptation setting, it has been suggested that the final adapted movement is reinforced in this way (Huang et al., 2011). This raises the possibility that reward is a composite function of both explicit and implicit factors.

Use-dependent learning

Recently, a third type of motor learning has been identified that makes movements more similar to past movements. This phenomenon, called use-dependent learning, was first observed in planar movements with task-redundant dimensions (Diedrichsen et al., 2010b). In a passive condition, participants had their hand guided along a tilted

trajectory to an elongated bar. Because all points on the bar were equally-valid goals, the titled motion occurred in a task-redundant dimension. Subsequent free movements were biased in the same direction as the imposed tilt, demonstrating use-dependent learning. The same bias was also seen when participants made active movements to the bar; this time, the tilted trajectory was imposed using a force channel (a spring wall that constrains movements to a straight line). When the force channel was removed, free movements were initially biased away from the imposed tilt—an after-effect of error-based learning. However, after several trials, the movements became biased towards the imposed tilt, revealing use-dependent learning. This study demonstrated that use-dependent learning and error-based learning can occur simultaneously, and that use-dependent learning is retained for longer than error-based learning.

Shortly afterwards, it was shown that use-dependent learning can also be induced along a task-relevant dimension of a movement in the setting of adaptation (Huang et al., 2011). Participants reached to multiple targets under different visuomotor rotations. Crucially, the relationship between the target locations and the visuomotor rotations was such that the ideal movement in hand space was the same for all target-rotation pairs. Hence, adaptation promoted repetition of this movement. This repetition induced a directional bias towards the repeated movement, which was comparable for trained and untrained targets and increased in absolute size for farther away targets. The authors hypothesised that successful error reduction produced an implicit reward that reinforced the ideal movement. Interestingly, this implies that use-dependent learning may in fact be reward-based learning. Use-dependent learning has also been interpreted as a bias induced by a prior belief over reach directions in a Bayesian framework (Verstynen and Sabes, 2011).

1.2.2 Dynamic learning

Skilled movement depends on the ability to compensate for the dynamics of the body and external perturbations. How this ability is acquired has been the focus of much research. So-called dynamic learning is typically investigated by having participants grasp the handle of a robotic manipulandum that can generate forces that depend on the state of the hand (Shadmehr and Mussa-Ivaldi, 1994). This imposes a physical perturbation that alters both the visual and proprioceptive consequences of motor commands. Robotic manipulanda have been used to simulate naturalistic object dynamics (Dingwell et al., 2002; Howard et al., 2009; Krakauer et al., 1999; Nasseroleslami et al., 2014) as well as more unusual dynamics. For example, one perturbation that has been extensively studied is a velocity-dependent curl field, which generates forces on the hand that act perpendicular to the velocity of the hand and proportional to its speed. Dynamic learning has also been studied using non-contact Coriolis forces by having participants make reaching movements in a rotating room (Lackner and Dizio, 1994; Lackner and DiZio, 1996).

Savings

It is often faster to relearn a task the second time around—a phenomenon known as savings (Brashers-Krug et al., 1996; Caithness et al., 2004; Klassen et al., 2005; Krakauer et al., 2005; Shadmehr and Brashers-Krug, 1997). Typically, savings is demonstrated by having a participant learn a task (e.g., a force field), and then, following either de-adaptation (i.e., washout) or an interval of time (e.g., 24 hours), relearn the same task. If the learning and relearning sessions are separated by time alone, savings can be induced by as few as five trials (Huberdeau et al., 2015) and can be observed after an interval of five months (Shadmehr and Brashers-Krug, 1997). In general, savings are more pronounced when a task is performed for longer (Huberdeau et al., 2015; Roemmich and Bastian, 2015), although this effect saturates. Interestingly,

if the learning session is divided into multiple smaller sessions separated by washout blocks, savings is enhanced (Day et al., 2018). This suggests that abrupt switches in the environment may play an important role in facilitating savings, perhaps by highlighting the presence of multiple contexts. Indeed, savings occurs when a task is introduced abruptly but not when it is introduce gradually (Roemmich and Bastian, 2015).

There have been several attempts to explain savings. One possibility is that the state of adaptation is not at baseline at re-exposure. This could be because some adaptation is retained by a slow learning process (see Modelling section below) (Körding et al., 2007b; Smith et al., 2006) or because the memory is protected from interference during de-adaptation (Criscimagna-Hemminger and Shadmehr, 2008; Oh and Schweighofer, 2018; Pekny et al., 2011). An alternative explanation is that the learning rate or retention parameters of the motor system are modified during exposure, enabling faster learning or greater retention on re-exposure (Herzfeld et al., 2014; Huang and Shadmehr, 2009). This explanation can account for the observation that savings occurs even after de-adaptation with a counter-perturbation or washout (Zarahn et al., 2008). Savings has also been attributed to reinforcement learning, such that successful adaptation reinforces the adapted movement, enabling it to be recalled on re-exposure (Huang et al., 2011).

Interference

When multiple conflicting tasks are learned in close temporal proximity, substantial interference often occurs (Brashers-Krug et al., 1996; Caithness et al., 2004; Gandolfo et al., 1996; Krakauer et al., 1999; Shadmehr and Brashers-Krug, 1997). Classically, interference has been studied using the A-B-A paradigm, where task A (e.g., a clockwise curl field) is learned first, followed by task B (e.g., a counterclockwise curl field) and then task A again. Two types of interference can be observed in this paradigm: anterograde

and retrograde interference. Anterograde interference occurs when experience of the first task A interferes with subsequent performance on task B and is assessed by comparing learning of task B with and without prior exposure to task A. Retrograde interference occurs when the experience of task B after A interferes with subsequent performance on task A (due to disruption of the previous memory of task A) and is assessed by comparing learning of task A before and after exposure to task B.

Early research found that when task B is performed immediately after task A, task A interferes with task B (anterograde interference), and no savings are observed when task A is relearned 24 hours later (retrograde interference). However, when task B is performed 6 hours after task A, task A does not interfere with task B, and savings are observed when task A is relearned 24 hours later (Brashers-Krug et al., 1996; Shadmehr and Brashers-Krug, 1997). These results provided evidence that task A had been consolidated in memory in the 6 hours after it had had first been learned; that is, the memory of task A had been transferred from a short-term labile state to a long-term fixed state that was resistant to disruption. A subsequent study found that if the amount of initial exposure to task A is doubled, interference can be prevented even if task B is performed as soon as 5 minutes after task A (Krakauer et al., 2005). Despite the appeal of the consolidation theory, one well-powered study failed to replicate these findings and found no evidence of consolidation (Caithness et al., 2004); for both visuomotor rotations and force fields (both position- and velocity-dependent), no evidence of savings was observed when task A was relearned, even though task B was performed 24 hours after task A. That is, task A and task B always interfered.

Early work on motor adaptation demonstrated that inertial loads (acceleration-dependent force fields) and position-dependent visuomotor rotations do not interfere (Krakauer et al., 1999). This was taken as evidence that movement dynamics and kinematics are represented by independent internal models (Krakauer et al., 1999). However, an alternative interpretation of this result is that perturbations interfere when they depend on the same kinematic variable. This was latter interpretation was

supported in an experiment that demonstrated interference between position-dependent force fields and position-dependent visuomotor rotations (Tong et al., 2002).

Generalisation

When people make reaching movements in a velocity-dependent curl field, learning acquired at trained movements generalises to untrained movements made in novel regions of the workspace. Importantly, learning decays smoothly as a function of the distance between the trained movement and untrained movements, revealing a fundamental property of motor adaptation, namely that it is local (Donchin et al., 2003; Gandolfo et al., 1996; Thoroughman and Shadmehr, 2000). Early studies suggested that learning was represented in intrinsic joint-based coordinates (Gandolfo et al., 1996; Malfait et al., 2002; Shadmehr and Mussa-Ivaldi, 1994). However, more recent work has revealed a more complex picture involving a mixture of intrinsic, extrinsic and object-centered coordinate frames (Berniker et al., 2013; Brayanov et al., 2012). Interestingly, learning does not always decay with generalisation. For example, when participants were trained to make reaching movements of a particular duration and amplitude in a velocity-dependent curl field, test movements performed at half the duration or twice the amplitude exhibited a linear extrapolation of forces as a function of velocity (Goodbody and Wolpert, 1998).

From a normative perspective, generalisation should reflect inductive biases about how dynamics vary across the state space (Williams and Rasmussen, 2006). For example, when dynamics vary slowly across state space, generalisation should be broad, whereas when dynamics vary rapidly, generalisation should be narrow. This principle was demonstrated in an experiment in which participants adapted to force fields of varying spatial frequency (Thoroughman and Taylor, 2005). Over minutes of training, the generalisation function became increasingly narrow for environments with higher spatial frequencies.

Context-dependent learning

The interference we have described above occurs when there are no cues to the dynamics other than the experience of the perturbation themselves. A number of studies have examined how context can enable multiple motor memories to exist by minimizing interference.

The dynamics of the body and the environment often change in a discrete manner, such as when we grasp an object, step onto ice or experience an injury. Such abrupt changes in dynamics occur during a context switch. To maintain accurate performance in such environments that switch between contexts, a person could learn multiple models of their environment and switch between these models whenever the context switches.

The dominant paradigm for investigating context-dependent learning is dual-adaptation in which two opposite perturbations (e.g., a clockwise and a counterclockwise curl field) are presented in either an alternating or random order. In contrast to adaptation to a single force field, which is almost complete (Shadmehr and Mussa-Ivaldi, 1994), adaptation to two opposing force fields is limited due to interference between the memories of each field (Brashers-Krug et al., 1996; Caithness et al., 2004; Gandolfo et al., 1996; Krakauer et al., 1999; Shadmehr and Brashers-Krug, 1997). However, if each force field is paired with an effective contextual cue, this interference can be reduced (Addou et al., 2011; Cothros et al., 2009; Howard et al., 2013; Osu et al., 2004a; Richter et al., 2004; Wada et al., 2003). For example, opposite force fields can be learned if each force field is associated with a different posture of the body (Gandolfo et al., 1996; Richter et al., 2004; Yeo et al., 2015) or a different end effector (Krakauer et al., 2006; Nozaki et al., 2006). Opposite force fields can also be learned if each force field is associated with a different lead-in or follow-through movement (Howard et al., 2012, 2015; Sarwary et al., 2015; Sheahan et al., 2016) or if each force field is associated with a different movement in visual space (Hirashima and Nozaki, 2012). It may even be possible to learn opposite force fields if they are presented in a predictable

sequence (Wainscott et al., 2005); although results on this question are conflicting (Karniel and Mussa-Ivaldi, 2002). On the whole, arbitrary visual cues such as color do not allow opposite force fields to be learned (Gandolfo et al., 1996; Howard et al., 2013). However, if a static visual cue provides an instruction to switch control policy (e.g., one-cursor vs. two-cursor control in a bimanual reaching task), opposite force fields can be learned (White and Diedrichsen, 2013). This suggests that the hallmark of an effective contextual cue may be its ability to engage different control policies before learning.

In Chapter 3 we examine whether the location on a tool can act as a contextual cue so that for the same movement of the tool, controlling different locations can be associated with difference motor memories.

Choosing which model to select and update on each time step, however, is not straightforward, as the context is often unknown. For example, when lifting a carton of milk, a person may not know whether the carton is full or empty (Wolpert and Ghahramani, 2000). One computational model that addressed context-dependent motor control and learning is the modular selection and identification for control (MOSAIC) model (Haruno et al., 2001; Wolpert and Kawato, 1998). This model postulates multiple context-specific pairs of forward (predictor) and inverse (controller) models known as modules. The goal of the MOSAIC model is to optimally partition control and learning between these modules on each time step using Bayesian context estimation. Before movement execution, sensory contextual cues are used to compute the prior probability of each context (e.g., empty or full carton). These prior probabilities then weight the outputs of each inverse model to compute the motor command (feedforward control). During movement execution, each forward model predicts the next state of the system based on the current state and the current motor command. This leads to a module-specific prediction error that is used to calculate the likelihood of each context. The prior and the likelihood are then combined using Bayes rule to calculate the posterior probability of each context. This posterior probability signals

the responsibility of each module and is used i) to re-weight the contribution of each inverse model to the final motor command (feedback control) and ii) to determine how much the forward and inverse models of each module should be updated based on their prediction error. During movement execution, context transition probabilities can also be learned and exploited to predict the current context using a hidden Markov model. In simulation, the MOSAIC model was able to learn to manipulate and switch between multiple objects, generalise to novel objects and perform online corrections when misleading cues were presented (Haruno et al., 2001; Wolpert and Kawato, 1998).

Importantly, the number of modules in the MOSAIC model is pre-specified, and there is no mechanism for adding or removing modules online with experience (e.g., as new objects appear or disappear). This is a significant limitation, as the number of tasks a person will perform in their lifetime is not fixed or known in advance. This issue was later addressed using an alternative architecture (locally weighted projection regression) to represent the inverse model as a weighted combination of linear models, where the number of linear models can grow as needed (Lonini et al., 2009). Using this architecture, this alternative model was able to learn to compensate for a novel velocity-dependent force field while still retaining a previously learned ability to manipulate three objects.

The question of whether new experiences overwrite existing memories or create new memories has recently been explored in sensorimotor learning (Oh and Schweighofer, 2018; Pekny et al., 2011; Roemmich and Bastian, 2015). One way in which this question has been investigated is by examining how people relearn a task they have previously learned and (ostensibly) unlearned via de-adaptation. In both visuomotor rotation (Herzfeld et al., 2014) and split-belt treadmill tasks (Roemmich and Bastian, 2015), faster re-adaptation (savings) has been observed when the initial perturbation was introduced abruptly but not when it was introduced gradually. One plausible interpretation of this result is that the abrupt introduction of the perturbation lead to the creation of a new memory and that apparent unlearning during washout was

in fact re-expression of the baseline memory (Oh and Schweighofer, 2018); according to this interpretation, the memory of the perturbation was preserved during washout, allowing faster re-adaptation. The impetus to create a new memory may have been provided by large errors, which the abrupt perturbation induced but the gradual perturbation did not. In support of this view, abrupt introduction of a large but not a small visuomotor rotation leads to savings (Morehead et al., 2015; Oh and Schweighofer, 2018). Interestingly, savings has also been observed in an experiment where a visuomotor rotation was introduced gradually but quickly (Turnham et al., 2011). This raises the possibility that the *rate of change* of a perturbation, which undergoes an abrupt step when the perturbation is introduced gradually, can also result in new memory creation.

In Chapter 4 we develop a novel normative model of learning that can automatically determine when to create a new memory vs. update existing memories, and in the latter case which memories and how to update them.

Modelling

State-space models have emerged as the dominant model of trial-by-trial motor adaptation (Albert and Shadmehr, 2017; Cheng and Sabes, 2006; Donchin et al., 2003; Thoroughman and Shadmehr, 2000). These models frame motor adaptation as sequential estimation of a task perturbation (e.g., a force field or visuomotor rotation). On trial t, an adaptive state (x_t) is updated by forgetting some of what has been learned and correcting for part of an error (e_t). The error is defined as the difference between the task perturbation (f_t) and the motor output: $e_t = f_t - x_t$. The amount of forgetting and learning are determined by a retention factor ($0 < a < 1$) and a learning rate ($0 < b < 1$), respectively. The update rule for the single-rate state-space model is

$$x_{t+1} = ax_t + be_t. \tag{1.2}$$

This basic state-space model can be generalised by adding additional states that forget and learn at different rates (Inoue et al., 2014; Smith et al., 2006). The most influential of these multi-rate models is the dual-rate model, which assumes that two hidden states, a slow state ($x_{\mathrm{s},t}$) and a fast state ($x_{\mathrm{f},t}$) , are updated in parallel. These states sum to give the motor output ($y_t = x_{\mathrm{s},t} + x_{\mathrm{f},t}$), which determines the error used for learning: $e_t = f_t - y_t$. The updates rules for the slow state and the fast state are

$$\begin{aligned} x_{\mathrm{s},t+1} &= a_{\mathrm{s}} x_{\mathrm{s},t} + b_{\mathrm{s}} e_t, \\ x_{\mathrm{f},t+1} &= a_{\mathrm{f}} x_{\mathrm{f},t} + b_{\mathrm{f}} e_t. \end{aligned} \tag{1.3}$$

where $a_{\mathrm{s}} > a_{\mathrm{f}}$ and $b_{\mathrm{s}} < b_{\mathrm{f}}$. These inequalities ensure that the slow state learns and forgets gradually and the fast state learns and forgets quickly. The dual-rate model is able to explain a range of adaptation phenomena that a single-rate model cannot capture, such as savings, anterograde interference and rapid unlearning (Smith et al., 2006). However, the phenomenon that is thought of as the strongest evidence for the dual-rate model is spontaneous recovery. Spontaneous recovery can be observed when a series of error-clamp trials are presented after an adaptation-extinction training paradigm which uses a prolonged period of field A followed by a short period of the opposite field (-A) to bring adaptation back to baseline. Although the motor output is at baseline at the start of the error-clamp phase, this arises from the combination of the slow state being negative and the fast state being positive. Because the fast state decays quicker than the slow state, the motor output exhibits spontaneous recovery with forces generated in the channel appropriate for field A. Interestingly, although the slow and fast states of the dual-rate model are hidden, they appear to map onto identified learning systems. For example, in visuomotor rotation experiments, the time courses of these states mirror implicit and explicit learning processes, respectively (McDougle et al., 2015; Taylor et al., 2014).

In Chapter 4 we will show how spontaneous recovery may not reflect a dual-rate model but instead reflects switching between existing motor memories, an idea we test and confirm in Chapter 5.

The state-space models presented so far also have Bayesian counterparts. Under the assumption that the task perturbation evolves as a linear-Gaussian dynamical system, optimal motor adaptation can be performed using a Kalman filter (Baddeley et al., 2003; Berniker and Kording, 2008; Burge et al., 2008; Korenberg and Ghahramani, 2002; Wei and Körding, 2010). Like the dual-rate model, the Kalman filter can also incorporate multiple states, allowing the learner to adapt to a body that changes over multiple timescales, such as during growth, fatigue and injury (Körding et al., 2007b). The Kalman filter model of motor adaptation makes a number of predictions pertaining to the rate of learning. First, learning should be slower under higher measurement noise. This has been confirmed in experiments where the feedback of a visual perturbation is blurred (Burge et al., 2008; Wei and Körding, 2010). Second, learning should be faster under higher state uncertainty. This has also been confirmed in an experiment where the volatility of a perturbation was varied (Burge et al., 2008), as well as an experiment where participants sat idle without vision prior to experiencing a perturbation (Wei and Körding, 2010).

Not all predictions of the Kalman filter model agree with experimental data. For example, whereas the learning rate of the Kalman filter is independent of error size, the motor system is more sensitive to smaller errors (Fine and Thoroughman, 2007; Marko et al., 2012; Wei and Kording, 2009), and learning saturates for large errors (Herzfeld et al., 2014; Robinson et al., 2003). These discrepancies have been accounted for by assuming multiple causes of errors, some of which are more relevant (i.e., worthy of correction) than others (Körding et al., 2007a). Accordingly, when an error is experienced, the learner must assign credit between relevant and irrelevant causes of the error. Under the assumption that larger errors are less likely to be relevant (e.g.,

caused by the motor plant), this framework can account for the observed dependence of learning rate on error size.

In environments that switch between multiple perturbations, the rate of learning adapts to the consistency of the environment (Castro et al., 2014; Herzfeld et al., 2014; Takiyama et al., 2015). For example, during rapid switching between a clockwise and a counterclockwise curl field, the rate of learning decreases, and vice versa for a slowly switching environment (Herzfeld et al., 2014). To explain this finding, the authors proposed a modified state-space model where the learning rate adapted online based on the signs of consecutive errors. A similar result was also observed in a study where the persistence of a single force field was varied. When the force field was presented for 1, 7 or 20 trials, the rate of learning decreased, stay the same or increased, respectively (Castro et al., 2014). In these experiments, the change in learning rate occurred over many tens to hundreds of trials, suggesting that the statistics of the environment were slowly internalised.

State-space models assume that the adaptive state decays to zero in the absence of error (Smith et al., 2006). However, adaptation often decays to a non-zero endpoint in error-clamp trials (Criscimagna-Hemminger and Shadmehr, 2008; Keisler and Shadmehr, 2010; Pekny et al., 2011; Vaswani and Shadmehr, 2013). This non-zero endpoint has been interpreted as a component of motor memory that is resistant to de-instantiation (Vaswani and Shadmehr, 2013). State-space models also assume that the adaptive state decays on every trial (Smith et al., 2006). This assumption was recently challenged following evidence of lags to decay when a change in context was masked (Vaswani and Shadmehr, 2013). However, these observed lags appeared to be an artifact of inappropriate analysis assumptions (Brennan and Smith, 2015). Under appropriate assumptions, decay appeared to be independent of context change detection, consistent with standard modelling assumptions (Brennan and Smith, 2015).

Neural correlates of control and learning

How the brain generates voluntary movement is a central question in motor control. The functional contribution of primary motor cortex to movement can be understood from the perspective of two models: the representational model and the dynamical systems model. The representational model proposes that single neurons in motor cortex encode movement-related parameters. Substantial controversy has centered around whether these parameters represent low-level dynamics variables such as force or high-level kinematics variables such as movement direction (for review see Omrani et al., 2017). Both these variables correlate with activity in motor cortex, although responses of individual neurons are often complex, labile and heterogeneous (Churchland and Shenoy, 2007; Scott, 2008). The dynamical systems model views motor cortex as a pattern generator for producing temporally and spatially complex patterns of muscle activity. This perspective emphasises two distinct periods in the production of movement. In the preparatory period, inputs to motor cortex drive the neural population to an initial state that subsequently determines the evolution of neural activity during the movement period (Churchland et al., 2010). These preparatory dynamics occur in a null space that does not produce movement (Kaufman et al., 2014). In the movement period, the state of the neural population exhibits oscillations that drive downstream multiphasic electromyographic signals (Churchland et al., 2012). Interestingly, neural network models trained to perform a motor task or reproduce muscle activity can exhibit both directional tuning at the single-unit level (Lillicrap and Scott, 2013) and oscillatory dynamics at the population level (Sussillo et al., 2015). This suggests that the dynamical systems model may offer a unifying view of motor cortex. Whether or not motor cortex implements an optimal feedback controller remains an open question (Scott, 2008).

Substantial evidence suggests that fast trial-by-trial motor adaptation depends on the cerebellum. The principal site of learning in the cerebellum is the synapse between parallel fibres and Purkinje cells (Albus, 1971; Marr and Thach, 1991). The Purkinje

cell is the sole output cell of the cerebellar cortex and receives convergent inputs from two main excitatory sources: parallel fibers and climbing fibers. It is thought that parallel fibers convey sensorimotor contextual information regarding the state of the body, the goal of the movement and an efference copy of the motor command, and climbing fibers convey sensory error signals in motor coordinates (Kitazawa et al., 1998; Kobayashi et al., 1998).

It has been widely hypothesized that the cerebellum is the anatomical locus of a forward model. Several lines of evidence from single cell recordings in monkeys support this hypothesis. First, the cerebellum receives the necessary inputs to implement a forward model, namely the current state of the plant and the motor command (Stein and Glickstein, 1992). Second, simple spike discharges of Purkinje cells of the cerebellum precede movement kinematics and encode the position, direction of movement and speed of arm movements (Hewitt et al., 2011; Roitman et al., 2005), consistent with a role in predicting the next state. Third, deep cerebellar output neurons encode sensory prediction errors—a key signal for learning a forward model—and these signals evolve on the same time course as behavioural learning during motor adaptation (Brooks et al., 2015). In addition, the cerebellar-like structure of the electric fish has a similar architecture to the cerebellum and is known to implement a forward model that cancels the predictable sensory consequences of self-generated electric fields (for review see Bell et al., 2008).

The role of the cerebellum as a forward model is also supported by clinical studies. For example, cerebellar patients have proprioceptive deficits during active movement, but not when the arm is moved passively (Bhanpuri et al., 2013), and cerebellar patients who adapt to a gradual visuomotor rotation do not show recalibration in a proprioceptive localisation task (Izawa et al., 2012). Further evidence implicating the cerebellum in forward state estimation also comes from functional neuroimaging (e.g., Blakemore et al., 1998) and non-invasive stimulation studies (Miall et al., 2007).

Chapter 2

Increasing muscle co-contraction speeds up internal model acquisition during dynamic motor learning

2.1 Introduction

The reduction of kinematic errors during reaching movements in the presence of novel dynamics occurs through at least two complementary mechanisms (Franklin et al., 2003; Osu et al., 2003). First, participants can learn an internal model of the dynamics and use this model to anticipate and compensate for perturbations (Conditt et al., 1997; Donchin et al., 2003; Lackner and Dizio, 1994; Shadmehr and Mussa-Ivaldi, 1994). Second, participants can use a less-specific strategy of stiffening up the limb through muscle co-contraction, that is use impedance control (Burdet et al., 2006; Hogan, 1985; Mitrovic et al., 2010; Shadmehr and Mussa-Ivaldi, 1994) to reduce the kinematic errors that result from perturbing dynamics (Burdet et al., 2001; Lametti et al., 2007; Osu et al., 2004b). Several studies have shown that both these mechanisms contribute to the early compensation for novel dynamics (Albert and Shadmehr, 2016; Darainy

and Ostry, 2008; Franklin et al., 2008; Milner and Franklin, 2005; Sing et al., 2009) and that muscle co-contraction decreases as the internal model is learned (Franklin et al., 2003, 2012; Huang et al., 2012; Thoroughman and Shadmehr, 1999). Therefore, muscle co-contraction offers a temporary, non-specific strategy to compensate for novel dynamics until they are learned by the internal model. Here we ask whether the increase in muscle co-contraction during exposure to novel dynamics affects the learning of the internal model.

Muscle co-contraction could affect dynamic motor learning in three contrasting ways. First, it could facilitate learning by keeping the state of the limb close to the desired state required for task success (i.e., the target). By resisting force-field displacements, muscle co-contraction increases the alignment between the actual motions experienced and the motion to be learned. Importantly, because participants learn the mapping from visited states, rather than planned states, to dynamics (Castro et al., 2011), increased muscle co-contraction should accelerate learning by concentrating adaptation within the region of state space associated with the planned motion. Indeed, in artificial neural networks, it has been shown that the inverse dynamics of a plant can be learned faster when the actual trajectories executed are closer to the desired trajectories to be learned, as this ensures that local error information is applicable to the region of state space to be learned (Sanger, 1994). Conversely, reduced muscle co-contraction should slow down learning by increasing the range of actual motions experienced and thus spreading adaptation more thinly across state space. Consistent with this idea, adaptation to a visuomotor rotation is less complete when the rotation is introduced abruptly compared with gradually (Kagerer et al., 1997).

Second, muscle co-contraction could impede learning by reducing kinematic variability (Burdet et al., 2001; Gribble et al., 2003; Selen et al., 2005; Wong et al., 2009), which has been shown in some forms to enhance error-based learning, both in humans (Wu et al., 2014) and in songbirds (Charlesworth et al., 2012; Kao et al., 2005; Tumer and Brainard, 2007). Moreover, muscle co-contraction reduces kinematic errors, which are

a key training signal for internal model acquisition (Fine and Thoroughman, 2007; Kawato et al., 1987; Milner and Hinder, 2006; Thoroughman and Shadmehr, 2000). Although some saturation is seen for large errors, in general, trial-to-trial learning increases with the magnitude of the error on the previous trial (Körding and Wolpert, 2004; Wei and Kording, 2009). Therefore, by reducing kinematic errors, muscle co-contraction could reduce trial-to-trial learning and thus decrease the overall rate or extent of learning.

Third, muscle co-contraction could have no effect on learning if kinematic errors are normalized against the current level of muscle co-contraction. Indeed, incoming sensory information is often pre-processed on the basis of motor output. For example, reafferent feedback is attenuated by sensory predictions based on efference copies of motor commands (Blakemore et al., 2000, 1998; Shergill et al., 2003). It is therefore possible that error-based learning is driven by pre-processed kinematic errors. Consequently, even though kinematic errors are smaller during muscle co-contraction, the motor system may still adapt to the true internal model performance error. Alternatively, or concurrently, the opposing effects of concentrated learning in state space (hypothesis one) and smaller kinematic errors (hypothesis two) could cancel out such that muscle co-contraction has no effect on learning.

To differentiate between these hypotheses, we developed a novel paradigm in which participants were either pretrained to increase (stiff group) or decrease (relaxed group) muscle co-contraction in the presence of dynamic perturbations. A third group (control group) was not pretrained at all. We then examined force-field adaptation using channel trials, which constrain movements to a straight line and allow us to measure the force generated into the channel walls as a measure of predictive compensation. In the initial stage of exposure, the stiff group adapted more than both the relaxed group and the control group. In both the initial and final stage of exposure, we found a positive correlation between an individual's level of muscle co-contraction and their level of adaptation.

2.2 Methods

Thirty-six neurologically intact participants (16 men and 20 women; age 25.3 ± 5.3 yr, mean ± s.d.) were recruited to participate in the experiment, which had been approved by the Cambridge Psychology Research Ethics Committee and was performed in accordance with guidelines and regulations. All participants provided written informed consent and were right-handed according to the Edinburgh handedness inventory (Oldfield, 1971). Participants had not previously performed any experiments involving velocity-dependent curl fields.

2.2.1 Apparatus

Experiments were performed using a vBOT planar robotic manipulandum (Howard et al., 2009) with virtual-reality system and air table (Fig. 2.1a). The vBOT is a modular, general-purpose, two-dimensional planar manipulandum optimized for dynamic learning paradigms. Position was measured using optical encoders sampled at 1 kHz. Torque motors allowed forces to be generated at the endpoint. A monitor mounted above the vBOT projected visual feedback into the plane of movement via a horizontal mirror. The location of the mirror prevented direct vision of the hand and forearm. Participants were seated in front of the robotic manipulandum with their torso restrained by a four-point harness to reduce movement. The handle of the manipulandum was grasped with the right hand, while the right forearm was supported on an air sled, which constrained arm movements to the horizontal plane.

Movements were performed in either a null field (vBOT passive), a velocity-dependent curl force field or in a simulated channel. In the curl field, the force generated at the

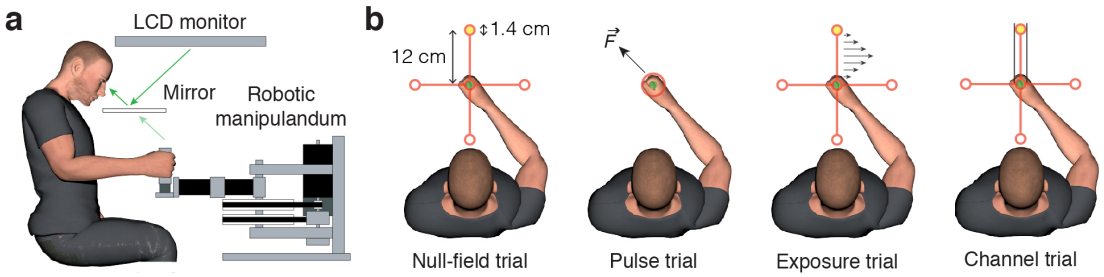


Figure 2.1. Experimental paradigm. a, The participant grasped the handle of the robotic manipulandum (vBOT) while seated. Visual feedback was presented veridically using a horizontally-mounted monitor viewed through a mirror. **b**, At the start of each reaching trial, a target was illuminated (yellow circle), cueing the participant to move their hand (green circle) from the center of the workspace to the illuminated target by making a fast reaching movement. In the pulse phase, a series of brief pulses of force (black arrow) was applied to the handle of the manipulandum in a random direction. Participants in the stiff group were instructed to keep their hand within a 1.5 cm radius of the home position (red ring) by co-contracting the muscles of their arm. Participants in the relaxed group were instructed to let their hand move freely by relaxing the muscles of their arm. In the exposure phase, a velocity-dependent force field was applied. Channel trials were used to assess adaptation for the 90° target only.

handle of the manipulandum was given by:

$$\begin{pmatrix} f_{\mathrm{x}} \\ f_{\mathrm{y}} \end{pmatrix} = b \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}, \tag{2.1}$$

where f_{x}, f_{y}, $\dot{x}$ and $\dot{y}$ are the forces and velocities at the handle in the x- and y-directions respectively. The gain b was set to $\pm$ 16 Ns/m, where the sign of b specified the direction of the curl (clockwise or counterclockwise). On channel trials, the hand was constrained to move along a straight line from the home position to the target, using stiffness (6,000 N/m) and damping (5 Ns/m) perpendicular to the movement direction. Channel trials clamped the kinematic error to zero, allowing forces that were generated in a feedforward manner to be measured orthogonal to the direction of reach (Milner and Franklin, 2005; Scheidt et al., 2000).

2.2.2 Electromyography

Surface electromyography (EMG) was recorded from two monoarticular shoulder muscles (pectoralis major and posterior deltoid), two biarticular shoulder-elbow muscles (biceps brachii and the long head of the triceps) and two monoarticular elbow muscles (the lateral head of the triceps and brachioradialis). The EMG was recorded using the Delsys Bagnoli 8 electromyography system (Boston, MA). The skin was cleaned with alcohol and prepared with an abrasive gel to increase conductance. After the abrasive gel was removed with cotton wool, conductive gel was applied to the electrodes, which were secured to the skin with tape. Electrode placement was chosen to maximize signal while minimizing crosstalk from adjacent muscles. EMG signals were analog band-pass filtered between 20 and 450 Hz (in the Delsys Bagnoli EMG system) and then sampled at 2 kHz.

2.2.3 Paradigm

With the exception of the pulse phase, in which a series of force pulses was applied to the handle of the robot while the hand was stationary, participants made 12 cm center-out reaching movements to one of four targets located at 0°, 90°, 180° and 270°. Visual feedback consisted of the home position (0.7 cm radius disk), the four targets (0.7 cm radius rings), a set of red lines connecting each target to the home position and the hand cursor (0.5 cm radius disk) (Fig. 2.1b). The four targets, the home position and the connecting lines were displayed at all times, except during the pulse phase. Each trial started when the hand cursor had remained within 0.3 cm of the home position at a speed below 0.5 cm/s for 100 ms. One of the targets was then filled in, followed 500 ms later by an auditory tone, which cued the participant to initiate their movement. The trial ended when the hand cursor had remained within 0.3 cm of the target position for 500 ms. If the peak speed of the movement was less than 40

cm/s or greater than 60 cm/s, a low-frequency auditory tone was played, and a 'Too Slow' or 'Too Fast' message was displayed, respectively. At the end of each trial, the hand was passively returned to the home position by the robotic manipulandum.

Participants were assigned to either a stiff group (n = 12), a relaxed group (n = 12) or a control group (n = 12). Each experiment began with a training phase to familiarize participants with the passive dynamics of the robot and the requirements of the task. Participants then performed a pre-exposure phase consisting of 100 null-field trials (20 blocks), a pulse phase consisting of 50 trials (stiff and relaxed groups only), a post-pulse phase consisting of four null-field trials, an exposure phase consisting of 410 force-field trials (83 blocks) and a post-exposure phase consisting of 100 null-field trials (20 blocks). Each block of five trials consisted of one null or field trial to each target and one channel trial to the 90° target. We only applied channel trials for one of the four targets (the 90° target for all participants to allow comparison) to provide enough force-field trials for learning without making the experiment excessively long. The order of trials within a block was pseudorandom, except that a channel trial never occurred on the first trial of a block. The force field direction was fixed for each participant but counterbalanced across participants.

The pulse phase was designed to increase muscle co-contraction in the stiff group and decrease muscle co-contraction in the relaxed group. Each trial started with the hand in the home position. A single force pulse (5 N for 500 ms) was then applied to the hand. The 50 force pulses (one for each trial) ranged uniformly in direction over 360° and were presented in a pseudorandom order. Participants in the stiff group were instructed to keep the handle positioned within a border (1.5 cm radius ring) around the home position by co-contracting the muscles of their limb (Fig. 2.1b). Participants in the relaxed group were instructed to let their hand move freely by relaxing the muscles of their limb. After the force pulse had ended, the handle of the manipulandum was passively returned to the home position in preparation for the next trial.

During the exposure phase, participants in the stiff group were instructed to maintain their arm stiff and to not deviate from the red lines connecting the home position to the target (described above), and participants in the relaxed group were instructed to relax their arm. The control group did not perform the pulse phase and were not given any explicit instructions regarding muscle co-contraction. The exact instructions given to participants can be found in Appendix A.

Participants were given a 45 second rest break after the pre-exposure phase and after every 13 blocks (65 trials) in the exposure phase, with the exception of the end of the exposure phase, which was followed directly by the post-exposure phase. At the end of each rest break, participants in the stiff group were reminded that co-contracting their muscles may help them to make straight reaching movements, and participants in the relaxed group were reminded that if they relaxed their arm their movements would naturally become straighter. To mitigate the effects of time-dependent decay of memory, a channel trial did not occur in the first four movements after each rest break.

2.2.4 Analysis

Data analysis was performed using MATLAB R2017b. Two measures of performance were computed. On exposure and null-field trials the maximum distance between the path of the hand and a straight line connecting the initial hand position and the target (maximum perpendicular error) was computed. On channel trials the percentage of the force field that was compensated for (adaptation) was computed by regressing the actual forces $f_{\mathrm{a}}(t)$ generated by participants in the channel on the ideal forces $f_{\mathrm{i}}(t)$ that would fully compensate for the forces on a force-field trial (defined in equation 2.1):

$$f_{\mathrm{a}}(t) = k \times f_{\mathrm{i}}(t) \tag{2.2}$$

$$\text{adaptation} = k \times 100\%, \tag{2.3}$$

where k is the regression coefficient and t is the discrete time step of the channel trial. The offset of the regression was constrained to zero. For this analysis, we used the portion of the movement where the hand speed was greater than 1 cm/s. At all stages of the experiment the ideal force term $f_{\mathrm{i}}(t)$ refers to the force required to fully compensate for the force field present in the exposure phase. Therefore, adaptation in the pre-exposure phase prior to learning should be close to zero. To combine maximum perpendicular error (MPE) and adaptation data across participants, data from individual participants was sign adjusted according to the direction of the field they were trained on.

To quantify oscillations on individual channel trials, we calculated an oscillation index that was the path length of each force profile (from when the hand speed exceeded 1 cm/s to when it dropped below 1 cm/s), where a greater path length indicates a more oscillatory profile. To control for differences in movement duration and adaptation level, each force profile was first normalized in amplitude by dividing by the force range (max - min) and normalized in time between 0 and 1.

The raw EMG signal was bandpass filtered (30–500 Hz) using a 10th-order Butterworth filter implemented with MATLAB's filtfilt function and then full-wave rectified. We examined the EMG over two periods: an early period comprised of feedforward signals and a later period comprised of both feedforward and feedback signals. These periods were defined separately for the pulse phase and for reaching movements. In the pulse phase, the early period was taken from 200 ms before the perturbation onset to the perturbation onset, and the later period was taken from the perturbation onset to 130 ms after the perturbation onset. During reaching movements, the early period was taken from 200 ms before movement onset to 130 ms after movement onset. This period can be used as a measure of the feedforward muscle activation as feedback components of EMG are first detectable around 130 ms after movement onset when the hand is smoothly perturbed by a velocity-dependent force field (Franklin et al., 2008). The later period was taken from 130 ms to 400 ms after movement onset. To normalize

EMG traces across the muscles, we first computed for each muscle the mean EMG over the early period for the non-channel trials in the pre-exposure phase. We then normalized the EMG trace for each trial by dividing the trace by the corresponding mean for that muscle (Franklin et al., 2012). Note that this only scales the EMG traces for each muscle and does not affect the temporal profile. A global measure of muscle activity was then computed for each trial by calculating the difference between the summed EMG across muscles (first averaged across time points on a trial) and the summed pre-exposure EMG across muscles (first averaged across time points and trials for the pre-exposure phase). Accordingly, negative global EMG signifies reduced activity relative to the pre-exposure phase, whereas positive global EMG signifies increased activity relative to the pre-exposure phase. For plotting purposes, EMG waveforms were time normalized by resampling 1000 data points between movement onset and movement end on each trial for each participant. Movement onset was defined as the time the cursor first moved 0.3 cm from the center of the home position. Movement end was defined as the time the cursor was first within 0.3 cm of the center of the target position. Each resampled data point was assigned a time point linearly spaced between 0 and the mean movement duration across participants.

To identify differences in global EMG, MPE and adaptation between the groups, between-subjects ANOVAs, mixed-design ANOVAs and unpaired t-tests were performed. We also performed linear regression and ANCOVAs to identify a linear relationship between global EMG and adaptation. All statistical tests were two-sided with significance set to $P < 0.05$. Data are reported as mean ± standard error of the mean (s.e.m.).

2.3 Results

Participants started the experiment by making center-out reaching movements in a null field (pre-exposure phase). Then, in the stiff and relaxed groups, voluntary changes to

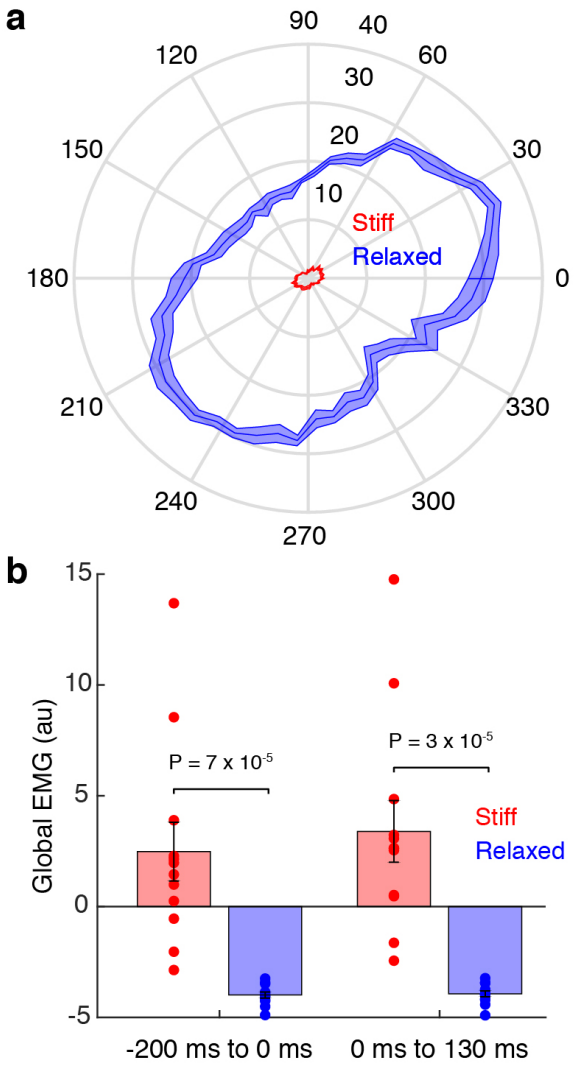


Figure 2.2. Performance in the force pulse phase. a, Polar plot of the maximum distance (cm) the hand moved as a function of the force pulse direction. Data show mean ± s.e.m. across participants. **b**, Global EMG before (left) and after (right) the force pulse for the stiff (red) and relaxed (blue) groups. Data show mean ± s.e.m. across participants.

muscle co-contraction were elicited by instructing participants to resist or comply with a sequence of force pulses while the hand was stationary (pulse phase). Figure 2.2a shows the maximum distance moved by the hand following force pulses in different directions. The maximum distance moved by the hand was smaller in the stiff group compared with the relaxed group (stiff: 1.7 cm ± 0.1, relaxed: 23.2 ± 0.9, unpaired t-test, $t_{22} = 22.81$, $P = 8 \times 10^{-17}$). This was due to greater muscle co-contraction, as

confirmed by global EMG (stiff: 2.9 ± 1.4, relaxed: -4.0 ± 0.1, mixed-design ANOVA, $F_{1,22} = 25.74$, $P = 4 \times 10^{-5}$), both in the period preceding the force pulse (Fig. 2.2b; stiff: 2.5 ± 1.3, relaxed: -4.0 ± 0.1, unpaired t-test, $t_{22} = 4.87$, $P = 7 \times 10^{-5}$) and in the period following pulse onset (Fig. 2.2b; stiff: 3.4 ± 1.4, relaxed: -3.9 ± 0.1, unpaired t-test, $t_{22} = 5.26$, $P = 3 \times 10^{-5}$). The variability of global EMG was greater for the stiff group (Fig. 2.2b). This is most likely because the relaxed group are bounded by an EMG of zero, whereas the stiff group can stiffen up to different extents due to different bounds across participants. Note that because joint impedance increases faster than force variability as a function of muscle co-contraction (Selen et al., 2005), co-contraction has a net stabilizing effect on endpoint kinematics (Gribble et al., 2003; Selen et al., 2009). The pulse phase was omitted for the control group.

2.3.1 Electromyography

After the pulse phase (stiff and relaxed groups) or the pre-exposure phase (control group), participants performed reaching movements in a velocity-dependent curl field (exposure phase). Figure 2.3 shows global EMG for all groups over the course of the experiment, split into periods separated by rest breaks. To identify differences in muscle co-contraction between the groups during the exposure phase, we performed a mixed-design ANOVA with stage of exposure (first vs. second half) and period of movement (early vs. later period) as within-subject factors and global EMG as the response variable. We found a main effect of group (mixed-design ANOVA, $F_{2,33} = 8.59$, $P = 0.001$) but no interaction between group and stage of exposure (mixed-design ANOVA, $F_{2,33} = 1.58$, $P = 0.220$). We therefore combined data from the first and second halves of the exposure phase. Importantly, global EMG was greater in the stiff group than the relaxed group, both in the early period of the movement (Fig. 2.3a; stiff: 2.3 ± 0.8, relaxed: -0.4 ± 0.2, unpaired t-test, $t_{22} = 3.42$, $P = 0.002$) and in the later period of the movement (Fig. 2.3b; stiff: 4.1 ± 1.1, relaxed: 0.1 ± 0.3, unpaired t-test, $t_{22} = 3.44$, $P = 0.002$). Global EMG was also greater in the stiff group than

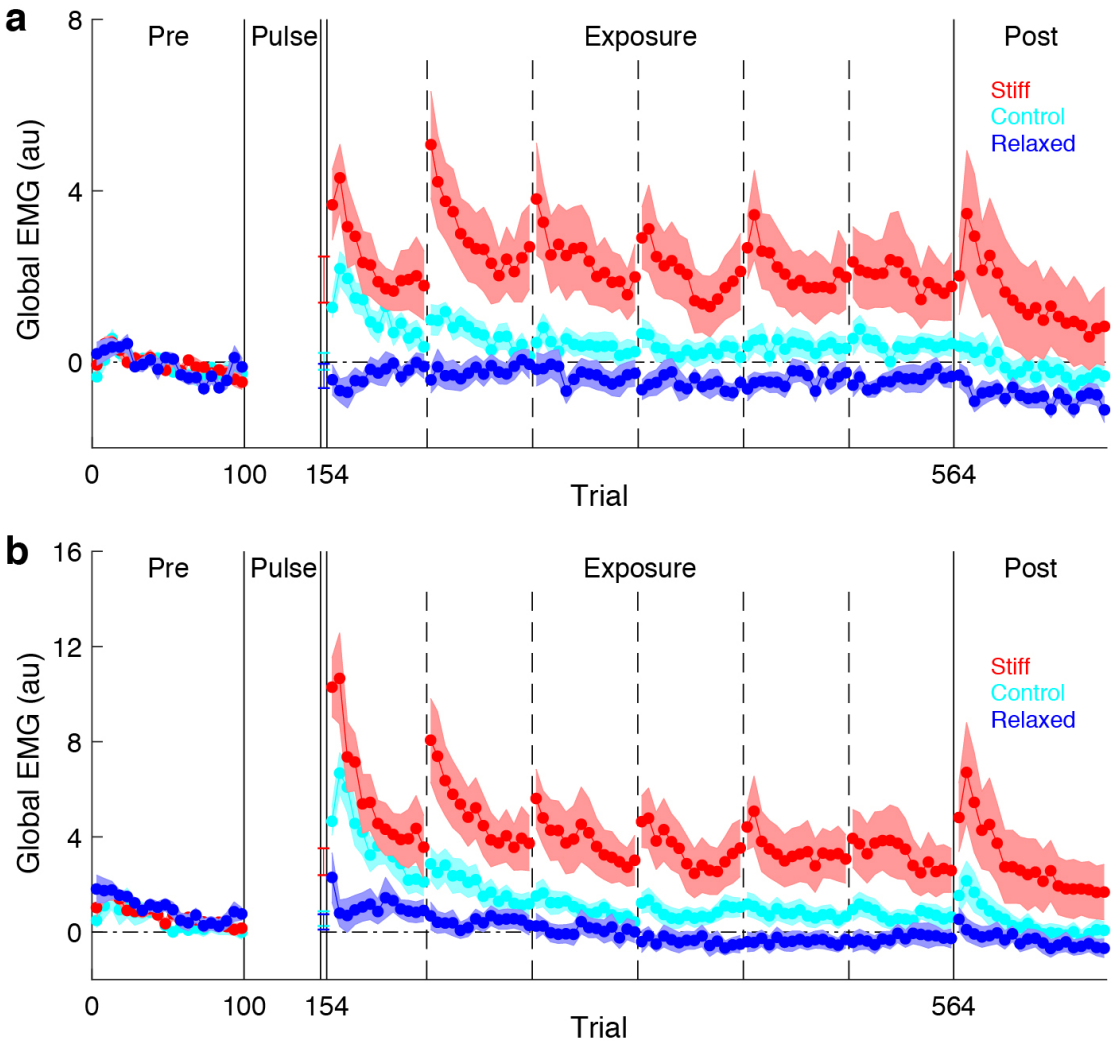


Figure 2.3. Global EMG over the course of the experiment **a**,**b**, Global EMG, which is a proxy for stiffness, was calculated on each trial and plotted as an average of a block of five trials. Data show mean ± s.e.m. across participants in the stiff (red), relaxed (blue) and control (cyan) groups. Vertical dashed lines indicate rest breaks. **a**, Global EMG from -200 ms to 130 ms relative to movement onset. **b**, Global EMG from 130 ms to 400 ms relative to movement onset.

the control group, both in the early period of the movement (Fig. 2.3a; control: 0.5 ± 0.2, two-tailed unpaired t-test, $t_{22} = 2.28$, $P = 0.033$) and in the later period of the movement (Fig. 2.3b; control: 1.5 ± 0.4, unpaired t-test, $t_{22} = 2.20$, $P = 0.038$). Finally, Global EMG was greater in the control group than the relaxed group, both in the early period of the movement (Fig. 2.3a; unpaired t-test, $t_{22} = 3.15$, $P = 0.005$)

and in the later period of the movement (Fig. 2.3b; unpaired t-test, $t_{22} = 2.70$, $P = 0.013$). Therefore, as expected, the stiff group exhibited the greatest degree of muscle co-contraction, followed by the control group and then the relaxed group.

Figure 2.4 shows the average EMG of the stiff (red), relaxed (blue) and control (cyan) groups during the first half of the exposure phase. This shows an increase in EMG both prior to and during movement in many of the muscles. Overall, a clear voluntary increase in EMG can be seen for the stiff group compared with the relaxed group. This increase is seen in all muscles pairs, suggesting that muscle co-contraction is not specific to a single joint. The EMG of the control group is intermediate between the stiff and relaxed groups.

2.3.2 Movement Analysis

In the pre-exposure phase, reaching movements were relatively straight in all groups, as demonstrated by small MPEs (Fig. 2.5a). During the exposure phase, there was no difference in peak velocity between the groups (stiff: 48.8 ± 0.4 cm/s, relaxed: 48.4 ± 0.5, control: 48.0 ± 0.5, between-subjects ANOVA, $F_{2,33} = 0.61$, $P = 0.549$). Therefore, each group experienced similar forces during learning. In the first block (5 trials) of the exposure phase, all groups showed large deviations from a straight line, and these deviations were different between the groups (between-subjects ANOVA, $F_{2,33} = 9.94$, $P = 4 \times 10^{-4}$). Importantly, MPE was, as expected, smaller in the stiff group compared with the relaxed group (stiff: 3.3 cm ± 0.2, relaxed: 4.5 ± 0.2, unpaired t-test, $t_{22} = 5.42$, $P = 2 \times 10^{-5}$). Furthermore, MPE was also smaller in the control group compared with the relaxed group (control: 3.2 ± 0.3, unpaired t-test, $t_{22} = 3.49$, $P = 0.002$). However, there was no difference in MPE between the stiff group and the control group (unpaired t-test, $t_{22} = 0.03$, $P = 0.977$). As the exposure phase progressed, MPE decreased in all groups. By the final 13 blocks (65 trials) of the exposure phase, there was no difference in MPE between the groups (stiff: 0.4 cm ±

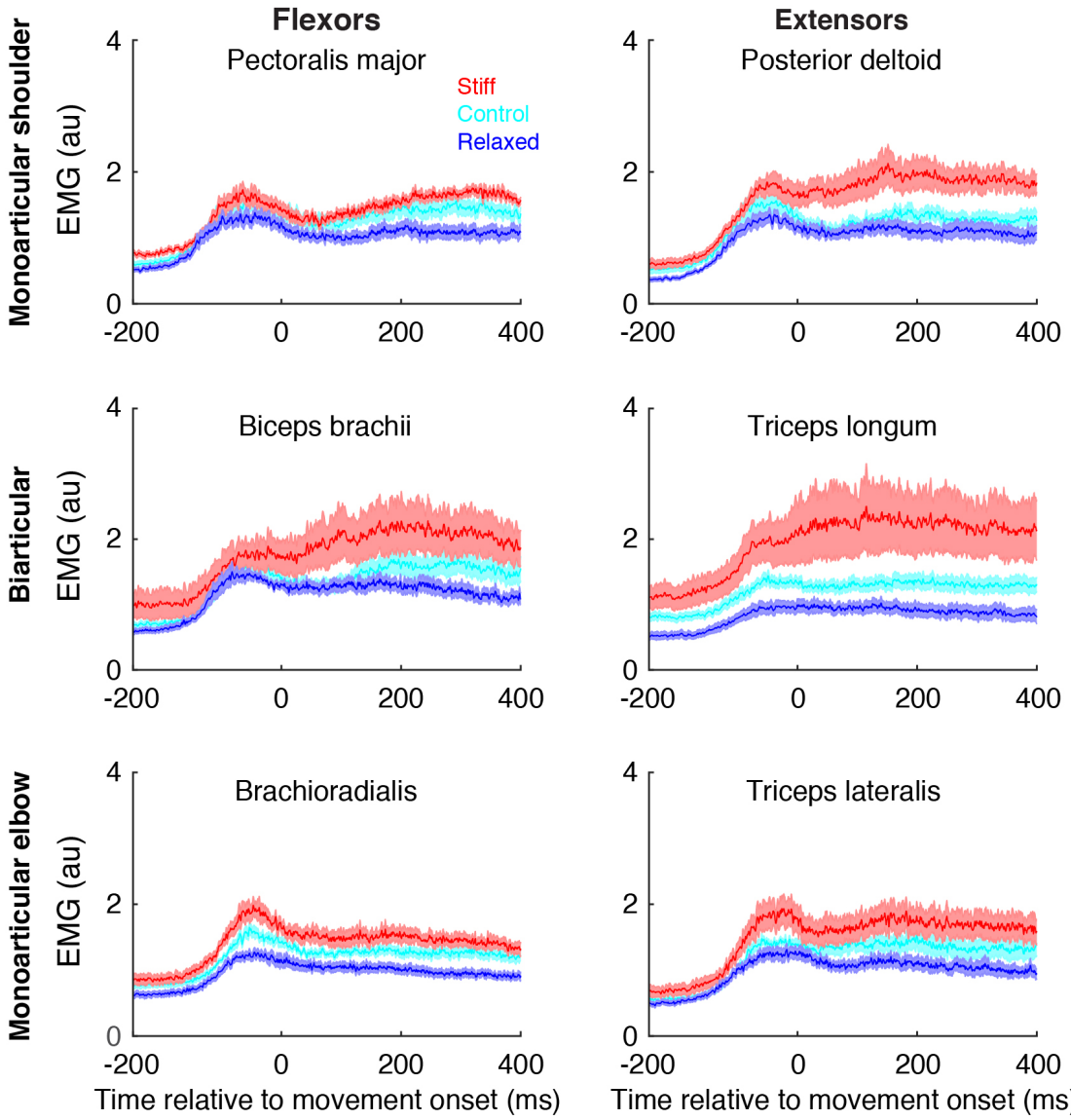


Figure 2.4. Muscle activity during the first half of the exposure phase. The EMG is averaged across all directions on force field trials. The period after movement onset has been time normalized. The time from 0 ms to 400 ms corresponds to approximately 64% of the mean movement duration across participants. The period before movement onset is shown in real time. Data show mean ± s.e.m. across all participants in the stiff (red), relaxed (blue) and control (cyan) groups.

0.1, relaxed: 0.7 ± 0.1, control: 0.7 ± 0.2, between-subjects ANOVA, $F_{2,33} = 1.94$, $P = 0.160$). In the null field of the post-exposure phase, MPE rose again, demonstrating after-effects.

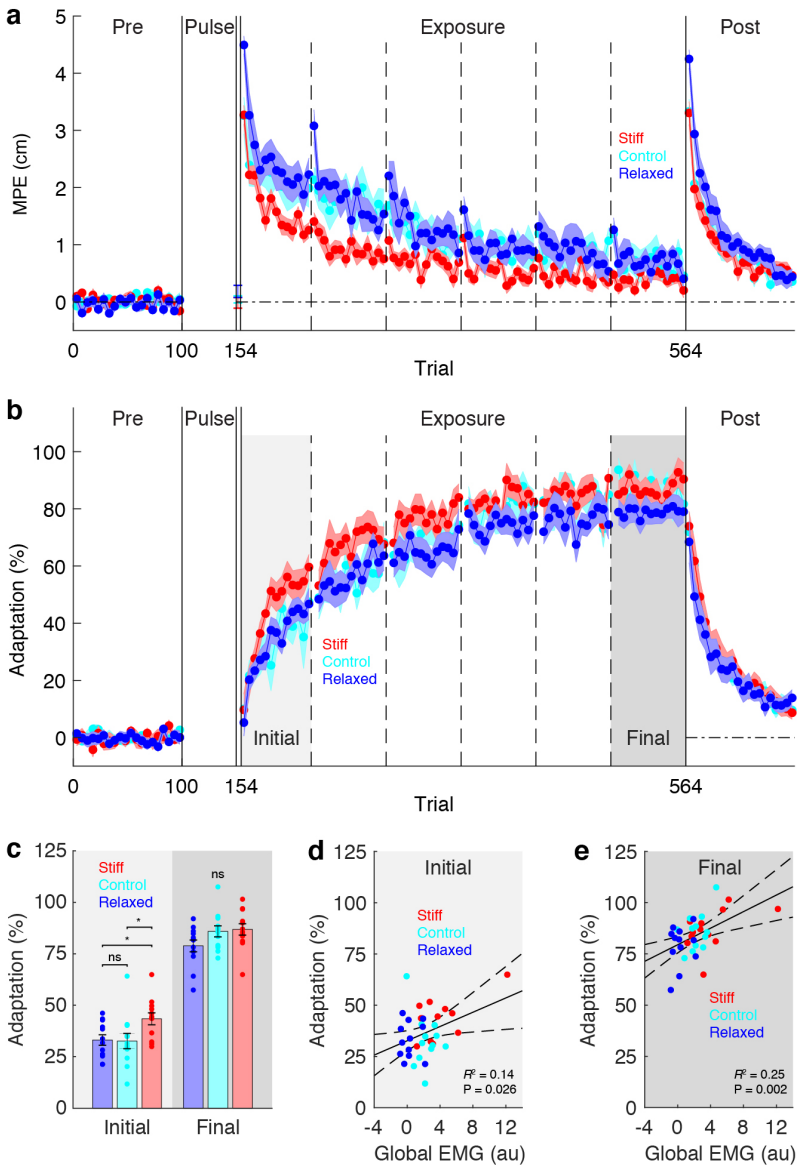


Figure 2.5. Kinematic and dynamic adaptation over the course of the experiment. **a**, Maximum perpendicular error (MPE) was calculated on each trial and plotted as an average of a block of five trials. Data show mean ± s.e.m. across participants in the stiff (red), relaxed (blue) and control (cyan) groups. Vertical dashed lines indicate rest breaks. **b**, is the same as **a** but for adaptation measured on channel trials. **c**, Mean ± s.e.m. initial adaptation (first 13 blocks of the exposure phase) and final adaptation (last 13 blocks of the exposure phase) for all groups. Asterisks indicate significant ($^*P < 0.05$), ns indicates not significant. **d**,**e**, Least-squares regression between each participant's global EMG (first 13 blocks of the exposure phase, early and later periods of movement combined) and their initial adaptation (**d**) and their final adaptation (**e**). Dashed curves show 95% confidence intervals for the fitted linear model. Data from all groups have been pooled.

In the relaxed group, increases in MPE can be seen at the start of each rest break (Fig. 5a; vertical dashed lines) when participants were reminded that their movements would naturally become straighter if they relaxed their arm. In the stiff group, changes in MPE cannot be seen at the start of each rest break. This is likely because decreases in MPE due to increased co-contraction are offset by increases in MPE due to forgetting of learning during rest breaks.

2.3.3 Force compensation

In the pre-exposure phase, adaptation measured on channel trials was, as expected, close to zero in all groups (Fig. 2.5b). During exposure, adaptation increased in all groups, eventually reaching asymptote. To assess learning, we compared initial (first 13 blocks of the exposure phase) and final (last 13 blocks of the exposure phase) adaptation between the groups. This revealed a main effect of group (mixed-design ANOVA, $F_{2,33} = 3.75$, $P = 0.034$) and no interaction (mixed-design ANOVA, $F_{2,33} = 2.35$, $P = 0.111$). Initial adaptation was greater in the stiff group than both the relaxed group (Fig. 2.5c; stiff: 43.5 ± 2.9, relaxed: 33.2 ± 2.5, unpaired t-test, $t_{22} = 2.68$, $P = 0.014$) and the control group (Fig. 2.5c; control: 32.6 ± 3.7, unpaired t-test, $t_{22} = 2.31$, $P = 0.031$). However, there was no difference in initial adaptation between the relaxed group and the control group (Fig. 2.5c; unpaired t-test, $t_{22} = 0.12$, $P = 0.904$). Final adaptation was not significantly different between the groups (Fig. 2.5c; stiff: 86.9 ± 2.8, relaxed: 78.9 ± 2.9, control: 85.9 ± 2.7, between-subjects ANOVA, $F_{2,33} = 2.49$, $P = 0.099$).

We examined data at the individual level pooling across all three groups. This revealed a significant linear relationship between a participant's global EMG (first 13 blocks of the exposure phase, early and later periods of the movement combined) and both their initial adaptation (Fig. 2.5d; $R^2 = 0.14$, $\beta = 1.74$, F-test, $F_{1,34} = 5.45$, $P = 0.026$) and their final adaptation (Fig. 2.5e; $R^2 = 0.25$, $\beta = 2.02$, F-test, $F_{1,34} = 11.17$, $P =$

0.002). To rule out the possibility that muscle co-contraction influences adaptation by increasing/decreasing oscillations on channel trials, we performed multiple linear regression with global EMG and oscillation index (see methods) as predictors and adaptation as the response variable. Importantly, oscillation index was not a significant predictor of either initial adaptation (β = -0.17, t-test, t_{33} = 0.04, P = 0.971) or final adaptation (β = -7.58, t-test, t_{33} = 1.00, P = 0.327). In contrast, global EMG was a significant predictor of both initial adaptation (β = 1.75, t-test, t_{33} = 2.19, P = 0.036) and final adaptation (β = 2.25, t-test, t_{33} = 3.24, P = 0.003). Therefore, the relationship between global EMG and adaptation cannot be explained in terms of increased/ decreased oscillations on channel trials due to muscle co-contraction.

To investigate whether the significant linear relationship between global EMG and adaptation (Fig. 2.5d-e) is a consequence of the previously reported inter-group differences in adaptation (Fig. 2.5c), we performed an ANCOVA with adaptation as the dependent variable, group as the independent variable and global EMG as the covariate. In initial adaptation, there was no main effect of global EMG (ANCOVA, $F_{1,30}$ = 1.68, P = 0.205) and no interaction between global EMG and group (ANCOVA, $F_{2,30}$ = 1.67, P = 0.206). In contrast, in final adaptation, there was a main effect of global EMG (ANCOVA, $F_{1,30}$ = 5.59, P = 0.025) and no interaction between global EMG and group (ANCOVA, $F_{2,30}$ = 0.32, P = 0.731). Therefore, global EMG remained a significant predictor of final adaptation at the individual level.

2.3.4 Discussion

We examined force-field adaptation under increased, decreased and baseline levels of muscle co-contraction. Participants were randomly assigned to either a stiff, relaxed or control group. The stiff and relaxed groups were pretrained using a sequence of force pulses to either increase (stiff group) or decrease (relaxed group) muscle co-contraction in the presence of perturbations. The control group was not pretrained. All groups

performed reaching movements in a velocity-dependent curl field. We assessed learning by measuring adaptation, that is predictive force compensation using channels trials. In the initial stage of exposure, the stiff group adapted more than the relaxed and control groups, despite experiencing smaller kinematic errors, an important training signal for error-based learning. Furthermore, at the individual level, muscle co-contraction was significantly correlated with adaptation in the final stage of exposure. These results show that, in addition to improving kinematic accuracy, muscle co-contraction also increases the rate of acquisition of an internal model.

There are a number of reasons why increased muscle co-contraction could lead to an increase in the rate of dynamic motor learning. First, when motor errors are experienced, the internal model is updated with respect to the actual states visited rather than the intended states, a process termed motion-referenced learning (Castro et al., 2011). Consequently, any intervention that increases the overlap between the actual motions experienced and the motion required to reach the target, such as increased muscle co-contraction, could increase the rate of adaptation. Second, error sensitivity is greater for smaller errors (Fine and Thoroughman, 2007; Marko et al., 2012; Wei and Kording, 2009). This could explain why muscle co-contraction accelerates adaptation, despite decreasing the size of errors. In a causal inference framework (Körding et al., 2007a), larger errors are more likely to have been caused by factors extrinsic to the motor plant and so should be discounted to a greater extent. Consequently, larger errors do not guarantee greater learning. Indeed, single-trial adaptation (the product of error sensitivity and error magnitude) appears to be non-monotonic with a peak at an intermediate error magnitude (Herzfeld et al., 2014; Robinson et al., 2003). This non-monotonic relationship can be appreciated in studies of error augmentation. For example, during reach adaptation to a visuomotor rotation, the rate of learning almost doubles when visual errors are amplified by a gain of 2 (Patton et al., 2013). However, when visual errors are amplified by a gain of 3, the rate of learning returns to the level associated with veridical visual feedback. Given this non-monotonic relationship between adaptation and error magnitude, it is possible that muscle co-contraction

increases the rate of adaptation by reducing errors to a size that induces greater learning.

Previous studies of error sensitivity (Fine and Thoroughman, 2007; Franklin et al., 2008; Herzfeld et al., 2014; Marko et al., 2012; Robinson et al., 2003; Wei and Kording, 2009) did not examine its dependence on muscle co-contraction. Therefore, it is possible that error sensitivity is a function of muscle co-contraction, such that as muscle co-contraction increases, single-trial adaptation is maximized by progressively smaller errors. For the sensorimotor system to implement such an error-sensitivity function, it would need to have knowledge of the effects of muscle co-contraction on the statistics of kinematic errors. Such knowledge has been postulated in normative models of impedance control, in which muscle co-contraction emerges as a strategy to minimize the expected cost of movement by reducing uncertainty in the forward dynamics model (Mitrovic et al., 2010). If error sensitivity were a function of muscle co-contraction, when kinematic errors are experienced, the sensorimotor system could appropriately assign credit between the forward dynamics model and sensorimotor noise, controlling for the level of muscle co-contraction. Interestingly, the forward dynamics model, which maps current states and motor commands to next states, is also a function of muscle co-contraction (a state determined by a motor command), and so it is an open question to what degree adaptation under increased co-contraction generalizes to other co-contraction levels.

Muscle co-contraction could affect motor performance indirectly via an effect on proprioception. Proprioception is primarily mediated by muscle spindle receptors, which encode the length and velocity of muscles (Dimitriou, 2014; Ribot-Ciscar and Roll, 1998). The sensitivity of muscle spindles to stretch is maintained by alpha-gamma coactivation (Vallbo, 1970, 1974). Alpha motor neurons activate extrafusal muscle fibres, while gamma motor neurons activate the contractile polar ends of muscle spindles to prevent spindle slackening during contraction of the extrafusal muscle. If a contracting muscle does not shorten (e.g., during co-contraction) but the muscle

spindle does, the spindle will become taught, potentially increasing its sensitivity to stretch. Hence, in principle, increased alpha-gamma coactivation during muscle co-contraction could enhance proprioceptive acuity. Indeed, early work suggested that the threshold for detecting movements imposed at the elbow is lower when muscles are contracted compared to relaxed (Colebatch and McCloskey, 1987; Taylor and McCloskey, 1992). This gave rise to the common view that proprioception is more accurate under active than passive conditions (Gandevia et al., 1992). However, these experiments were subsequently criticised on the basis that the muscles had not been appropriately conditioned to remove slack in the spindles that arises due to thixotropy (the dependence of a muscle's state on its contraction and length change history) (Proske and Gandevia, 2012). Critically, subsequent experiments that did condition the muscles found the opposite result; the ability to detect limb movement (Wise et al., 1998) and to match limb position (Wise and Fallon, 2002) at the elbow was reduced when muscles were co-contracted compared to relaxed. Hence proprioceptive acuity was diminished when muscles were co-contracted. It has been suggested that this reduction in acuity could arise from a decrease in muscle spindle responsiveness to stretch during co-contraction (Wise et al., 1999), which may help avoid oscillations in agonist-antagonist muscle pairs that exhibit reciprocal inhibition (Dideriksen et al., 2015; Nielsen and Kagamihara, 1993). It should be noted that to measure movement detection thresholds, these experiments imposed small angular displacements of $< 1°$, which are substantially smaller than the displacements imposed by the force field in our experiment. Nevertheless, given these findings, it seems unlikely that the greater adaptation we observed in the stiff group is due to enhanced proprioception arising from increased muscle co-contraction.

In the pulse phase, the stiff group were instructed to keep the handle as still as possible, whereas the relaxed group were instructed to let their hand move freely to wherever the handle took it. Because the direction of each force pulse was unpredictable, participants in the stiff group had to use impedance control and feedback mechanisms to minimise displacement of their hand. Therefore, it is possible that participants in the stiff

group upregulated their feedback gains (Scott, 2004; Scott et al., 2015), and that these feedback gains persisted into the exposure phase. Importantly, however, the channel trials used to assess adaptation measure feedforward rather than feedback force compensation, so it is unlikely that enhanced feedback gains promoted by instructions in the pulse phase are responsible for the greater adaptation that we observed in the stiff group. It is also unlikely that the instructions given to the relaxed group in the pulse phase reduced their engagement with the task, as adaptation was not different between the relaxed and control groups.

In the exposure phase, the stiff group were instructed to to stay as close as possible to the red lines leading from the home position to the targets, whereas the relaxed group were not. From the perspective of optimal feedback control, this instruction defines a cost function that penalises deviations of the hand from the red line throughout the movement. Importantly, this cost function will lead to less (not more) adaptation than the typical cost function used for point-to-point reaching movements that only penalises the terminal state of the movement. This is because—when the terminal state alone is penalised—the optimal policy in a viscous curl field leads to hand paths that overcompensate for the field, and humans exhibit such overcompensation (Izawa et al., 2008). Therefore, the instruction to stay as close as possible to the red line will likely have diminished or even eliminated this overcompensation and thus produced less adaptation according to our measure (the proportion of the force field compensated for). Finally, it is possible that participants in the stiff group adopted an explicit strategy to stay close to the red lines. However, the role of explicit strategy use in force-field adaptation has received little attention in the literature and is not well understood (Schween et al., 2019).

All participants were able to form an internal model of the force field, as demonstrated by adaptation on channel trials and after-effects in the post-exposure phase. This is consistent with previous studies that have demonstrated that people can adapt to viscous force fields that are superimposed on background loads (Kurtzer et al.,

2005; Liu and Reinkensmeyer, 2007). It has even been suggested that people partition force fields into dynamic and static components (Kurtzer et al., 2005), which could allow velocity-dependent dynamics to be learned across a range of muscle load states. However, not all muscle states appear to be available to the internal model. For example, after recovery from fatigue, muscles initially generate excessive force, as if the internal model is still producing motor commands appropriate for fatigued muscles (Takahashi et al., 2006).

In our study, although participants made movements in four directions, we only measured adaptation for one direction to limit the length of the experiment. Given that endpoint stiffness is in general anisotropic (Burdet et al., 2001; Franklin et al., 2007, 2004), adaptation may similarly have been anisotropic, giving rise to an adaptation ellipse. Moreover, the geometry (size, shape, orientation) of this adaptation ellipse, like the geometry of the stiffness ellipse (Burdet et al., 2001; Franklin et al., 2007, 2004), may be modifiable, allowing the rate of adaptation to be modulated in a direction-dependent manner. For example, by orienting the major axis of the stiffness ellipse in the direction of greatest unpredictability in the environment, it may be possible to maximize the rate of adaptation in the direction it is needed most: where errors are largest.

It remains an open question whether muscle co-contraction can modulate adaptation independently of the errors experienced. Feedback error learning (FEL) (Kawato, 1990; Kawato et al., 1987) provides one possible mechanism by which this could occur. According to FEL, the feedback response to error acts as a teaching signal for motor adaptation. Consistent with FEL, single-trial adaptation has been shown to correlate with the magnitude of the feedback response (Albert and Shadmehr, 2016). Given that feedback gains scale with muscle activity (Pruszynski et al., 2009) and feedback responses are potentiated when muscles are co-contracted (Akazawa et al., 1983; Doemges and Rack, 1992; Krutky et al., 2009), adaptation to an error of a given magnitude should be larger when muscles are co-contracted.

In conclusion, we have shown that muscle co-contraction accelerates dynamic motor learning, despite reducing kinematic errors, which are the main driving force behind adaptation. Muscle co-contraction therefore improves both kinematic accuracy and dynamic learning, simultaneously enhancing present and future motor performance. The modifiable nature of muscle co-contraction suggests that the rate of motor adaptation can be actively modulated.

Chapter 3

Multiple motor memories are learned to control different points on a tool

3.1 Introduction

In Chapter 2 we examined how a single novel dynamic environment is learned and how muscle co-contraction affects that learning. In this chapter we focus on how participants learn the dynamics of an object and, in particular, whether difference control locations on a tool can be associated with distinct dynamics.

Understanding how humans learn to use tools is a central challenge in neuroscience (Ingram and Wolpert, 2011). Tool use requires knowledge of the dynamics of objects (relating applied force to motion), which depends on both the physical properties of objects (e.g., mass, mass distribution and friction) and how objects are used to interact with the environment (e.g., for transport or percussion). Numerous paradigms have been developed to explore how humans learn the dynamics of objects. In conventional object-manipulation tasks, participants grasp and lift physical objects with familiar dynamics (Johansson and Westling, 1988, 1984; Westling and Johansson, 1984). When

lifting an object for the first time, visual cues about size and surface material strongly influence lift forces (Baugh et al., 2012; Flanagan and Beltzner, 2000; Gordon et al., 1993). However, after several lifts, participants learn the dynamics of the object, as demonstrated by adaptation of grip and lift forces measured at the fingertips before object lift off (Flanagan et al., 2006; Johansson and Flanagan, 2009). If the dynamics of the object is simple, learning can occur in a single trial (Gordon et al., 1993). However, even for objects with simple dynamics, interference can be seen when alternately lifting the same object and grasping it at different locations (Fu and Santello, 2012, 2014).

Robotic manipulanda offer an alternative testbed for object-manipulation studies (Wolpert and Flanagan, 2010). These interfaces can simulate objects with unfamiliar dynamics (Dingwell et al., 2002, 2004; Flanagan and Wing, 1997; Nagengast et al., 2009; Nasseroleslami et al., 2014; Shadmehr and Mussa-Ivaldi, 1994) and provide an opportunity to manipulate the relationship between haptic and visual feedback (Ahmed et al., 2008; Danion et al., 2011). In a typical adaptation experiment, participants grasp the handle of a robotic manipulandum and make reaching movements that are perturbed by the robot. For example, a perturbation that has been examined extensively is a viscous curl field that generates forces on the hand that act perpendicular to the velocity of the hand and proportional to its speed (Scheidt et al., 2000; Shadmehr and Mussa-Ivaldi, 1994; Sheahan et al., 2016; Smith et al., 2006). Over repeated movements, participants learn to generate the forces required to compensate for the perturbation, forming a motor memory of the grasped object (Cothros et al., 2006; Kluzik et al., 2008). However, the capacity to learn different dynamics in such paradigms appears to be surprisingly limited when the same movement is made (under veridical visual feedback) in force fields that act in opposite directions (Gandolfo et al., 1996; Howard et al., 2013).

In the standard force-field adaptation paradigm, the 'object' being moved to a target is a small disc and the task involves controlling the centre of the disc. However, in natural object-manipulation tasks involving objects with more complex geometries, we can

exert control over different locations, or 'control points', on the object. For example, when using a wide broom to sweep along a wall, we could control the edge of the broom that is nearest to the wall to make contact with the wall. Such contact will perturb the head of the broom, rotating it in opposite directions for walls on the left and right; that is, when controlling the left and right side of the broom. Similarly, when drinking from a wine glass, we can control the near rim as we lift the glass to our mouth and then control the base of the stem as we replace the glass. Numerous studies of object manipulation have examined how memories of object dynamics transfer across changes in grasp configurations (Bursztyn and Flanagan, 2008; Fu and Santello, 2012, 2014; Ingram et al., 2013, 2010, 2011; Salimi et al., 2000). However, none of these previous studies have explicitly manipulated the location of the control point independently of either the grasp location or the required movement of the object.

Here, we tested the hypothesis that the motor system can flexibly engage separate memories when controlling different points on a single object, allowing different dynamics to be learned and associated with these control points. Using a planar robotic interface and virtual-reality system, we linked the movement of a rectangular virtual object to the motion of the hand. The hand was clearly visible at all times and located at the centre of the object (Fig. 3.1). There were two control points, located on the left and right of the object, and two corresponding targets. The two control points were linked to opposing viscous curl fields. On each trial, the participant was required to generate a straight ahead hand movement to align either the left or right control point with its corresponding target. Thus, the required movement was the same for both control points.

We found that participants could learn opposing dynamics when it was linked to these different points. By testing generalization, we show that this learning is associated with control points, rather than hand or target locations. In addition, we show that the motor system only assigns distinct memories to control points if they are linked to different dynamics, allowing flexible and efficient use of motor memory. Finally,

we show that our results are best accounted for by a normative switching state-space model (SSSM). Unlike standard state-space models (SSMs), the SSSM can learn the probability of cues (that is, control points) given the context (that is, force field) and infer the state for each context based on motor errors. Our results uncover an important mechanism through which the human motor system can generate flexible and dextrous motor behaviour.

3.2 Methods

3.2.1 Participants

A total of 50 neurologically intact participants (22 males and 28 females; age 28.2 ± 7.8, mean ± s.d.) were recruited to participate in four experiments, which had been approved by the Cambridge Psychology Research Ethics Committee. All participants provided written informed consent and were right-handed according to the Edinburgh handedness inventory (Oldfield, 1971).

3.2.2 Apparatus

All experiments were performed using a vBOT planar robotic manipulandum with a virtual-reality system and air table, which we described in Chapter 2. A monitor mounted above the vBOT projected virtual images into the plane of movement via a horizontal semi-silvered mirror. (Fig. 3.1a). Participants were seated in front of the vBOT and grasped its handle with their right hand. A lamp was used to illuminate the hand below the mirror with the illumination adjusted so that both the vBOT, hand, arm and virtual images were clearly visible (Fig. 3.1b). This was done to ensure that participants always had an accurate estimate of the state of their hand and arm.

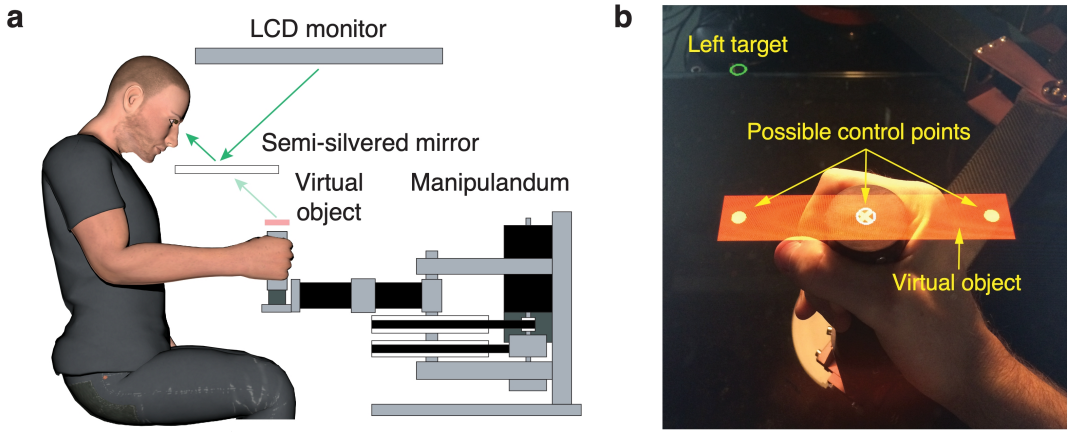


Figure 3.1. Experimental paradigm. **a**, Participants grasped the handle of a robotic manipulandum. A semi-silvered mirror allowed the participant to view their hand as well as a virtual object that was reflected from the monitor. LCD, liquid-crystal display. **b**, Photograph of the hand, virtual object and target as seen from the perspective of the participant. Depending on the experiment and trial type, one of three control points on the object had to be aligned with a target placed above it.

On each trial, the vBOT could generate forces associated with either a velocity-dependent curl field (force-field trials) or a force channel (channel trials). Alternatively, in the case of the null field, no forces were generated (null-field trials). For the curl field, the force generated on the hand was given by

$$\begin{pmatrix} f_{\mathrm{x}} \\ f_{\mathrm{y}} \end{pmatrix} = g \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}, \tag{3.1}$$

where f_{x}, f_{y}, $\dot{x}$ and $\dot{y}$ are the forces and velocities at the handle in the x- and y-directions (mediolateral and anteroposterior, respectively). The field gain g was set to $\pm$ 15 Nsm^{-1} and the sign of g specified whether the curl field was clockwise (CW) or counterclockwise (CCW).

In all experiments, participants first performed a familiarization phase of between 96 and 128 trials, consisting of null-field trials and channel trials for each control point in a pseudorandomized order.

To have precise control over vision and dynamics in our experiment, we used a robotic interface and virtual-reality system. Importantly, a number of previous object-manipulation studies using robotic manipulanda have been shown to produce similar results to experiments that use real-world physical objects (Fu and Santello, 2012; Howard et al., 2009; Ingram et al., 2010, 2017; Ingram and Wolpert, 2011).

3.2.3 Experiment 1: encoding of motor memories by control points

Two-control-points condition

In the two-control-points condition, we investigated whether participants could learn opposing (CW and CCW) force fields when each field was associated with a different control point on the object. The paradigm consisted of pre-exposure, exposure, generalization and de-adaptation ('washout') phases, as detailed below.

Pre-exposure, exposure and de-adaptation phases

Participants (n = 10) grasped the handle of the vBOT and a virtual object (solid red rectangle of 16 cm × 3 cm) was displayed centred on the hand (Fig. 3.1b). The position of the object translated with the hand. The object had two potential control points (green discs with a radius of 0.4 cm) ± 7 cm lateral to the centre of the object.

To start each trial, participants aligned the centre of the object (indicted by a yellow cross) with the home position (a circle with a radius of 0.5 cm) situated in the midline approximately 30 cm in front of the participant's chest. The trial started after the centre point was within 0.3 cm of the home position and had remained below a speed of 0.5 cm s^{-1} for 100 ms. After a 0.2 s delay, a target (a circle with a radius of 0.5 cm) appeared 12 cm away (anteriorly along the y axis) above either the left or right control point. A tone indicated that the participants should initiate a reaching movement to

the target. Participants were instructed to move the corresponding control point to the target. That is, if the target was aligned with the left control point, they should move the left control point to the target, and conversely for the target aligned with the right control point. Crucially, because each target was aligned with its respective control point, the hand moved to the same location to attain both targets.

A trial ended when the control point had remained within 0.3 cm of the target for 50 ms. If the peak speed of the movement was less than 50 cm s^{-1} or more than 70 cm s^{-1}, a low-pitch tone sounded and a 'too slow' or 'too fast' message was displayed, respectively. At the end of each trial, the vBOT actively returned the hand to the home position.

After familiarization, participants performed blocks of eight trials in which one of the trials (not the first) was randomly chosen as a channel trial and the remainder were either null-field trials or force-field trials. The paradigm consisted of a pre-exposure phase (4 blocks/32 trials), an exposure phase (52 blocks/416 trials), a generalization phase (71 blocks/568 trials; see below) and a de-adaptation phase (12 blocks/96 null-field trials). Each target appeared an equal number of times within a block in a pseudorandom order.

During the exposure phase, the location of the target (left or right) determined the direction of the force field (CW or CCW) applied during the movement. As such, the field direction was associated with the particular control point on the object (left or right) and its corresponding target (Fig. 3.3a). The direction of the force field for each control point was counterbalanced across participants. A 45 s rest break was given after blocks 12, 25, 38 and 51 of the exposure phase.

Generalization

The generalization phase (immediately after the exposure phase) was used to investigate the representation of the memories for the two force fields. Specifically, we tested

whether the representations were object-centred (associated with the location of the control points on the object), body-centred (associated with the locations of the targets/control points relative to the body) or hand-centred (associated with the location of the targets/control points relative to the hand). The generalization phase included force-field trials (to maintain adaptation), as well as channel trials (to monitor ongoing adaptation and examine generalization). To monitor adaptation, participants performed channel trials with the hand and object in the original exposure configuration. To examine generalization, participants performed channel trials during lateral translations ($\pm$ 14 cm) of either the object or the hand, or both the object and the hand. This gave three kinds of generalization trial (Fig. 3.7a).

On object-and-hand translation trials, participants translated the hand and the object $\pm$ 14 cm along the x axis before the trial started. For leftward translations of the object and hand, the left target was cued and the participant moved the right control point toward the target (which was now aligned with the left target). The converse was true for rightward translations. Importantly, as on all trials in the study, the control points and targets were always aligned along the y axis, thus requiring a straight movement.

On hand translation trials, participants translated their hand $\pm$ 14 cm along the x axis while the object remained in the centre of the workspace. For leftward translations of the hand, the left target was cued and the participant moved the left control point toward the target. The converse was true for rightward translations.

On object translation trials, the object was visually translated $\pm$ 14 cm along the x axis while the hand remained in the centre of the workspace. For leftward translations of the object, the left target was cued and the participant moved the right control point toward the target. The converse was true for rightward translations.

In all conditions, the object tracked the hand during the movement with the fixed offset associated with the particular generalization type maintained throughout the trial.

A 2.5 s delay was imposed at the start of each channel trial to allow time for the participant to translate their hand and/or the object, if required. A 15 s rest break was given every 9 blocks during the generalization phase starting at block 8. A block of force-field trials (with no channel trials) followed each rest break to mitigate the effects of any time-dependent decay of motor memory. Each of the channel types (2 to monitor ongoing adaptation and 6 for generalization) was repeated 8 times and there were 7 pure exposure blocks post-rest, giving 71 blocks (568 trials).

Note that the familiarization phase in the two-control-points condition included trials for each of the generalization conditions described above.

Single-control-point condition

In the single-control-point condition, we examined whether the visual targets alone could provide a contextual cue for learning. A separate group of participants (n = 8) performed the single-control-point condition task. As for the two-control-points condition (described above), the left and right control points were displayed on the object. Similarly, the left and right targets were displayed and were always associated with the particular field direction (CW or CCW). However, participants were instructed to move the central control point (yellow cross) on the object to a central target (which remained on the screen at all times) (Fig. 3.3b) and to ignore the lateral targets. Participants were thus required to control a single control point. The paradigm consisted of the same pre-exposure (4 blocks/32 trials), exposure (51 blocks/408 trials) and de-adaptation (12 blocks/96 trials) phases as for the two-control-points condition. However, in this condition the generalization phase was omitted.

3.2.4 Gaze fixation control

During target-directed reaches, people tend to fixate the target (Johansson et al., 2001; Land and Furneaux, 1997; Land et al., 1999). The gaze fixation experiment (n = 8) was performed to exclude the possibility that gaze direction, rather than the control point on the object, was the contextual cue responsible for adaptation to opposing force fields.

An EyeLink 1000 eye tracker was used to track the position of the left eye at 500 Hz. The eye tracker was mounted at an oblique angle between the handle of the manipulandum and a transparent reflective sheet of plastic that was used to project visual feedback (replacing the semi-silvered mirror). A forehead rest was used to stabilize the head. The eye tracker was calibrated at the start of the experiment and after every rest break using a nine-point grid.

The paradigm included both fixation and free-gaze trials. On fixation trials, participants had to fixate a small cross (located 8 cm above the home position) while performing the task (moving the left or right control point to the left or right target, respectively). Because movements were guided by peripheral vision, the spatial tolerance of the target was increased to 0.6 cm. If the participant broke fixation ($>$ 2 cm), the trial was aborted and repeated. The central control point on the object was not displayed to prevent participants from using it to guide the movement towards the fixation point. The home position was indicated by a rectangular frame that was visible at all times during the experiment. On free-gaze trials, no fixation cross was displayed and eye movements were unconstrained.

After familiarization (described above), participants performed a fixation pre-exposure phase (4 blocks/32 trials), a fixation exposure phase (65 blocks/520 trials), a free-gaze exposure phase (64 blocks/512 trials) and a free-gaze de-adaptation phase (12 blocks/96 trials). The direction of the force field associated with each control point

was counterbalanced across participants. Each block had the same structure as for the two-control-points condition, with seven exposure trials and one channel trial.

A 45 s rest break was given every 13 blocks during the exposure phase starting at block 12. A block of force-field trials (with no channel trials) followed each rest break to mitigate the effects of any time-dependent decay of motor memory.

3.2.5 Uniform object control

The object used in the main experiment consisted of different visual elements (green circular control points embedded in a red rectangle) that could conceivably be interpreted as distinct objects. We therefore repeated the exposure phase in the two-control-points group using an object that no longer contained different visual elements. A separate group of participants (n = 8) controlled a new uniform object with control points that were not discriminable from the rest of the object by colour (Fig. 3.6). The red rectangle was removed from the original object and the control points (green discs with a radius of 0.4 cm) were connected by a bar (solid green rectangle of 14 cm $\times$ 0.6 cm) that was the same colour as the control points. Importantly, these control points were the same size and in the same location as those of the original experiment. We reduced the size of the central cross so that it would fit inside the connecting bar.

As in the two-control-points experiment (described above), participants moved the corresponding control point to the displayed target, which was always associated with a particular field direction (CW or CCW). The paradigm consisted of the same pre-exposure (4 blocks/32 trials), exposure (51 blocks/408 trials) and de-adaptation (12 blocks/96 trials) phases as in the single-control-point experiment.

3.2.6 Experiment 2: are separate motor memories obligatory for different control points?

This experiment was performed to investigate whether separate motor memories are formed even when the dynamics experienced at the two control points on the object are the same.

After familiarization, participants performed a pre-exposure phase, an exposure phase and a de-adaptation phase. The pre-exposure (4 blocks/32 trials) and exposure phases (51 blocks/408 trials) were the same as in experiment 1. Participants performed blocks of eight trials consisting of seven null (preexposure) or field (exposure) trials and one channel trial. In the exposure phase, half the participants (n = 8) experienced opposing fields at each control point (as in experiment 1) and the other half (n = 8) experienced the same field at each control point. In both the opposing-field and same-field conditions, the direction of the force field for each control point was counterbalanced across participants. A 45 s rest break was given after blocks 12, 25 and 38 of the exposure phase. A block of force-field trials (with no channel trials) followed each rest break to mitigate the effects of any time-dependent decay of motor memory.

The exposure phase was followed by 50 blocks of de-adaptation, with each block consisting of 1 null-field trial and 1 channel trial. Importantly, one control point was cued throughout the first half of the de-adaptation phase (the first 25 blocks/50 trials) and the other control point was cued throughout the second half (the last 25 blocks/50 trials). The order in which each control point (left or right) was cued was counterbalanced across participants.

If the dynamics at both control points is encoded by a single representation, de-adaptation at one control point should also lead to de-adaptation at the other control point. Alternatively, if the representations are separate, de-adaptation at one control

point should have minimal effect on the representation associated with the other control point.

3.2.7 Data analysis

Data analysis was performed using MATLAB R2017a. Two measures of performance were calculated. On null-field and force-field trials, we calculated the maximum distance between the path of the hand and a straight line connecting the start position and the target (that is, the MPE). On channel trials, we calculated the percentage of the force field that was compensated for (adaptation), as described in Chapter 2.

To combine MPE and adaptation results across participants, data from individual participants were sign adjusted according to the direction of the field.

To identify changes in MPE and adaptation within and between experiments, mixed-design ANOVAs and paired t-tests were performed. Comparisons of the final level of adaptation between experiments were performed using between-subjects ANOVAs and unpaired t-tests. All statistical tests were two-sided with significance set to $P < 0.05$. Where values are reported, they represent the mean $\pm$ s.e.m.

3.2.8 Models

We fit two classes of model to experiment 2. The first class was a deterministic, context-dependent state-space model (SSM) based on (Kim et al., 2015; Lee and Schweighofer, 2009; Nozaki and Scott, 2009). The second class was a probabilistic switching state-space model (SSSM). We did not fit the models to experiment 1 as this experiment was not designed to distinguish between the model classes. Note, however, that the initial part of experiment 1 was identical to the initial part for the opposing-field group in experiment 2.

Context-dependent SSM

We used a context-dependent SSM that included separate states for different contexts and also different rates of adaptation (slow versus fast adaptive processes). In the current study, there were two contexts (the left and right control points) and each context could include states that updated with slow and fast rates. This gave a total of four states (two contexts and two rates) that could be represented by the elements of a state vector

$$\boldsymbol{x} = (x_{\mathrm{s}}^{(1)}\ x_{\mathrm{f}}^{(1)}\ x_{\mathrm{s}}^{(2)}\ x_{\mathrm{f}}^{(2)})^{\intercal}, \tag{3.2}$$

where the superscripts represent which context the state is associated with (for example, 1 = left and 2 = right control points), the subscripts represent the slow (s) and fast (f) adaptation processes, and $^{\intercal}$ is the transpose operator. The motor output on trial t was a weighted sum of the elements in the state vector, with the weighting determined by the context:

$$y(t) = \boldsymbol{x}(t)^{\mathrm{T}} \cdot \boldsymbol{c}(t), \tag{3.3}$$

where the context weighting vector, $\boldsymbol{c}$, varies according to the context for the trial (left or right control point). The error on a trial is the difference between the motor output and the task perturbation:

$$e(t) = f(t) - y(t). \tag{3.4}$$

In the current study, the task perturbation, f, is zero for null-field trials, +1 for CW field trials and -1 for CCW field trials. The state vector is updated across trials:

$$\boldsymbol{x}(t+1) = \boldsymbol{a} \odot \boldsymbol{x}(t) + e(t) \cdot \boldsymbol{b} \odot \boldsymbol{c}(t), \tag{3.5}$$

where $\odot$ represents element-wise multiplication. Trial-by-trial decay is determined by the retention vector $\boldsymbol{a}$. Error-dependent adaptation of states is determined by the context weighting vector and the learning-rate vector $\boldsymbol{b}$.

Table 3.1. Variants of the context-dependent state-space model

Model	DOF	Retention vector $\boldsymbol{a}$	Learning-rate vector $\boldsymbol{b}$	Context vector $\boldsymbol{c}$	Effective states
SSM1	3	$\begin{pmatrix} a_\mathrm{s} & 0 & a_\mathrm{s} & 0 \end{pmatrix}^\intercal$	$\begin{pmatrix} b_\mathrm{s} & 0 & b_\mathrm{s} & 0 \end{pmatrix}^\intercal$	$\boldsymbol{c}^{(1)} = \begin{pmatrix} 1 & 0 & c_\mathrm{s} & 0 \end{pmatrix}^\intercal$ $\boldsymbol{c}^{(2)} = \begin{pmatrix} c_\mathrm{s} & 0 & 1 & 0 \end{pmatrix}^\intercal$	2
SSM2	3	$\begin{pmatrix} a & a & a & 0 \end{pmatrix}^\intercal$	$\begin{pmatrix} b_\mathrm{s} & b_\mathrm{f} & b_\mathrm{s} & 0 \end{pmatrix}^\intercal$	$\boldsymbol{c}^{(1)} = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix}^\intercal$ $\boldsymbol{c}^{(2)} = \begin{pmatrix} 0 & 1 & 1 & 0 \end{pmatrix}^\intercal$	3
SSM3	4	$\begin{pmatrix} a_\mathrm{s} & a_\mathrm{f} & a_\mathrm{s} & 0 \end{pmatrix}^\intercal$	$\begin{pmatrix} b_\mathrm{s} & b_\mathrm{f} & b_\mathrm{s} & 0 \end{pmatrix}^\intercal$	$\boldsymbol{c}^{(1)} = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix}^\intercal$ $\boldsymbol{c}^{(2)} = \begin{pmatrix} 0 & 1 & 1 & 0 \end{pmatrix}^\intercal$	3
SSM4	5	$\begin{pmatrix} a_\mathrm{s} & a_\mathrm{f} & a_\mathrm{s} & 0 \end{pmatrix}^\intercal$	$\begin{pmatrix} b_\mathrm{s} & b_\mathrm{f} & b_\mathrm{s} & 0 \end{pmatrix}^\intercal$	$\boldsymbol{c}^{(1)} = \begin{pmatrix} 1 & 1 & c_\mathrm{s} & 0 \end{pmatrix}^\intercal$ $\boldsymbol{c}^{(2)} = \begin{pmatrix} c_\mathrm{s} & 1 & 1 & 0 \end{pmatrix}^\intercal$	3
SSM5	5	$\begin{pmatrix} a_\mathrm{s} & a_\mathrm{f} & a_\mathrm{s} & a_\mathrm{f} \end{pmatrix}^\intercal$	$\begin{pmatrix} b_\mathrm{s} & b_\mathrm{f} & b_\mathrm{s} & b_\mathrm{f} \end{pmatrix}^\intercal$	$\boldsymbol{c}^{(1)} = \begin{pmatrix} 1 & 1 & 0 & c_\mathrm{f} \end{pmatrix}^\intercal$ $\boldsymbol{c}^{(2)} = \begin{pmatrix} 0 & c_\mathrm{f} & 1 & 1 \end{pmatrix}^\intercal$	4
SSM6	5	$\begin{pmatrix} a_\mathrm{s} & a_\mathrm{f} & a_\mathrm{s} & a_\mathrm{f} \end{pmatrix}^\intercal$	$\begin{pmatrix} b_\mathrm{s} & b_\mathrm{f} & b_\mathrm{s} & b_\mathrm{f} \end{pmatrix}^\intercal$	$\boldsymbol{c}^{(1)} = \begin{pmatrix} 1 & 1 & c_\mathrm{s} & 0 \end{pmatrix}^\intercal$ $\boldsymbol{c}^{(2)} = \begin{pmatrix} c_\mathrm{s} & 0 & 1 & 1 \end{pmatrix}^\intercal$	4
SSM7	6	$\begin{pmatrix} a_\mathrm{s} & a_\mathrm{f} & a_\mathrm{s} & a_\mathrm{f} \end{pmatrix}^\intercal$	$\begin{pmatrix} b_\mathrm{s} & b_\mathrm{f} & b_\mathrm{s} & b_\mathrm{f} \end{pmatrix}^\intercal$	$\boldsymbol{c}^{(1)} = \begin{pmatrix} 1 & 1 & c_\mathrm{s} & c_\mathrm{f} \end{pmatrix}^\intercal$ $\boldsymbol{c}^{(2)} = \begin{pmatrix} c_\mathrm{s} & c_\mathrm{f} & 1 & 1 \end{pmatrix}^\intercal$	4

We examined 7 context-dependent state-space models, which varied in their number of parameters (DOF). The models could vary in the retention parameters, learning rates and coupling. The setting of the parameters determine the effective number of states in each model. Note that models 2-4 do not map the states on to contexts 1 and 2 in equation 3.2 as one of the states is shared by both contexts (i.e., context independent). Furthermore, model 2 does not necessarily map the states on to the fast and slow states in equation 3.2 as the context-independent state can be either fast or slow.

We considered seven variants of the model, which determine the way $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ are parameterized (Table 3.1). The models vary as to whether there is a single context-dependent state (SSM 1), a slow/fast context-independent state and a fast/slow context-dependent state (SSMs 2-4), or both fast and slow context-dependent states (SSMs 5-7). For context-dependent states, the values in the context weighting vector determine the coupling (if any) between contexts. Models that include parameters

c_s and c_f can have coupling between the contexts for the slow and/or fast processes, respectively.

SSSM

The generative model

Here, we develop a probabilistic SSSM that allows the sensorimotor system to optimally partition learning across contexts. We assume that the environment consists of distinct contexts that switch in a probabilistic manner according to a hidden Markov model (Fig. 3.2). Each context is associated with its own perturbations, which evolve according to linear-Gaussian dynamics. We assume that participants possess an internal generative model of perturbations that matches the environment. We characterize this model as an SSSM—a generalization of the SSM to systems with multiple operating regimes or modes. SSSMs define a probability density over a time series of discrete and continuous states. Here, the discrete states represent contexts and the continuous states represent perturbations. We implement the model with two contexts, although the extension to more than two contexts is straightforward.

On trial t, the environment is in one of two contexts, $c_t \in \{1, 2\}$, and between trials the context either remains the same or switches in a stochastic manner. The context transition probabilities are defined by the matrix $\mathbf{\Pi}$, where

$$\pi_{ij} = p(c_t = j | c_{t-1} = i). \tag{3.6}$$

To reduce the number of free parameters in the model, we assume that the context remains the same or switches with equal probability, that is $\pi_{ij} = 0.5$ for all i, j. This is consistent with the exposure phase of our experiment in which each force field is presented with equal frequency and in a pseudorandom order.

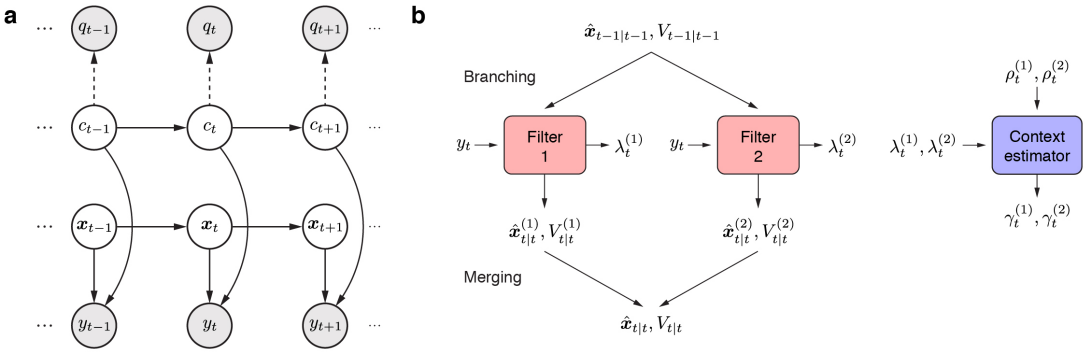


Figure 3.2. Motor adaptation as online state and parameter estimation in a switching state-space model. **a**, The graphical model. The discrete context, c_t, and the continuous perturbations, x_t, evolve according to Markovian dynamics. The context determines both which observable cue, q_t, is emitted and how each perturbation contributes to the continuous observation y_t. Observed and latent variables are represented by gray and white nodes, respectively. The dashed arrows indicate dependencies that are learned online. **b**, The GPB1 algorithm. Given a single estimate of the state on trial $t-1$, each Kalman filter predicts the state on trial t and then updates its prediction when y_t is observed. The updated estimates of each filter are then merged into a single estimate. The probability of each context, which depends on both the contextual cue and the prediction error of each filter, determines how to merge the estimates in an optimal manner.

On trial t, the context gives rise to one of two observable cues, $q_t \in \{1, 2\}$, corresponding to the location of the control point. The cue emission probabilities are defined by the matrix $\boldsymbol{\Phi}$, where

$$\phi_{kj} = p(q_t = k | c_t = j). \tag{3.7}$$

The perturbations in each context evolve independently according to linear-Gaussian dynamics, as defined by the state transition function

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{w}_t, \tag{3.8}$$

where $\boldsymbol{x}_t = (x_{\mathrm{s},t}^{(1)} \; x_{\mathrm{f},t}^{(1)} \; x_{\mathrm{s},t}^{(2)} \; x_{\mathrm{f},t}^{(2)})^\intercal$ is a vector of slow and fast states (subscripts) for contexts 1 and 2 (superscripts), $\boldsymbol{A}$ is a diagonal matrix of the form $\mathrm{diag}(a_\mathrm{s}, a_\mathrm{a}, a_\mathrm{s}, a_\mathrm{a})$ and $\boldsymbol{w}_t$ is zero-mean Gaussian process noise with covariance $\boldsymbol{Q} = \mathrm{diag}(q_\mathrm{s}, q_\mathrm{f}, q_\mathrm{s}, q_\mathrm{f})$. Thus, we implement a probabilistic context-dependent dual-rate model.

The context determines which perturbations are observed, as defined by the output function

$$y_t = \boldsymbol{h}^{(c_t)} \boldsymbol{x}_t + v_t, \tag{3.9}$$

where $\boldsymbol{h}^{(1)} = (1\ 1\ 0\ 0)$ is the observation vector for context 1, $\boldsymbol{h}^{(2)} = (0\ 0\ 1\ 1)$ is the observation vector for context 2 and v_t is zero-mean Gaussian observation noise with variance r.

State inference

To operate effectively, the sensorimotor system must keep track of perturbations in the environment, as well as the contexts in which they occur. However, both perturbations and contexts are latent variables, which must be inferred from a sequence of observations $\{y_{1:t}, q_{1:t}\}$. In Appendix B, we present the algorithm that we use to perform state inference in the switching state-space model. Figure 3.2b shows a schematic of the algorithm.

Parameter learning

To perform state inference, the sensorimotor system must set the parameters, $\theta = \{\boldsymbol{A}, \boldsymbol{Q}, \boldsymbol{h}^{(1)}, \boldsymbol{h}^{(2)}, r, \boldsymbol{\Pi}, \boldsymbol{\Phi}\}$, of the SSSM. Importantly, these parameters need not be fixed but may be learned online. In general, the sensorimotor system could learn all of these parameters online. However, to reduce the degrees of freedom of the model for our data set, we assume that the sensorimotor system only learns the cue emission probabilities and that all other parameter estimates are fixed. In Appendix B, we present the algorithm that we use to perform parameter learning in the switching state-space model.

Model fitting

We simulated each model separately for each participant using the particular sequence of cues they observed. We then calculated the mean model data across participants for each group and each control point. We assume that adaptation measured on channel trials represents predicted observations that have been corrupted by independent and identically distributed Gaussian motor noise. Maximum-likelihood estimates of the free parameters of the models (Table 3.2) were therefore obtained by minimizing the sum-of-squared errors between the mean model data (equation 3.3 for the SSMs and equation B.2 for the SSSM) and the mean experimental data measured on channel trials. As each model involved fitting both same- and opposing-field groups jointly with the same set of parameters, we chose to fit the mean adaptation data across participants to reduce variability in the data set.

Model comparison

Model selection was performed using the Bayesian Information Criterion (BIC):

$$\mathrm{BIC} = n \cdot \ln[(1 - R^2)/n] + k \cdot \ln(n) + \text{const.}, \tag{3.10}$$

where n (204) is the number of data points, k is the number of free parameters and the constant does not depend on the model. The first term in the BIC penalizes underfitting, whereas the second term penalizes model complexity, as measured by the number of free parameters in the model. Taking the difference in BIC values for two competing models approximates twice the log of the Bayes factor (Kass and Raftery, 1995). A BIC difference of greater than 4.6 (a Bayes factor of greater than 10) is considered to provide strong evidence in favour of the model with the lower BIC value (Jeffreys, 1998).

3.3 Results

We carried out two experiments to examine the mapping of motor memories to control points. In experiment 1, we examined the ability of participants to learn opposing dynamics when they are associated with either different control points or a single control point. We then examined the reference frame in which control points are represented by testing generalization of learning. In experiment 2, we tested whether the assignment of separate memories to control points is obligatory or whether this only occurs when necessary.

Participants grasped the handle of a robotic interface (Fig. 3.1a), and a semi-silvered mirror was used to display a rectangular virtual object centred on the hand. The object translated with the hand. The workspace under the mirror was illuminated so that both the hand and the object were always visible (Fig. 3.1b). The object had three visible control points: circles displayed on the left and right of the object and a cross at the centre of the object. For participants in the two-control-points group (n = 10), a target appeared above one of the two lateral control points at the start of each trial and participants were instructed to move the corresponding control point to the target (Fig. 3.3a). The target location (left or right) cued the direction of the force field on exposure trials (see below). For participants in the single-control-point group (n = 8), a central target appeared above the central control point and participants were instructed to move the central control point to this target (Fig. 3.3b). In this group, as in the two-control-points group, a lateral 'target' was also displayed to cue the direction of the force field. Crucially, for both groups, the start position of the hand and the required final position of the hand were the same in all trials.

In the pre-exposure phase, participants made movements in a null field. In the exposure phase, opposing force fields were associated with the two lateral targets. Hence, for participants in the two-control- points group, each force field was also associated with

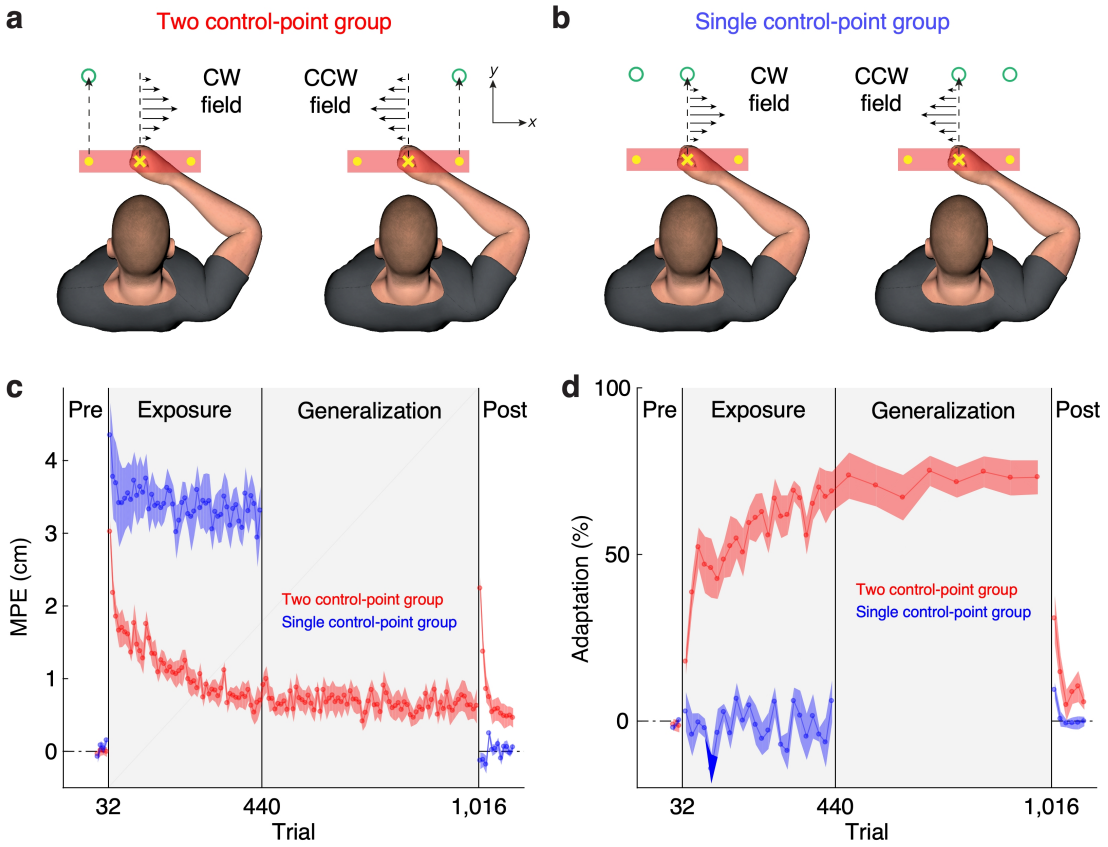


Figure 3.3. Separate motor memories are formed for different control points. Learning curves for experiment 1. **a**, Participants in the two-control-points group moved the left or right control point to its corresponding target. The direction of the force field applied to the hand was determined by the target/control point. **b**, Participants in the single-control-point group moved the central control point to a central target. The direction of the field applied to the hand was cued by which lateral 'target' was displayed. **c**, Kinematic error measured as the MPE (deviation from a straight line to the target) over the course of the experiment. Data were first averaged over blocks of eight exposure trials and then plotted as mean ± s.e.m. across participants. The grey background shows the period when the force field was turned on. **d**, Same as **c**, but for percentage adaptation measured on channel trials. The generalization phase was performed by the two-control-points group (n = 10), but omitted for the single-control-point group (n = 8).

the point on the object that had to be controlled. To produce the same movement of the hand in both force fields, participants needed to learn the opposing dynamics.

Participants performed 52 blocks of 8 trials (416 trials) in the exposure phase with an equal number of trials for each target in a pseudorandom order. So that predictive

force compensation could be assessed independently from co-contraction, one randomly selected trial in each block was a channel trial (Milner and Franklin, 2005; Scheidt et al., 2000), in which the movement was confined to a simulated mechanical channel to the target. On these trials, the forces generated by participants into the wall of the channel could be measured.

To assess performance, we measured kinematic error, defined as the maximum perpendicular error (MPE), between the actual and ideal (that is, straight) hand path. We also measured force compensation on channel trials, defined as the percentage of the force required to fully compensate for the force field (adaptation). During the pre-exposure phase, as expected, the kinematic error and adaptation were close to zero (Fig. 3.3c,d). At the start of the exposure phase, the force fields produced substantial deviations of the hand path from a straight line (Fig. 3.4). To assess learning, we compared adaptation, which unlike MPE is not influenced by stiffness, in the final two blocks of the pre-exposure and exposure phases. This revealed a significant interaction between group and epoch (mixed-design analysis of variance (ANOVA), $F_{1,16} = 73.31, P = 2 \times 10^{-7}$). Adaptation reached 68.1 ± 6.6% and -0.2 ± 4.3% of full force compensation by the end of the exposure phase for the two- and single-control-point groups (Fig. 3.3d), respectively. For the two-control-points group, the increase in adaptation was significant (paired t-test, 69.4 ± 6.7%, $t_9 = 10.99, P = 2 \times 10^{-6}$).

The failure to learn when controlling a single control point is consistent with previous studies showing that arbitrary visual cues do not facilitate learning of opposing force fields (Gandolfo et al., 1996; Howard et al., 2012, 2013), and indicates that the learning observed when controlling two different control points—despite making similar hand movements—is not simply due to the contextual cues provided by the lateral targets.

We performed two additional control experiments. In the first (Fig. 3.5), we show that participants (n = 8) can learn opposing fields, linked to different control points, even when they are required to fixate a central fixation point. This rules out the possibility

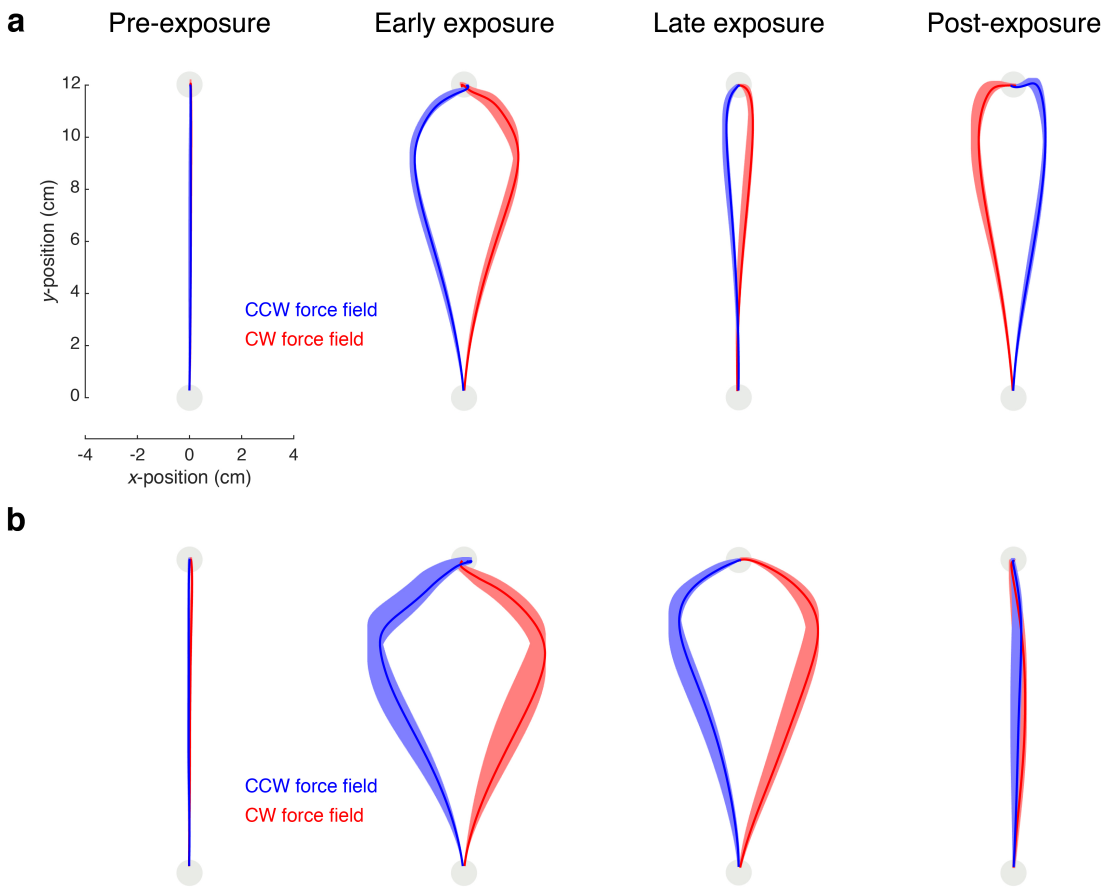


Figure 3.4. Hand paths at different stages of Experiment 1. a,b, In early exposure, the force fields produced substantial deviations of the hand path from a straight line. For the two-control-points groups (n = 10) (**a**), hand paths were significantly straighter by late exposure, and clear after-effects were seen in early post-exposure. For the single-control-point group (n = 8) (**b**), hand paths were not significantly straighter by late exposure, and no after-effects were seen. Traces show mean ± s.e.m. across participants hand paths over two blocks.

that learning is due to associating field direction with gaze position. The second control was designed to rule out the possibility that subjects might be treating the two control points (small circles) as separate objects, despite the fact that they move coherently and are bound by a rectangle. In this control, we used an object where the control points were defined by the geometry of the object and there were no distinct visual elements (Fig. 3.6). We show that participants (n = 8) can still learn opposing fields linked to these different control points (Fig. 3.6).

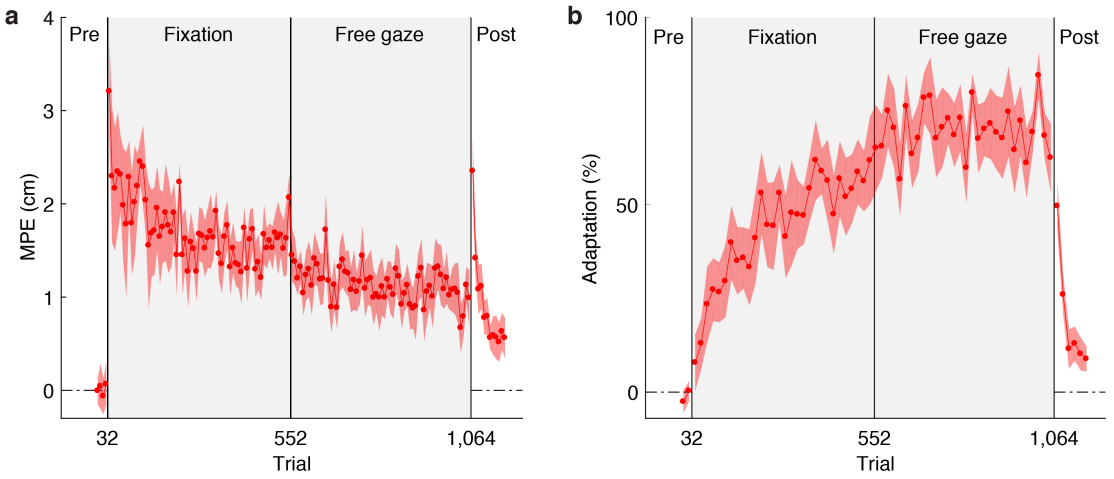


Figure 3.5. Adaptation is not gaze dependent. In the fixation stage of the exposure phase, participants (n = 8) were required to fixate a cross in the midline of the workspace. In the free-gaze stage of the exposure phase, no fixation cross was displayed and gaze was unconstrained. **a,b**, Mean ± s.e.m. (**a**) MPE and (**b**) adaptation across all participants. Each MPE data point represents the mean of a block of eight trials. To assess learning, we compared adaptation in the final two blocks of the pre-exposure phase with the final two blocks of the fixation stage. We found a significant increase in adaptation (paired t-test, $60.2 \pm 8.4\%, t_7 = 7.70, P = 1 \times 10^{-4}$), which reached $59.0 \pm 7.9\%$ (mean ± s.e.m.) of full compensation by the final two blocks of the fixation stage. We contrasted the adaptation reached by blocks 23-24 of the exposure phase with the two-control-points and opposing-field groups who experienced the same task with free gaze. This revealed no significant difference (between-subjects ANOVA, $F_{2,23} = 2.07$, $P = 0.149$), indicating that fixation did not detriment learning. We allowed the fixation group to continue exposure after the fixation phase but with free gaze. There was no significant additional learning after 616 further trials (paired t-test, $6.4 \pm 10.4\%, t_7 = 0.67, P = 0.526$).

Our hypothesis that separate memories can be formed for different control points on manipulated objects assumes that these points are represented in object-centred coordinates; that is, that translating the object would not interfere with the learned memories associated with these points. However, other interpretations are possible based on the results for the two-control-points group. Specifically, whereas our hypothesis assumes that memories formed for the two fields are linked to the locations of the control points relative to the object (object-centred representation), it is possible that these memories are linked to the locations of the targets relative to the body

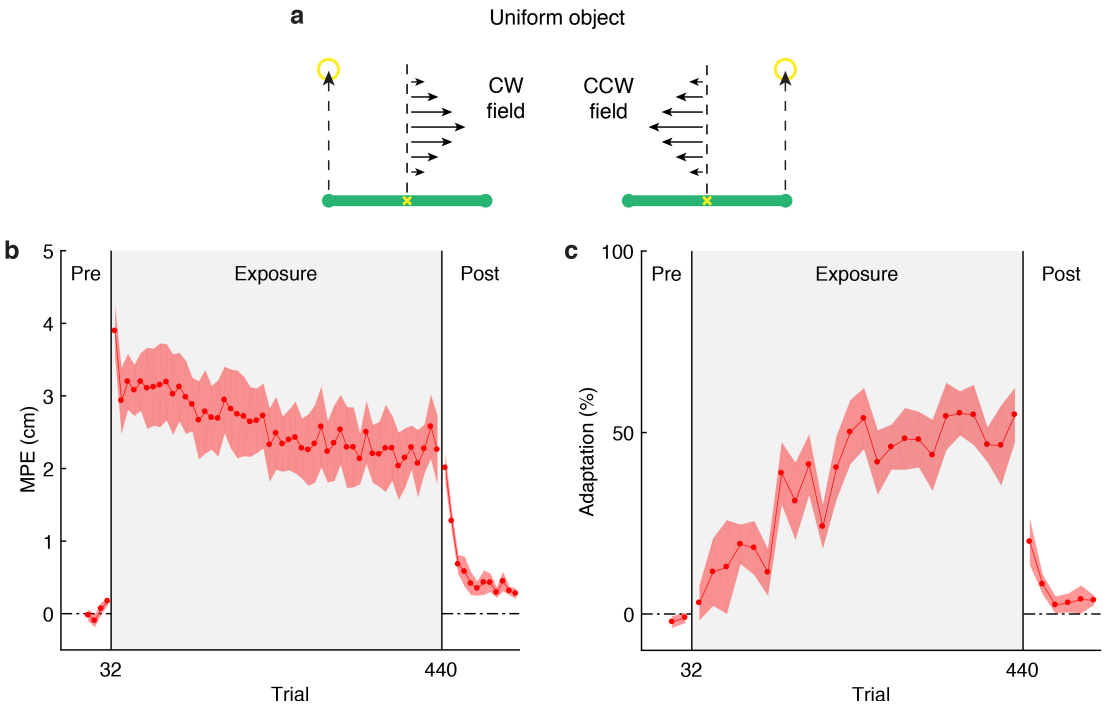


Figure 3.6. Adaptation does not require control points to be distinct from the rest of the object. a, Participants (n = 8) controlled a uniform object composed of two green disks connected by a green bar. Because the disks and bar were the same color, the object did not contain any distinct visual elements. **b,c**, Mean ± s.e.m. (**b**) MPE and (**c**) adaptation across all participants. Each MPE data point represents the mean of a block of eight trials. To assess learning, we compared adaptation in the final two blocks of the pre-exposure phase with the final two blocks of the exposure phase. We found a significant increase in adaptation (paired t-test, $52.3 \pm 9.3\%, t_7 = 6.08, P = 5 \times 10^{-4}$), which reached $50.7 \pm 8.5\%$ (mean ± s.e.m.) of full compensation by the final two blocks of the exposure phase. We contrasted the adaptation reached by this stage with the two-control-point and opposing-field groups who experienced the same task with control points displayed as green disks on a red rectangle. There was no significant difference (between-subjects ANOVA, $F_{2,23} = 2.11, P = 0.144$), indicating that the uniform object did not detriment learning.

(body-centred representation) or the locations of the targets and control points relative to the hand (hand-centred representation).

Based on our hypothesis and pilot data, we anticipated that participants in the two-control-points group would learn the two force fields. Therefore, for this group only, we included a generalization phase following the exposure phase in which we tested

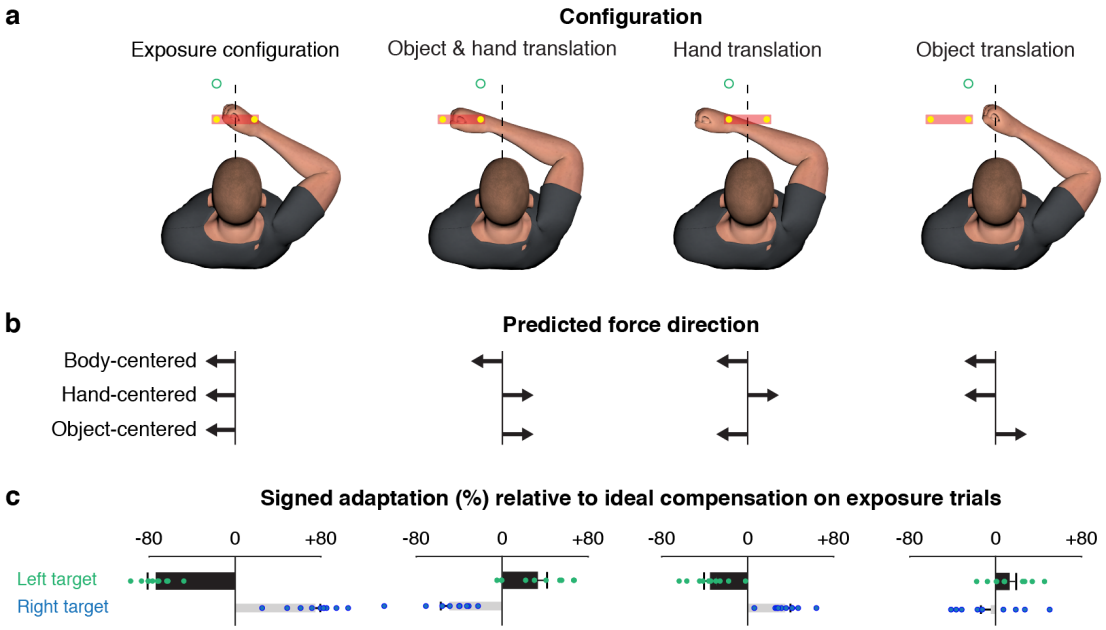


Figure 3.7. Control points are represented in an object-centred frame of reference. **a**, Configurations used to examine generalization of learning for trials with the left target. For trials with the right target all translations were reversed. All these trials were channel trials. **b**, Predictions of the direction of force according to the three possible representations. Note that these are for the scenario in which the left and right targets/control points were associated with the CW and CCW fields, respectively. The predictions for the right target (not shown) are in the opposite directions. **c**, Mean + s.e.m. signed adaptation across participants (n = 10) relative to the ideal compensation on exposure trials. The sign of the bar (+ or −) indicates compensatory forces appropriate for CCW and CW fields, respectively. Black and grey bars correspond to trials with the left and right targets, respectively. Dots represent data from individual participants.

the alternative interpretations outlined above. As shown in Fig. 3.3c,d, performance—measured by both the kinematic error and adaptation—was stable during this phase.

In the generalization phase, we used channel trials to measure the forces produced in four configurations (Fig. 3.7a) where: (1) the hand and object were in the original exposure configuration; (2) both the object and hand translated; (3) only the hand translated; and (4) only the object translated. Note that Fig. 3.7a shows the generalization trials involving the left target. In trials involving the right target, the hand and/or object were translated in the opposite direction. In all configurations, the object moved with

the hand with the appropriate offset throughout the movement. Fig. 3.7b shows the direction of force we expect the participants to generate for each possible representation (that is, body-, hand- or object-centred) for each configuration. All predictions are for the case in which the left and right targets/control points were associated with clockwise and counterclockwise, respectively. In the exposure configuration, we expect participants to generate forces in the appropriate (in this case, leftward) direction, and this is consistent with all three possible representations (Fig. 3.7b, left). Importantly, however, the pattern across the other three configurations is different for each possible representation.

The thick black bars in Fig. 3.7c show the measured force compensation (signed adaptation) in all four configurations relative to the force required to fully compensate for the force field in the exposure configuration. The thin grey bars show the force compensation in trials involving the right target, in which all predicted forces were in the opposite direction. Separate t-tests revealed that adaptation was significantly different from zero for both the left and right targets in the object-and-hand (one-sample t-test, left target: $32.6 \pm 9.2\%, t_9 = 3.75, P = 0.005$; right target: $-48.0 \pm 8.3\%, t_9 = 6.12, P = 2 \times 10^{-4}$) and hand translation configurations (one-sample t-test, left target: $34.4 \pm 5.9\%, t_9 = 6.14, P = 2 \times 10^{-4}$; right target: $34.5 \pm 5.0\%, t_9 = 7.28, P = 5 \times 10^{-5}$). Furthermore, the direction of adaptation in these configurations was only consistent with the object-based predictions. Although adaptation was not significantly different from zero for either target in the object translation configuration (one-sample t-test, left target: $12.5 \pm 6.5\%, t_9 = 2.03, P = 0.072$; right target: $-3.54 \pm 49.9\%, t_9 = 0.38, P = 0.714$), the sign of adaptation was again consistent with the object-based predictions. Note that the magnitude of force compensation (that is, length of the bars) relative to the exposure configuration decreased in all three generalization configurations, involving translation of the hand and object, object or hand. This decrease is consistent with a large body of literature showing that motor learning is generally local, with graded reduction with changes in context (Berniker et al., 2013; Wolpert et al., 2011). Overall,

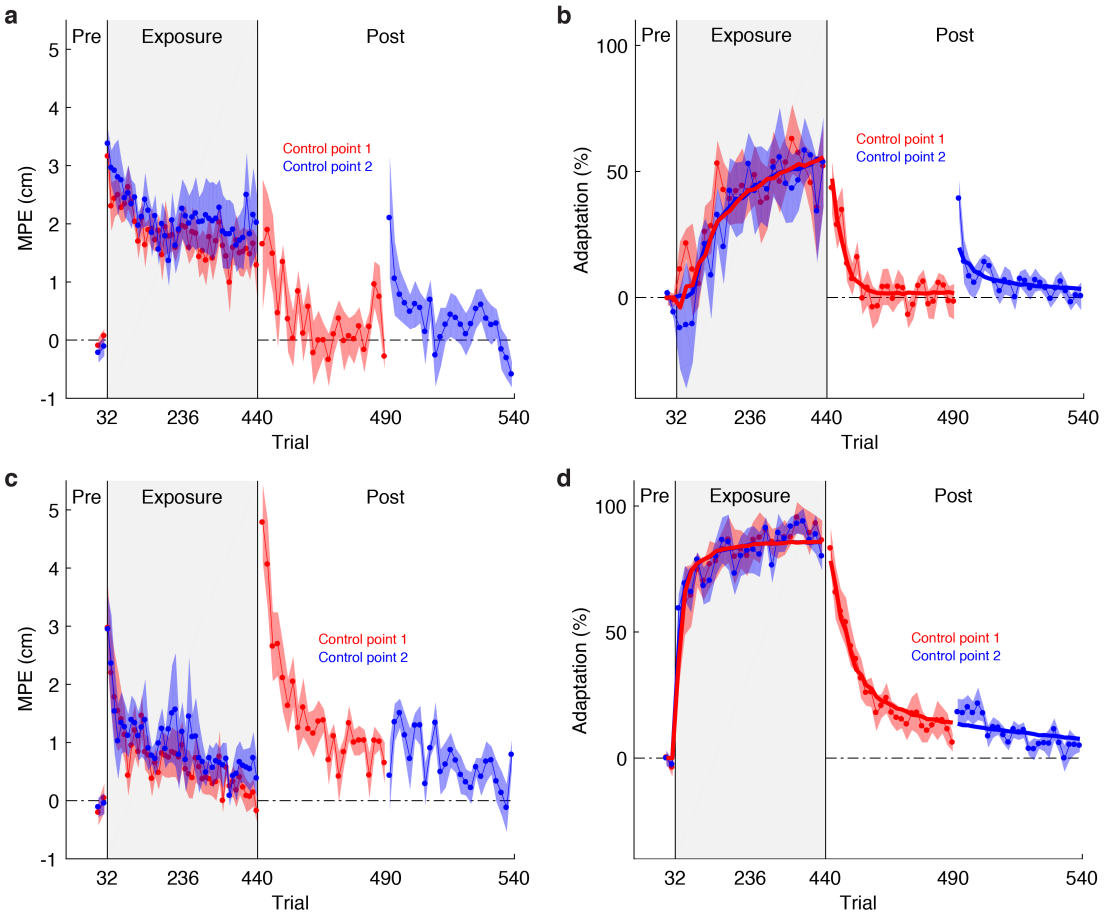


Figure 3.8. The formation of distinct motor memories for different control points is not obligatory. a,b, MPE ± s.e.m. (**a**) and adaptation ± s.e.m. (**b**) for each control point in the group that experienced opposing force fields at each control point (n = 8). Data in the pre-exposure (Pre) and exposure phases were first averaged over blocks of eight trials. The post-exposure phase (Post) has been expanded for clarity. For simplicity, the sign of adaptation for control point 2 has been inverted. **c,d**, Same as **a** and **b**, respectively, but for the group that experienced the same force field at each control point (n = 8). Solid lines in **b** and **d** show the mean SSSM fits across participants.

these results clearly support the hypothesis that separate memories are formed for control points represented in object-centred coordinates.

Having established that participants can form separate motor memories for different control points, we asked whether this process is obligatory or flexible. Specifically, we tested whether separate memories are formed even when the dynamics experienced

for the different control points is the same. In contrast with experiment 1, here, a single motor memory would be sufficient for representing the dynamics of both control points and would also support efficient allocation of neural resources. Two groups were examined: participants in the opposing-field group (n = 8) experienced a different (opposing) force field for each control point, whereas participants in the same-field group (n = 8) experienced the same force field for each control point. To assess the separation of motor memories, after learning we de-adapted one control point with null-field trials and then assessed the state of adaptation when controlling the other control point (Fig. 3.8). If the memories are separate, de-adapting one control point should not de-adapt the memory for the other control point.

Both groups learned to compensate for the perturbation (Fig. 3.8a,c, MPE; Fig. 3.8b,d, adaptation) and, as expected, learning was greater for the group who learned the same field for both control points. By the final block of the exposure phase, there was no significant difference in adaptation between the opposing-field group ($52.8 \pm 12.1\%$) and the two-control-points group (experiment 1) ($68.9 \pm 5.9\%$) who performed the same task (unpaired t-test, 16.1 $\pm 13.1\%, t_1 6 = 1.27, P = 0.222$). To assess whether de-adapting one memory led to de-adaptation of the other memory, we compared the mean of the last two channel trials in the post-exposure phase for the first control point with the first channel trial in the post-exposure phase for the second control point. For the group who learned opposing force fields, after de-adapting one control point, substantial adaptation re-emerged (Fig. 3.8b; paired t-test, 37.6 $\pm 8.6\%, t_7 = 4.69, P = 0.002$) when controlling the other control point. Although this effect is primarily driven by the first channel trial, this first trial provides a critical test for the separation of memories. This suggests that separate representations existed for each control point. In contrast, for the group that experienced the same force field for both control points, when we de-adapted one control point there was minimal adaptation remaining (Fig. 3.8d; paired t-test, 9.4 $\pm 7.8\%, t_7 = 1.31, P = 0.232$) when switching to the other control point. This was despite the greater adaptation observed in the same-field group, which increased the power to detect a difference in adaptation between the two control points

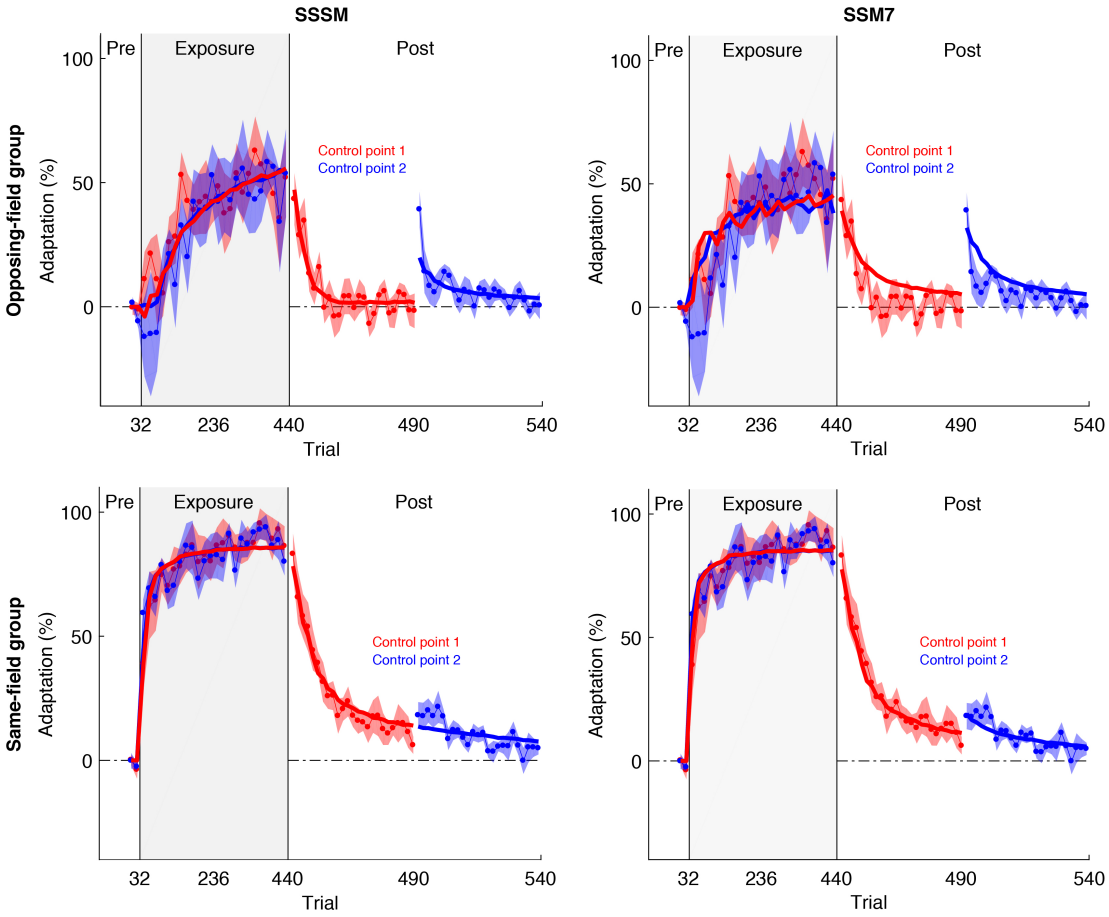


Figure 3.9. Model fits to adaptation data from Experiment 2. Solid lines show the mean fits across participants for the switching state-space model (SSSM) and state-space model 7 (SSM7), which had the next best BIC. Experimental data is shown as in Fig. 3.8b,d for the opposing-field (n = 8) and same-field (n = 8) groups. The superiority of the SSSM is most apparent in the fit to data from the opposing-field group.

during de-adaptation. This suggests that a single representation existed for both control points in this case.

We fit several candidate models to the data and performed model comparison. Context-dependent SSMs have been used to account for motor learning of opposing perturbations (Kim et al., 2015; Lee and Schweighofer, 2009; Nozaki and Scott, 2009). Briefly, these models assume that each context (force field) is associated with separate states of adaptation that are updated based on the error from the previous trial. Contexts

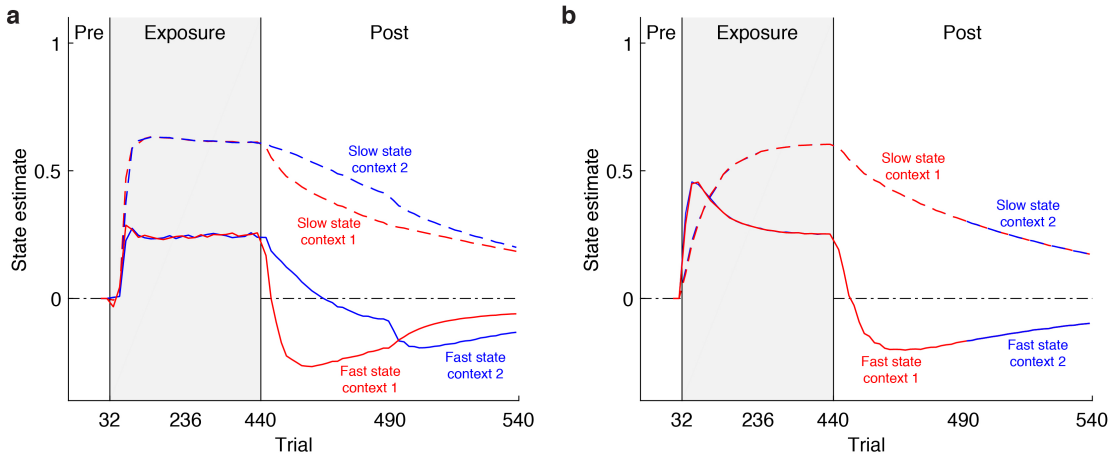


Figure 3.10. Slow and fast states of the switching state-space model. **a**, Mean state estimates across participants for the group that experienced opposing force fields at each control point. The slow and fast states of each context learn to represent different perturbations and, due to the learned association between contexts and cues, de-adapt at different rates during washout. For simplicity, the signs of the slow and fast states for context 2 have been inverted. **b**, For the group that experienced the same force field at each control point, the slow and fast states of each context represent the same perturbation and de-adapt at the same rate during washout.

can be coupled such that errors in one context update the states associated with the other context, and states associated with both contexts contribute to the final output. However, critically, these models assume that the coupling between contexts is known in advance and remains fixed over time. Therefore, they predict that: (1) with no coupling, two separate memories will be formed; (2) with complete coupling, a single memory will be formed; and (3) with partial coupling, two overlapping memories will be formed. While it is difficult to account for our results with either no coupling or complete coupling, it is possible that a context-dependent SSM with partial coupling could explain the data.

Context-dependent SSMs can be either single-rate (Ingram et al., 2011; Thoroughman and Shadmehr, 2000) or dual-rate, in which the total adaptation involves two memory processes: (1) a fast process that both adapts and decays quickly; and (2) a slow process that adapts and decays more gradually (Smith et al., 2006; Trewartha et al., 2014). We examined seven variants of the context-dependent SSM (Table 3.1) (Kim

Table 3.2. Comparison of model fits for the state-space (SSM) and switching state-space (SSSM) models

State-space models	R^2	ΔBIC	DOF	a_s	b_s	c_s	a_f	b_f	c_f
SSM1	0.92	135.3	3	0.9858	0.0730	0.4578	-	-	-
SSM2	0.92	135.3	3	0.9858	0.0215	[0]	0.9858	0.0668	[1]
SSM3	0.92	140.6	4	0.9858	0.0215	[0]	0.9858	0.0668	[1]
SSM4	0.93	117.3	5	0.9905	0.0445	0.4040	0.9229	0.0922	[1]
SSM5	0.92	145.9	5	0.9862	0.0007	[0]	0.9858	0.0721	0.4636
SSM6	0.93	129.6	5	0.9895	0.0502	0.5234	0.8860	0.0553	[0]
SSM7	0.94	112.4	6	0.9932	0.0295	0.4660	0.9405	0.0888	0.4340
Switching state-space model	R^2	ΔBIC	DOF	a_s	q_s	a_f	q_f	r	α
SSSM	0.96	0	6	0.9946	0.0003	0.9404	0.0068	0.0792	0.6067

Numbers in square brackets are fixed and not fit to the data. Note that although the R^2 values are similar these are independent of the number of data points, whereas the BIC scales with the number of data points. Therefore, the small improvement in R^2 for a model can translate into a large difference in likelihood and hence BIC. In SSM2, the fast state corresponds to the context-independent state shared by both contexts (see Table 3.1 for details).

et al., 2015; Nozaki and Scott, 2009), which vary as to whether single- or dual-rate states are included and also which states are coupled (see Methods) for details).

As an alternative, we also developed a normative SSSM (Ackerson and Fu, 1970; Chang and Athans, 1978; Shumway and Stoffer, 1991) of motor learning. The SSSM is a generalization of the SSM to systems with multiple operating modes (here, contexts), each of which can be associated with different, and even evolving, dynamics. On each trial, the system is in one context that determines both the dynamics experienced and the cues observed. Between trials, contexts switch stochastically. On each trial, the learner must infer the probability of each context (from control-point cues and prediction errors) and use this information to partition control and learning between two context-specific dual-rate states (see Methods) for details). Importantly, although the association between cues and contexts is deterministic (the control point perfectly

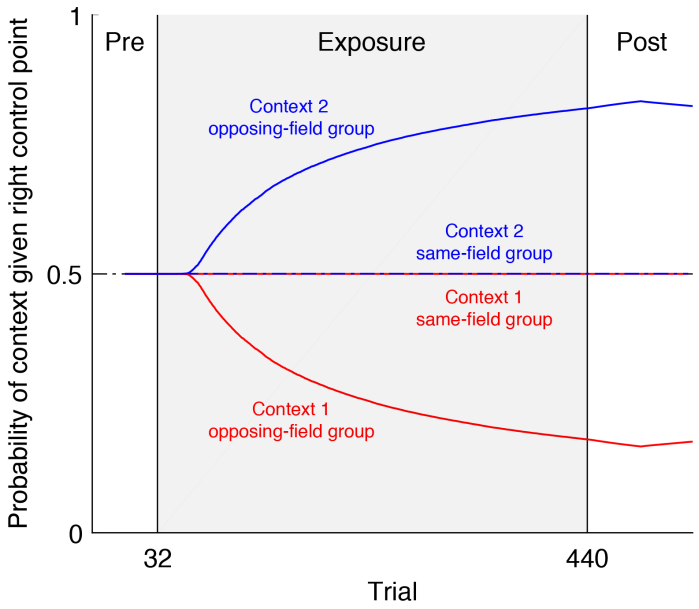


Figure 3.11. Learning to associate contexts with cues in the SSSM. The SSSM estimates the probability of each context given the cue (plotted here for the right control point cue before movement). Following exposure to opposing force fields, participants learn to differentiate between the two contexts given the right control point. Participants trained on the same force field at each control point do not learn to differentiate between contexts, leading to a single motor memory. Traces show mean model fits across participants. The plot for the left control point (not shown) is the mirror image about the line 0.5.

predicts the force-field direction), this association must be learned through experience. We therefore reasoned that the formation of separate motor memories may arise from participants learning the association between cues and contexts. Critically, in the SSSM, the probabilities of cues (control points) given contexts (force fields) are therefore learned online.

We fit each model to the data from both the opposing- and same-field groups simultaneously, and performed model comparison using the Bayesian information criterion (BIC) (Table 3.2). Although the context-dependent SSMs qualitatively capture the difference in post-washout adaptation between the same- and opposing-field groups (Fig. 3.9), the SSSM fits the overall time course of adaptation and de-adaptation in the

opposing-field group better (Fig. 3.8b,d, solid lines; slow and fast states shown in Fig. 3.10) and has much stronger support (Δ BIC of 112.4 relative to the next-best model).

To understand how the SSSM works, we examined how the estimated probability of each context given a cue evolves for each group. As shown in Fig. 3.11, at the start of the experiment both contexts are activated equally by the right control point (that is, the probability of each context given the right control point is 0.5). Therefore, the model starts naive with respect to how control points relate to contexts. In the same-field group, both contexts are activated equally throughout the experiment, regardless of the control point, leading effectively to a single motor memory. In contrast, in the opposing-field group, each context becomes paired with a different control point, resulting in separate motor memories.

3.4 Discussion

Our results demonstrate that the motor system can form separate memories for different control points on the same object, even when the movements associated with these points are the same. By examining generalization, we found that learning was associated with control points on the object and not with the location of the object in extrinsic space or the location of the control point relative to the hand. Moreover, we found that the motor system only forms distinct memories if the dynamics of the control points differs, allowing efficient and flexible allocation of motor memory. Finally, we developed a normative model of context-dependent motor adaptation that can account for such flexible parcellation of motor memories.

Humans have a remarkable capacity to manipulate objects with different dynamics, and a central focus of research in motor control has been to elucidate how the brain learns and represents these dynamics. To investigate how contextual information can be used to learn different dynamics, numerous studies have used opposing novel

dynamics (e.g., viscous curl fields) that perturb the hand in opposite directions. This work has shown that the ability to learn opposing dynamics, when each is associated with a visual cue, such as the colour of the cursor or background, is generally highly limited (Gandolfo et al., 1996; Howard et al., 2012, 2013). Similarly, when lifting physical objects with different mass distributions, substantial interference is observed when each mass distribution is associated with a different visual geometric cue (Fu and Santello, 2012, 2014). In agreement with these findings, we show that when controlling a single control point, participants cannot learn opposing fields when they are associated with the location of peripheral visual cues that correspond to the targets in the two-control-points group.

Several studies have shown that opposing fields can be learned if the perturbed movement is preceded or followed by a separate movement that differs for each field, or if the perturbed movement is accompanied by a concurrent movement of the other hand that again differs for each field (Howard et al., 2012, 2015; Nozaki et al., 2006; Sheahan et al., 2016; Yokoi et al., 2011). Furthermore, previous studies have shown that opposing fields can be learned if they are associated with different states of the limb, as when the arm operates in a different region of the workspace for each field (Howard et al., 2013; Hwang et al., 2006; Richter et al., 2004; Yeo et al., 2015) or a visual–proprioceptive discrepancy is introduced so that the perceived state of the limb is different for each field (Hirashima and Nozaki, 2012; Howard et al., 2013). Critically, in our experiment, and in contrast with these previous studies, veridical visual feedback of the hand was provided throughout the movement, and the movement start and end points were the same for the opposing fields. Consequently, participants' estimates of the state of the limb did not differ for the two fields. Therefore, we can be confident that it is the control of different points on the object that allows the formation of separate memories for different dynamical contexts. Moreover, we show clearly that the learning of each control point is linked to the location on the object.

Tool use generally involves controlling a specific location, or 'control point', on the tool as it interacts with the environment, such as the face of a hammer striking a nail or the eraser of a pencil rubbing a page. Skilled performance when using tools requires knowledge of the dynamics of the tool (relating applied force to motion) as well as contact mechanics (Casadio et al., 2015; Chib et al., 2009), which may vary across control points. Most natural manipulation tasks involve a series of action phases separated by contact events (Flanagan et al., 2006; Johansson and Flanagan, 2009). Consider, for example, grasping a pencil, lifting it out of a container and transporting it to a piece of paper, using it to erase a pencil mark and then replacing it in the container. Here, contact between the fingertips and pencil marks the end of the reach phase, contact between the pencil and paper marks the end of the transport phase, the breaking of contact between the pencil and paper marks the end of the erase phase, and so on. By comparing predicted and actual sensory signals associated with these contact events, the motor system can monitor task progress, launch corrective actions as needed, and calibrate internal models of objects and their interactions (Flanagan et al., 2006; Johansson and Flanagan, 2009). Critically, both the control point on the object and the dynamics may change from phase to phase. For example, while erasing, the eraser end of the pencil is controlled and the dynamics includes the friction between the eraser and paper. The ability to form separate motor memories for different control points may thus be an essential component of skilled object manipulation and tool use.

Recent studies in neurophysiology, neuroimaging and neuropsychology support the idea that tools are incorporated into a body schema. For example, monkeys have been trained to use a rake to retrieve a pellet of food that could not be reached by hand (Iriki et al., 1996). Following tool use, the visual receptive fields of intraparietal neurons expanded beyond the space surrounding the hand to include the tip of the rake. Similar incorporation of tools is supported by studies in humans making temporal order judgements of tactile stimuli on the hand (Yamamoto and Kitazawa, 2001). Importantly, such incorporation is seen only for the functional part of a tool (Farnè et al., 2005) and a tool must be actively controlled for assimilation to occur (Iriki

et al., 1996; Witt et al., 2005). This suggests that the point being controlled on a tool determines how the body schema changes with tool use. Interestingly, the body schema is thought to consist of distinct neural modules that represent different body parts (Benton, 1959; Haggard and Wolpert, 2005). It is therefore conceivable that different control points are assimilated by different body schema modules. Indeed, one may expect to find intraparietal neurons that respond differently for different control points on a tool. Furthermore, following incorporation of control points into the body schema, the effective 'state' of the body may include the control point, allowing different dynamics to be associated with different states linked to control points.

We found that the formation of distinct motor memories for different control points is not obligatory, but only occurs if the dynamics for each control point is different. Therefore, the allocation of motor memory is both flexible and efficient. To explain this result, we developed a model of motor adaptation based on an SSSM (Ackerson and Fu, 1970; Chang and Athans, 1978; Shumway and Stoffer, 1991), which is a generalization of the mixture-of-experts architecture in neural networks (Jacobs et al., 1991; Wolpert and Kawato, 1998) to dynamical systems (Ghahramani and Hinton, 2000). The SSSM is, unlike these previous models, normative and therefore derived from first principles. Moreover, although the context-dependent SSMs could qualitatively capture the difference in post-washout adaptation between the same- and opposing-field groups, model comparison convincingly selected the SSSM over existing models. SSSMs combine a hidden Markov model with linear dynamical systems and can be used to model systems with multiple discrete operating regimes, termed modes. An SSSM with one mode is an SSM. The key innovation of the SSSM is that only one of the state-space modes is active at any one time and, over time, the system can transition between modes. For example, an SSSM would be appropriate when handling different objects. The transition between the use of different objects can be considered as a Markov process specifying the probability of switching from one object to another (e.g., from kettle to cup) and the dynamics of each object can be different and change over time (e.g., as the kettle empties and the cup fills).

Within this framework, the problem for the sensorimotor system in our experiment is twofold. First, it must infer the current mode (e.g., the context) and the state of the dynamics (for example, the force-field perturbation) based on the parameters of the SSSM. The expected perturbation is then a weighted sum of the perturbations expected for each mode, with weights given by the probability of each mode. Second, it must update the parameters of the SSSM online to ensure that these inferences are as accurate as possible. Here, we provide evidence to support the idea that motor adaptation involves not only state inference but also parameter learning or system identification (Baddeley et al., 2003; Burge et al., 2008; Huang and Shadmehr, 2009; Zarahn et al., 2008).

For simplicity, the SSSM we used had two modes. However, in general, SSSMs can have any number of modes. Central to the performance of the model here is its capacity to learn through experience (that is, during the experiment) the associations between cues (e.g., control points) and contexts (e.g., force fields) via online parameter estimation. To reduce the degrees of freedom of the model, we restricted the simulation to only update the parameters that determine the probabilities of the cues given the contexts. However, in general, all parameters of the model can be updated online. The fits of the SSSM suggest that participants only develop distinct memories for each control point if the dynamics are different (Fig. 3.11). In non-parametric extensions of the standard SSSM, the number of modes is learned through experience, rather than fixed a priori (Fox et al., 2009). Such a model would allow the motor system to be efficient by recruiting additional context-dependent memories (that is, modes) when control points have different dynamics and, conversely, removing modes that are redundant when control points have the same dynamics.

Our results suggest that objects are not represented by the motor system as holistic entities, but are instead parcellated in a task-dependent manner according to control points. Interestingly, the results of the SSSM suggest that effective contextual cues, such as control points, do not necessarily engage distinct motor memories a priori.

Instead, they engage online parameter-learning processes that form distinct motor memories only if and when needed.

Chapter 4

A nonparametric switching state-space model of sensorimotor learning

4.1 Introduction

In Chapter 3 we proposed a switching state-space model to explain dynamic motor learning during tool use. In the model we assumed that the number of contexts and cues are known in advance and fixed. Hence, the number of memories (one per context) do not change over time. However, in the real world, a learner does not know, *a priori*, how many contexts or cues exist. Therefore, a key element missing from current models (including the one proposed in Chapter 3) is an understanding of when sensorimotor experience leads to the formation of a new memory as opposed to the refinement of an existing memory. Current frameworks primarily deal with adaptation of existing memories, with mechanistic models designed to explain overall trial-to-trial learning behavior (e.g., Smith et al., 2006; Thoroughman and Shadmehr, 2000). The few models that have examined multiple memories have done so, as in the previous chapter,

by hard-coding the number of memories and cues as well as how much each memory is expressed and updated on each trial (Lee and Schweighofer, 2009; Nozaki and Scott, 2009).

Here we develop a novel normative model of sensorimotor learning, which departs substantially from previous models of sensorimotor learning. To do this we extend the switching state-space framework to a nonparametric version—the nonparametric switching state-space model (NSSSM). In the parametric version of the SSSM, the complexity of the model is fixed by specifying the number of contexts (i.e., tasks). Specifying too many contexts can lead to overfitting of the data, whereas specifying too few contexts can lead to interference. Nonparametric models have recently been used in the machine learning field to allow the complexity of the model to adapt to the data (Gershman and Blei, 2012). Despite their name, nonparametric models are not models without parameters. Rather, they are infinite-capacity models defined by a finite number of hyperparameters. The hyperparameters control factors such as the willingness to increase the number of contexts based on the data.

In this chapter we will introduce some of the key features of the model. We will then show that the NSSSM can produce spontaneous recovery, which is typically considered a hallmark of the dual-rate model (Smith et al., 2006). According to the NSSSM, spontaneous recovery reflects the re-expression of a memory acquired in a previous context due to Bayesian context estimation, and there is no need to posit multiple processes with different rates. For ease of reading, in this chapter we give an overview of the features that are critical to understanding how the model works. We relegate the extensive details (which are uncontroversial in the machine learning field) to Appendix C.

4.2 Model

The relation between motor outputs and sensory consequences (e.g., dynamics) can change abruptly based on the sensorimotor context (e.g., as we grasp a glass, walk onto a slippery surface, or let go of a rehabilitation robotic device) or gradually as the state changes within a context (e.g., as we drink from the glass or rotate it in our grasp). In the NSSSM, we assume that the world can be in one of a potentially large (or even infinite) set of contexts (Fig 4.1). Each context is associated with a perturbation to the motor-to-sensory mapping (e.g., based on the weight of an object or the level of fatigue) (Ackerson and Fu, 1970; Chang and Athans, 1978; Shumway and Stoffer, 1991). Each perturbation evolves over time according to its own state transition properties (drift, decay and process noise). The context stochastically determines sensory observations in the form of state feedback that arises during movement (e.g., the perturbed state of the limb) and discrete contextual cues (e.g., the visual geometry of an object). Across time, the environment can switch stochastically between contexts.

Critically, the number of contexts is not fixed but can increase over time (this is the meaning of nonparametric in the model title). Therefore, at each point in time, the environment can either transition to a previously experienced context (including the current context) or to a new context. Similarly, at each point in time, the environment can either emit a previously observed cue or emit a new cue.

Specifying the model requires placing priors over the parameters that govern the context transitions, the cue emissions, the state update and state observations. The context transition matrix is Markovian with elements π_{ij} representing the probability of transitioning from context i to context j and includes transitions to previously experienced contexts as well as to new contexts yet to be experienced. We use a prior for this matrix based on the hierarchical Dirichlet Process (or Chinese restaurant franchise) (Teh et al., 2006). This prior is specified by 3 hyperparameters: one (γ)

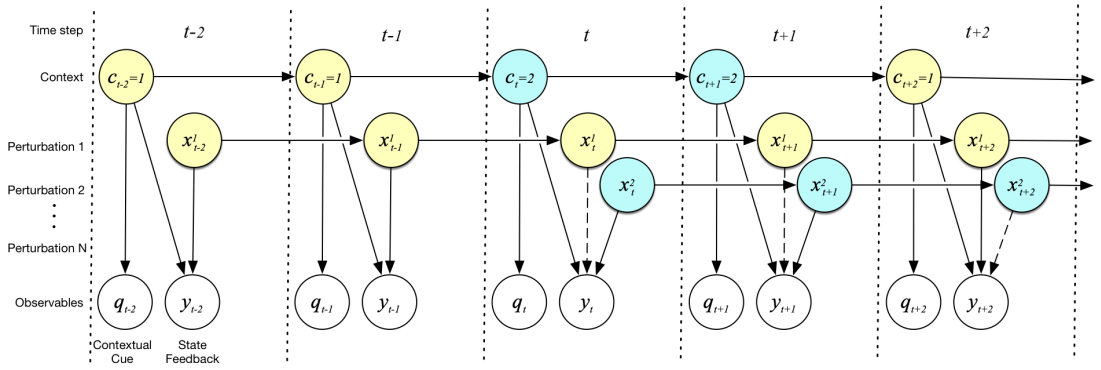


Figure 4.1. Graphical model of the NSSSM. The environment can transition between an unknown, potentially infinite number of contexts. Only two contexts are shown here for clarity. Each context is associated with a perturbation that evolves independent from all other perturbations. The current context influences which cue is emitted and which perturbation contributes to the observed state feedback (solid vs. dashed arrows). In this example, the environment transitions between context 1 (yellow) and context 2 (blue), with context 2 first experienced at time t.

governs how the number of experienced contexts is expected to grow with time, another (α) determines how similar rows of the transition matrix are (i.e., how strongly the probability of transitioning to a particular context depends on the current context) and the other (κ) imposes a self-transition bias (i.e., makes self-transitions more probable than transitions between contexts).

The cue emission matrix has elements ϕ_{jk} specifying the probability of experiencing cue k in context j. The prior for this matrix is also based on a hierarchical Dirichlet Process, but there is no notion of a self-transition. Therefore, the prior is specified by 2 hyperparameters: one (γ') governs the expected number of cues associated with each context (as a function of the number of times the context has been visited) and the other (α') determines how similar the rows of the cue emission matrix are (i.e., how strongly the probability of a particular cue depends on the current context).

Each context (j) is associated with a perturbation (x_t^j) that evolves independent from all other perturbations as a linear dynamical system:

$$x_t^j = a_j x_{t-1}^j + d_j + w_t^j \quad w_t^j \sim \mathcal{N}(0, s_j), \tag{4.1}$$

where a_j is the state transition coefficient, d_j is the state drift and s_j is the variance of the process noise, each associated with context j. The perturbation associated with the current context (c_t) gives rise to state feedback:

$$y_t = x_t^{c_t} + v_t \quad v_t \sim \mathcal{N}(0, r_{c_t}), \tag{4.2}$$

where r_{c_t} is the variance of the measurement noise associated with context c_t.

In the most general form of the NSSSM, all parameters of the state update and observation equations are learned online. In this setting, we place a normal-inverse-gamma prior over the parameters (a_j, d_j, s_k) of equation 4.1 and an inverse-gamma prior over the parameter (r_j) of equation 4.2. These are the conjugate priors to the normal distribution with unknown mean and/or variance that facilitate exact computation of the posteriors. The hyperparameters of these distributions (4 for equation 4.1 and 2 for equation 4.2) govern the mean and variance of the parameters.

Within this framework, the problem for the sensorimotor system is to use the sequence of sensory cues and state feedback to infer i) the number of contexts that have been experienced, ii) the context transition and cue emission matrices, iii) the parameters of the state update and observation equations for each perturbation and iv) the current context and the states of all perturbations. Therefore, unlike standard state-space models, the parameters governing the behavior of the learner change with experience (i.e. within an experiment).

This is a hard inference problem as the states and parameters are coupled. We perform online state inference and parameter learning to simulate the learning process and fit

the model to experimental data. To this end, we use a Monte Carlo simulation strategy based on a resample-sampling framework known as particle learning (PL) (Carvalho et al., 2010). See Appendix C for full details of the model.

4.3 Results

4.3.1 Spontaneous Recovery

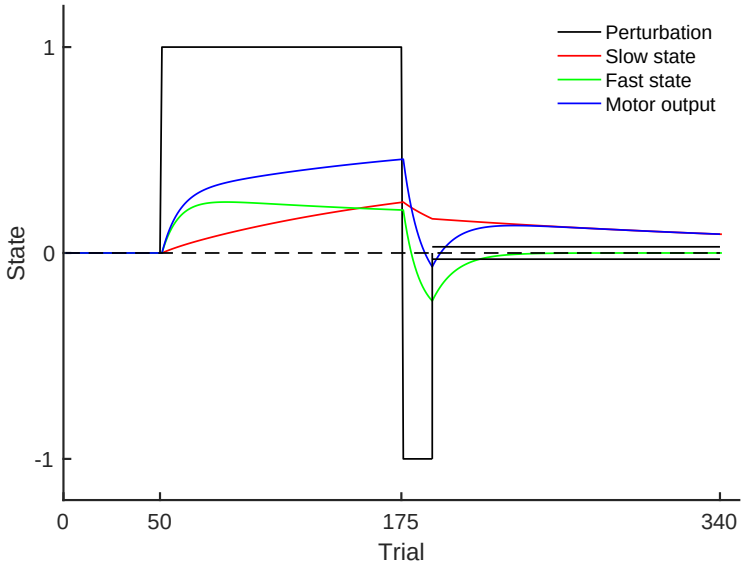


Figure 4.2. Spontaneous recovery can arise from a dual-rate model. The perturbation schedule used to elicit spontaneous recovery (black). After a series of null-field trials (P^0), the participant is exposed to a series of P^+ trials followed by brief exposure to P^-. This is followed by a sequence of channel trials (parallel horizontal lines). The fast (green) and slow (red) processes of the dual-rate model sum to produce the motor output (blue). The parameters of the dual-rate model ($a_\mathrm{f} = 0.92$, $a_\mathrm{s} = 0.996$, $b_\mathrm{f} = 0.03$, $b_\mathrm{s} = 0.004$) are taken from Smith et al. (2006).

Conventional wisdom is that motor adaptation consists of two (or more) parallel processes operating on different timescales (slow and fast learners), with spontaneous recovery taken as the most compelling piece of evidence. Spontaneous recovery is seen after adaptation to a positive perturbation (which we term P^+, e.g., a clockwise viscous

curl field) followed by a brief period of negative perturbation (which we term P^-, e.g., a counter-clockwise viscous curl field) that brings performance back to baseline. When presented with a long series of channel trials, spontaneous recovery, or rebound, is observed (Keisler and Shadmehr, 2010; Smith et al., 2006). According to multi-rate models, this rebound results from the interaction of the slow and fast processes decaying at different rates in the absence of error (Fig 4.2). The initial adaptation to P^+ is achieved by both the slow and fast processes contributing to learning. The rapid de-adaptation in P^- is achieved by the fast process becoming negative (green curve) while the slow process remains positive (red curve), thereby leading to no apparent adaptation (blue curve). During the channel trials the fast process rapidly decays to zero revealing the slow process, which is still partially adapted to the initial perturbation.

The NSSSM provides a contrasting interpretation of spontaneous recovery in terms of memory re-expression due to Bayesian context estimation. Fig 4.3 shows the NSSSM fit to a typical spontaneous recovery paradigm (Fig 4.3a) and data set (Fig 4.3b). The NSSSM treats P^0, P^+ and P^- as separate contexts. Over the course of the experiment, the number of inferred contexts increases (Fig 4.3c). The model starts with one memory instantiated for P^0 (Fig 4.3e). Then on the introduction of P^+ and P^-, a second and then memory are instantiated (Fig 4.3f-g). Importantly, the addition of new memories is an automatic Bayesian process based on whether or not the model can explain the observed data. The memories are instantiated close to the correct value for the perturbations (Fig 4.3e-g) and tend to decay if the force fields are not presented. When a force field is not presented, the associated memory becomes more uncertain (Fig 4.3e-g note spreading of colors).

To explain spontaneous recovery in the NSSSM, we first need to explain how the motor output depends on the estimated context. In the NSSSM, the motor output is a weighted sum of the perturbations expected in each context, with weights given by the prior probability of each context (i.e., the probability of each context prior to moving).

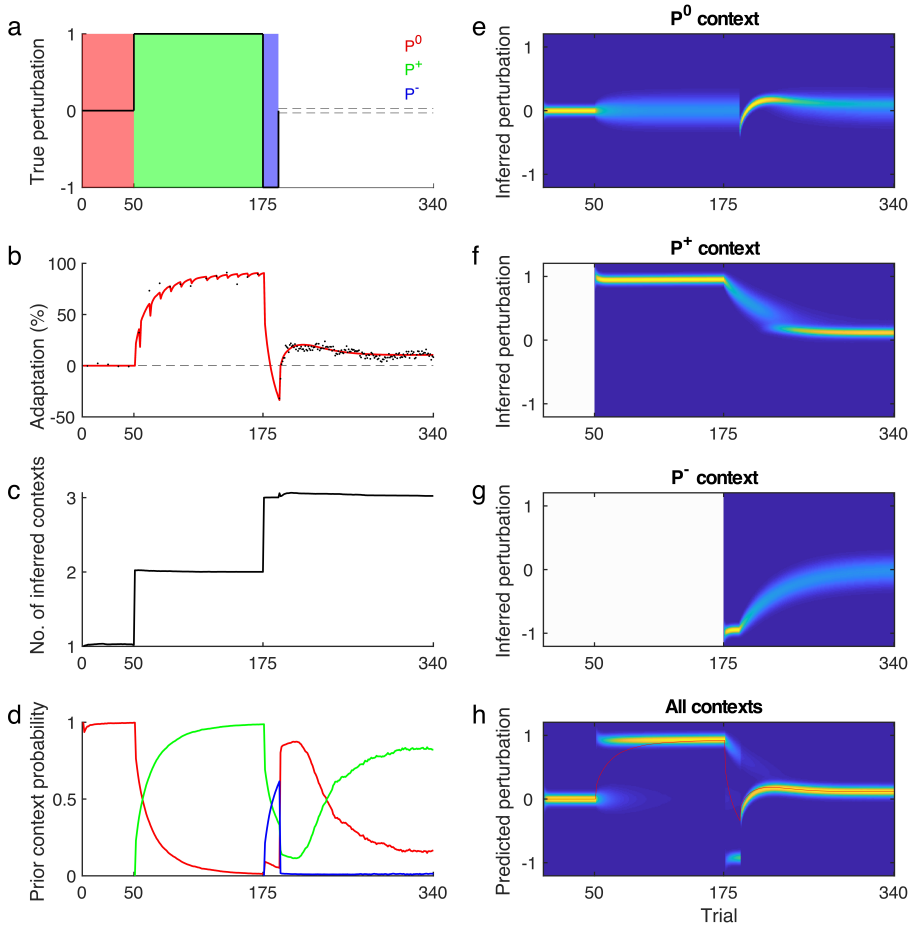


Figure 4.3. Spontaneous recovery in the NSSSM. a, The perturbation schedule used to elicit spontaneous recovery. Occasional channel trials are presented during the P^0 and P^+ phases to assess learning. **b**, Adaptation to P^+ measured on channel trials shows learning of P^+ and de-adaptation to baseline after P^-. In the series of channel trials there is re-expression of positive adaptation termed spontaneous recovery. The data shown are from a typical participant (circles). The fit of the NSSSM is shown in red. **c**, The average number of instantiated contexts across particles show that a new context is inferred following the introduction of P^+ and P^-. **d**, The prior probability of context P^0 (red), P^+ (green) and P^- (blue). **e-g**, The three states are shown in the order they are instantiated (grey prior to instantiation) and are labelled by the context that led to their instantiation. The colormap indicates the probability of the state of each perturbation. **h**, The predicted perturbation is a weighted sum of the individual perturbation estimates (**e-g**), where the weights are given by the prior probabilities (**d**). The motor output (red) is the mean of the predicted perturbation. Note that in **c-h**, we simulated the NSSSM without channel trials during the exposure phase for clarity.

The prior probability of each context depends on three components. First, it depends on how common each context is, which is estimated by inferring the context counts (i.e., the number of times each context has occurred). Second, it depends how common each context is given the previous context. This can estimated by inferring the context transition counts (i.e., the number of times each context has followed each previous context). Third, it depends on a self-transition bias, which determines the tendency of the previous context (regardless of its identity) to transition to itself. Because the context counts and context transition counts are estimated and updated online, the prior probability of each context varies across trials (Fig 4.3d).

One additional aspect of the model behaviour that is key to its behavior is what happens on channel trials (one-dimensional spring-like force field that constrains the hand to a straight-line path to the target). The classical view of channel trials in a state-space model is that there is no learning as the channel leads to no performance error. In the absence of errors, states simply decay over repeated trials. In contrast, in the NSSSM, perfect performance in the channel trial arises as the motor output exactly matches the perturbation. This is equivalent to observing a perturbation that is the opposite of the motor output, which we will term a fictitious perturbation (in that it is self generated and would be absent if the participant chose to move as in a null field). Because the motor output is a weighted sum of the perturbations expected in each context, the fictitious perturbation will be different from at least one perturbation estimate (and in general from all perturbation estimates), unless all perturbation estimates are identical. Therefore, in the NSSSM, there will be learning on channel trials. Hence, we use the term 'channel trial' rather than 'error-clamp' trial. This alternative perspective predicts that in channel trials there will be active unlearning (in addition to decay), whereas with the passage of time alone there will only be decay. Consistent with this prediction, de-adaptation is faster when movements are performed in channel trials compared with the passage of time alone (Criscimagna-Hemminger and Shadmehr, 2008; Ingram et al., 2013; Kitago et al., 2013).

Having given some intuitions about the behaviour of the NSSSM, we now return to the simulations. Although the introduction of P^+ leads to a new state that is instantiated close to P^+, the motor output only slowly transitions to P^+ (Fig 4.3b trials from 50 onwards). This is because the weight for P^0 is initially greater than the weight for P^+ (Fig 4.3d), as P^0 has been experienced more often. As P^+ is repeatedly encountered, the weight becomes greater for P^+, leading to a gradual expression of adaptation to the force field. When P^- is introduced abruptly, another state is instantiated close to P^-. Again, as P^- is repeatedly encountered, the weight for P^- increases. The motor output is close to zero at the end of the counter-exposure phase as the weights for P^+ and P^- are approximately equal.

Once the channel trial phase starts, the prior probabilities of contexts P^0, P^+ and P^- and their associated state estimates determine the motor output. On the first trial of the channel phase, the motor output is slightly negative, inducing a fictitious observation that most closely resembles the P^0 perturbation. This causes the memory of P^0 to become negative (Fig 4.3e) and the prior probability of P^0 on the second trial of the channel phase to increase (Fig 4.3d). The prior probability of P^- drops precipitously (Fig 4.3d), as this context has been encountered the least throughout the experiment. This precipitous drop means that the memory for P^- contributes less to the motor output on the second trial of the channel phase, and hence the second fictitious observation is less negative than the first. Importantly, the memory for P^0 tracks this fictitious observation and so becomes less negative itself (Fig 4.3e). This in turn causes the motor output (and fictitious observation) on the third trial of the channel phase to become even less negative. And so the cycle continues. Eventually, the motor output becomes positive due to the contribution of the memory of P^+. This produces spontaneous recovery, which eventually decays due to decay of the memories of P^0 and P^+. In Chapter 5 we will test this interpretation of spontaneous recovery directly.

4.4 Discussion

The NSSSM improves upon the switching state-space model (SSSM) described in Chapter 3 in several respects. First, we estimate the number of contexts and cues online and instantiate new memories in an automatic, Bayesian manner. Second, we learn the context transition probabilities (as well as the cue emission probabilities). Third, we extend parameter learning to the state-space parameters that govern how perturbations evolve over time (i.e., state transition coefficient, drift, process noise) and how perturbations give rise to state feedback (i.e., measurement noise). Fourth, we introduce the concept of fictitious observations, which leads to learning in channel trials.

The dual-rate model assumes that there are multiple states with different learning and retention parameters. Although the NSSSM also posits multiple states (one for each context), each state is instantiated with the same parameters (when the parameters are assumed to be known and shared across contexts) or with the same prior belief over parameters (when the parameters are assumed to be unknown and need to be learned online). Depending on the values of hyperparameters, states and parameters can be learned on different timescales in the NSSSM. We saw an example of this in the simulations of spontaneous recovery, where the state was learned quickly (each state estimate was instantiated close to the true perturbation), and the context transition distribution was learned slowly (it took many trials for the learner to believe with high confidence that the P^+ context would persist). Thus, in the NSSSM, learning on multiple timescales is achieved not by endowing states with different parameters but by performing combined state and parameter estimation.

The NSSSM can be contrasted against two related models in the literature. The first is an extension of the dual-rate model to environments with multiple contexts (Lee and Schweighofer, 2009; Pekny et al., 2011). This model consists of a single

context-independent fast process and multiple context-dependent slow processes, the number of which is specified in advance. The model assumes that the first trial of a perturbation engages a slow process specific to that context. This slow process—and no other slow process—is then expressed and updated on each subsequent trial. This implies that the learner is certain that the perturbation experienced on one trial will persist to the next trial. In contrast, in the NSSSM, the learner initially thinks that P^0 is more probable and only gradually comes to expect that P^+ will persist. The models also behave differently on channel trials. For example, in the multiple-context state-space model, each slow process contributes an equal amount to the motor output (Pekny et al., 2011), whereas in the NSSSM, the amount that each memory contributes to the motor output depends on prior experience and varies across trials.

The second model worthy of comparison is a switching Kalman filter model (Oh and Schweighofer, 2018). This model consists of two Kalman filters—a baseline filter and a novice filter that can develop into an expert perturbation filter. In contrast to the NSSSM, which can infer new contexts, the number of filters in the switching Kalman filter model is specified in advance, and there is no mechanism for adding new filters online. The switching Kalman filter model places a prior probability on each filter. However, these probabilities are fixed and they do not depend on the previous context. In contrast, in the NSSSM, the probability of each context is learned from experience and depends on the learned context transition frequencies. The switching Kalman filter model also uses a 'winner-take-all' strategy to decide which filter to engage and train on each trial. In contrast, in the NSSSM, all memories are engaged and trained on each trial by an amount proportional to the probability of their respective contexts. Thus, the NSSSM represents uncertainty about the current context, whereas the switching Kalman filter model does not.

The NSSSM infers the number of contexts as well as the frequencies of contexts and the statistics of contexts transitions. Therefore, even in the absence of contextual cues, the NSSSM can construct rich models of the environment. There is some

experimental evidence to suggest that humans also do this. For example, people can adapt to forces acting on the hand when they depend on the order of a movement in a predefined sequence (Wainscott et al., 2005), and people can learn the transitions between visuomotor transformations that switch mid-movement to improve online control (Vetter and Wolpert, 2000).

In the next two chapters we test novel predictions of the model related to spontaneous recovery and Bayesian context estimation.

Chapter 5

Spontaneous recovery as Bayesian context estimation under uncertainty

5.1 Introduction

In the previous chapter we proposed a nonparametric switching state-space model of dynamic motor learning. This model was able to produce spontaneous recovery, which is typically accounted for with a dual-rate model in the motor learning literature (Keisler and Shadmehr, 2010; Smith et al., 2006; Trewartha et al., 2014). The mechanism by which spontaneous recovery occurs in the NSSSM is very different from in the dual-rate model. In the dual-rate model, spontaneous recovery arises from the interaction of a fast process and a slow process that adapt and forget on different timescales. In the NSSSM, spontaneous recovery arises from the re-expression of a memory acquired in a previous context. In this chapter, we perform an experiment to test the NSSSM interpretation of spontaneous recovery.

We wanted to modify the standard spontaneous recovery paradigm as little as possible to distinguish the NSSSM from the dual-rate model. Simulations suggested that this

could be achieved by replacing the 3rd and 4th trials in the channel trial phase with P^+ trials. The NSSSM predicts that these trials should cause participants to believe with high confidence that the P^+ context has returned. This in turn should lead to more rapid and complete re-expression of the memory of P^+. We refer to this rapid re-expression as 'evoked recovery', to emphasise its dependence on re-exposure to the force field. Our experimental findings confirmed this prediction. Moreover, model comparison revealed that the NSSSM is superior to the dual-rate model at describing both spontaneous recovery and evoked recovery.

5.2 Methods

5.2.1 Participants

A total of 16 neurologically intact participants (8 males and 8 females; age 29.7 ± 6.8, mean ± s.d.) were recruited to participate in the experiment, which had been approved by the Columbia University IRB. All participants provided written informed consent and were right-handed according to the Edinburgh handedness inventory (Oldfield, 1971).

5.2.2 Apparatus

All experiments were performed using a vBOT planar robotic manipulandum with a virtual-reality system and air table, which we have described in Chapters 2 and 3.

5.2.3 Spontaneous recovery condition

Participants (n = 8) in the spontaneous recovery condition performed the standard spontaneous recovery paradigm. The paradigm consisted of pre-exposure, exposure, counter-exposure and channel phases, as detailed below.

The paradigm consisted of a pre-exposure phase (5 blocks/50 trials), an exposure phase (12 blocks/120 trials), a counter-exposure phase (15 trials) and a channel phase (150 trials) (Fig. 5.1a). In the pre-exposure and exposure phases, participants performed blocks of ten trials in which one of the trials (not the first) was randomly chosen as a channel trial and the remainder were either null-field trials (pre-exposure) or force-field trials (exposure). The counter-exposure phase consisted only of force-field trials and the channel phase consisted only of channel trials. A 45 s rest break was given after block 6 of the exposure phase. A block of 5 force-field trials (with no channel trials) followed this rest break to mitigate the effects of any time-dependent decay of motor memory.

The direction of the force field (CW or CCW) in the exposure phase was counterbalanced across participants. The direction of the force field in the counter-exposure phase was opposite to that of the exposure phase. Throughout this chapter, we refer to the null field, the exposure field and the counter-exposure field as P^0, P^+ and P^-, respectively.

5.2.4 Evoked recovery condition

Participants (n = 8) in the evoked recovery condition performed a modified version of the spontaneous recovery paradigm. The modified version was identical to the original paradigm except that trials 3 and 4 of the channel phase were replaced with P^+ trials (Fig. 5.1b). We chose to replace trials 3 and 4 of the channel phase (as opposed to trials

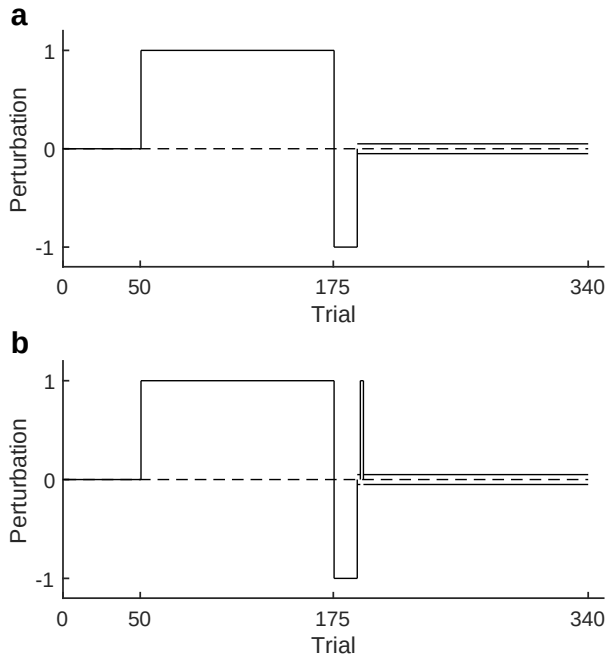


Figure 5.1. Experimental paradigm. a, The perturbation schedule used to elicit spontaneous recovery. After a series of P^0 trials the participant is exposed to series of P^+ trials followed by brief counter-exposure to P^-. This is followed by a sequence of channel trials (parallel horizontal lines). **b**, In the evoked recovery condition, the 3$^{\text{rd}}$ and 4$^{\text{th}}$ trials of the channel phase are replaced with P^+ trials.

1 and 2) so that adaptation could be measured before and after the P^+ intervention trials.

5.2.5 Data analysis

Data analysis was performed using MATLAB R2018b. Adaptation was measured on channel trials as the percentage of the force field that was compensated for, which have described in Chapter 2.

To identify changes in adaptation between conditions, a mixed-design ANOVA was performed with significance set to $P < 0.05$.

5.2.6 Model fitting

We fit several candidate models to the data of individual participants. These models include the NSSSM, the dual-rate model (Smith et al., 2006) and a 3-rate model (Inoue et al., 2014). The 3-rate model is a generalisation of the dual-rate model to 3 processes. We do not apply inequality constraints when fitting the 3-rate model so that the model can perform to its fullest potential. Henceforth, we refer to the dual-rate model as the 2-rate model, to emphasise that it is a member of a wider n-rate model class that also includes the 3-rate model.

We fit each model separately for each participant using the particular sequence of perturbations they experienced. Under the assumption that the model errors are independent and identically distributed according to a normal distribution, we obtained maximum likelihood estimates of the free parameters/hyperparameters of the models by minimising the mean squared error between model output and adaptation measured on channel trials.

To fit the NSSSM, we used Bayesian adaptive direct search (BADS; Acerbi and Ma, 2017), a Bayesian optimization algorithm that is suitable when no gradient information is available and the objective function is stochastic. BADS alternates between a series of fast, local Bayesian optimization steps and a systematic, slower exploration of a mesh grid. To ensure that our results did not arise from using different optimisers for different models, we fit each model with both BADS and fmincon (from 20 different initial parameters settings) and we report the best solution found by either optimiser.

Model comparison

The aim of model comparison is often to select a single model that best explains the data collected from multiple participants. This 'best model' approach assumes that all the participants' data are generated by the same model. An alternative approach is to

assume that each participant's data is generated from a model that is sampled from an unknown multinomial distribution over a finite set of candidate models (Rigoux et al., 2014; Stephan et al., 2009). Here, the aim of model comparison is to i) infer the posterior probability, for each participant, of belonging to each model; and ii) infer the posterior distribution over the unknown multinomial distribution. Importantly, this hierarchical approach is more robust to the presence of outlier participants when the population is heterogeneous. Furthermore, the set of candidate models can be partitioned into disjoint subsets known as families and inference can be performed at the level of families of models rather than at the level of individual models (Penny et al., 2010; Stephan et al., 2009). In so-called 'family-level inference', the aim of model comparison is to i) infer the posterior probability, for each participant, of belonging to each family; and ii) infer the posterior distribution over the unknown multinomial distribution over families of models. Importantly, family-level inference allows one to compare characteristics of interest that are shared across models while still leveraging the statistical strength of all models.

We performed family-level inference by considering two families of models: an NSSSM family and a multi-rate SSM family. For the NSSSM family, we defined a full model and three nested models, which are special cases of the full model that lack certain attributes or with a subset of parameters set to special values. The full model has a non-zero self-transition bias and learns the drift parameter online. This requires specifying two free parameters: the self-transition bias and the precision of the Gaussian prior on the drift parameter (the mean is set to zero). The three nested models set the self-transition bias and/or the drift parameter to zero. Collectively, the full and nested models form the NSSSM family. The two-rate and three-rate SSM models form the multi-rate SSM family. Importantly, although these families contain different numbers of models (4 vs. 2), family-level inference is unbiased as a uniform prior is placed over the family probabilities; this is achieved by placing smaller prior probability on individual models that belong to larger families (Penny et al., 2010).

We performed family-level inference using the Variational Bayesian Analysis (VBA) toolbox, which is available online at http://mbb-team.github.io/VBA-toolbox. This inference procedure uses variational Bayes to estimate the parameters of a Dirichlet distribution over the probabilities of all candidate models (for details see Stephan et al., 2009). The procedure requires only the log model evidences for each model and each participant. The log model evidence for each model and each participant can be approximated as half the negative of either the Akaike information criterion (AIC; Akaike, 1974) or the Bayesian information criterion (BIC; Schwarz et al., 1978) (Murphy, 2012). Under the assumption that the model errors are independent and identically distributed according to a normal distribution, the AIC is (up to an additive constant that does not depend on the model)

$$\mathrm{AIC} = n \cdot \ln[(1 - R^2)/n] + 2 \cdot k, \tag{5.1}$$

where n (167 for the spontaneous recovery condition and 165 for the evoked recovery condition) is the number of data points and k (4 for the 2-rate model, 6 for the 3-rate model and 4-6 for the variants of the NSSSM) is the number of free parameters. The first term in the AIC rewards goodness of fit, and the second term penalizes model complexity, as measured by the number of free parameters in the model. Under the same Gaussian i.i.d assumption, the BIC is (up to an additive constant)

$$\mathrm{BIC} = n \cdot \ln[(1 - R^2)/n] + k \cdot \ln(n). \tag{5.2}$$

The AIC and the BIC differ only in the size of the penalty term—the AIC has a fixed penalty that does not scale with the number of data points. When $n > e^2 \approx 7$, the BIC penalises model complexity more heavily than the AIC. Thus, the AIC tends to choose models that are too complex, and the BIC tends to choose models that are too simple (Hastie et al., 2005). For model selection purposes, there is no clear choice between the

AIC and the BIC. Therefore, following Penny et al. (2004), we perform family-level inference using both AIC and BIC (separately) and draw conclusions regarding the superiority of one particular family only when both metrics agree, that is, we choose the most conservative of the two metrics.

We compare families of models at both the level of individual participants and at the group level. At the level of individual participants, we report the posterior probability, for each participant, of belonging to each family. At the group level, we report the exceedance probability, which is the posterior probability that one family (here the NSSSM) is more frequent in the population than the other family (here the multi-rate SSM) (Stephan et al., 2009).

5.3 Results

In this experiment, we investigated whether spontaneous recovery arises from the re-expression of a memory acquired in a previous context. We recruited 16 participants to perform either a spontaneous recovery (n = 8) or an evoked recovery paradigm (n = 8). We begin by describing the spontaneous recovery condition.

Spontaneous recovery

Participants started the experiment in a null field (P^0) and then performed 12 blocks of 10 trials (120 trials) in a force field (P^+) (Fig. 5.1a). In these pre-exposure and exposure phases, one randomly selected trial in each block of 10 trials was a channel trial, in which the movement was constrained to a straight line between the home position and the target (Milner and Franklin, 2005; Scheidt et al., 2000). On these trials, the lateral forces generated by participants were recorded to provide a measure of adaptation (Fig. 5.2 blue). By the end of the exposure phase, participants had

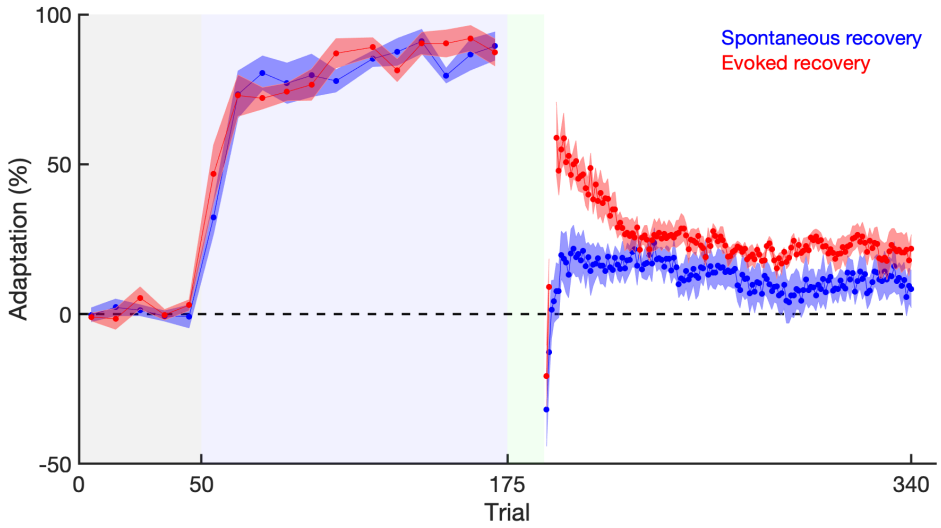


Figure 5.2. The memory of P^+ is not destroyed during counter-adaptation to P^- and can be re-expressed. Adaptation measured on channel trials over the course of the experiment. Trials associated with P^0, P^+ and P^- perturbations are indicated by grey, blue and green backgrounds, respectively. The channel phase is indicated by a white background. Data are plotted as mean ± s.e.m. for the spontaneous recovery group ($n = 8$) and the evoked recovery group ($n = 8$).

learned to compensate for the force field (Fig. 5.2), and then performed 15 trials in the opposite force field (P^-). During this counter-exposure phase, adaptation reduced to slightly below baseline. Participants then performed 150 channel trials. During the first ten trials of this channel phase, adaptation to P^+ re-emerged—demonstrating spontaneous recovery—and then slowly declined.

We fit several candidate models to the data of the individual participants. These models were the 2-rate model, the 3-rate model and 4 variants of the NSSSM (a full model and 3 nested models, see Model comparison). We divided these 6 models into 2 families (a multi-rate SSM family and an NSSSM family) and performed Bayesian model comparison of these families. Note that rather than assume that all participants' data was generated by the same model, we treat the best model as a random effect that can vary across participants. This approach is considerably more robust than conventional fixed-effects analyses (e.g., based on the group Bayes Factor), especially in the presence of outliers (Stephan et al., 2009).

As expected, both the NSSSM (Fig. 5.3a solid blue line) and the 2-rate model (Fig. 5.3b solid blue line) were able to qualitatively capture the time course of adaptation throughout the experiment, although the NSSSM provided a better fit than the 2-rate model to the initial adaptation to P^+ and the initial rebound. Bayesian model comparison showed that the NSSSM family was superior to the multi-rate family. For all eight participants, the posterior probability of the NSSSM family was greater than 0.9995 based on AIC and greater than 0.9999 based on BIC.

Evoked recovery

In an attempt to induce more complete recovery of the memory of P^+, we performed a separate evoked recovery condition (n = 8), where we replaced the 3^{rd} and 4^{th} trials of the channel phase with P^+ trials (Fig. 5.1b). An ideal observer should predict that the 5^{th} trial of the channel phase is highly likely to be a P^+ trial, as nearly all of the P^+ trials presented earlier in the experiment were followed by a P^+ trial (only the last trial of the exposure phase was followed by a P^- trial). Therefore, according to the NSSSM, the P^+ trials presented on the 3^{rd} and 4^{th} trials of the channel phase should increase the prior probability of P^+ on the 5^{th} of the channel phase (as well as on the trials beyond), leading to greater recovery of the memory of P^+.

As expected, the evoked recovery group adapted and de-adapted similarly to the spontaneous recovery group during the exposure and counter-exposure phases (Fig. 5.2 red). There was no significant difference in adaptation between the groups in the final two blocks of the exposure phase or in the first 2 trials of the channel phase (mixed-design ANOVA, main effect: $F_{1,14} = 1.37, P = 0.261$, interaction effect: $F_{1,14} = 0.94, P = 0.348$). As predicted by the NSSSM, the two P^+ trials in the channel phase induced rapid recovery of the memory of P^+, as demonstrated by an abrupt step in adaptation towards P^+. Following this abrupt step, adaptation decayed over $\sim$30 trials to a non-zero endpoint. Interestingly, the adaptation trace in the evoked

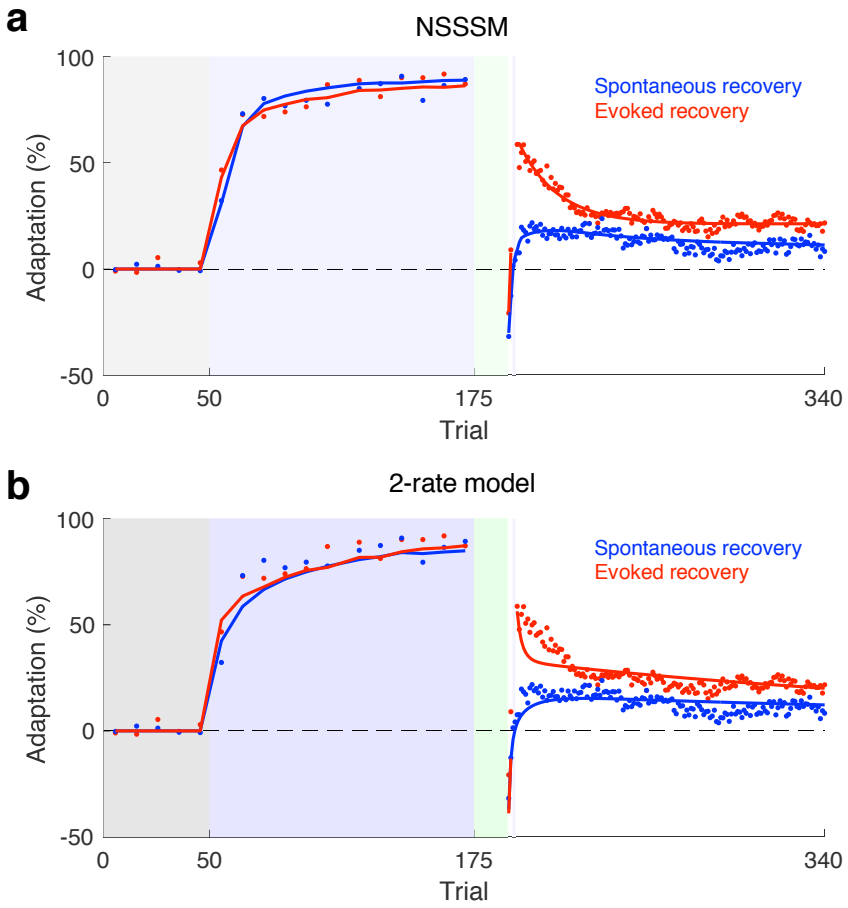
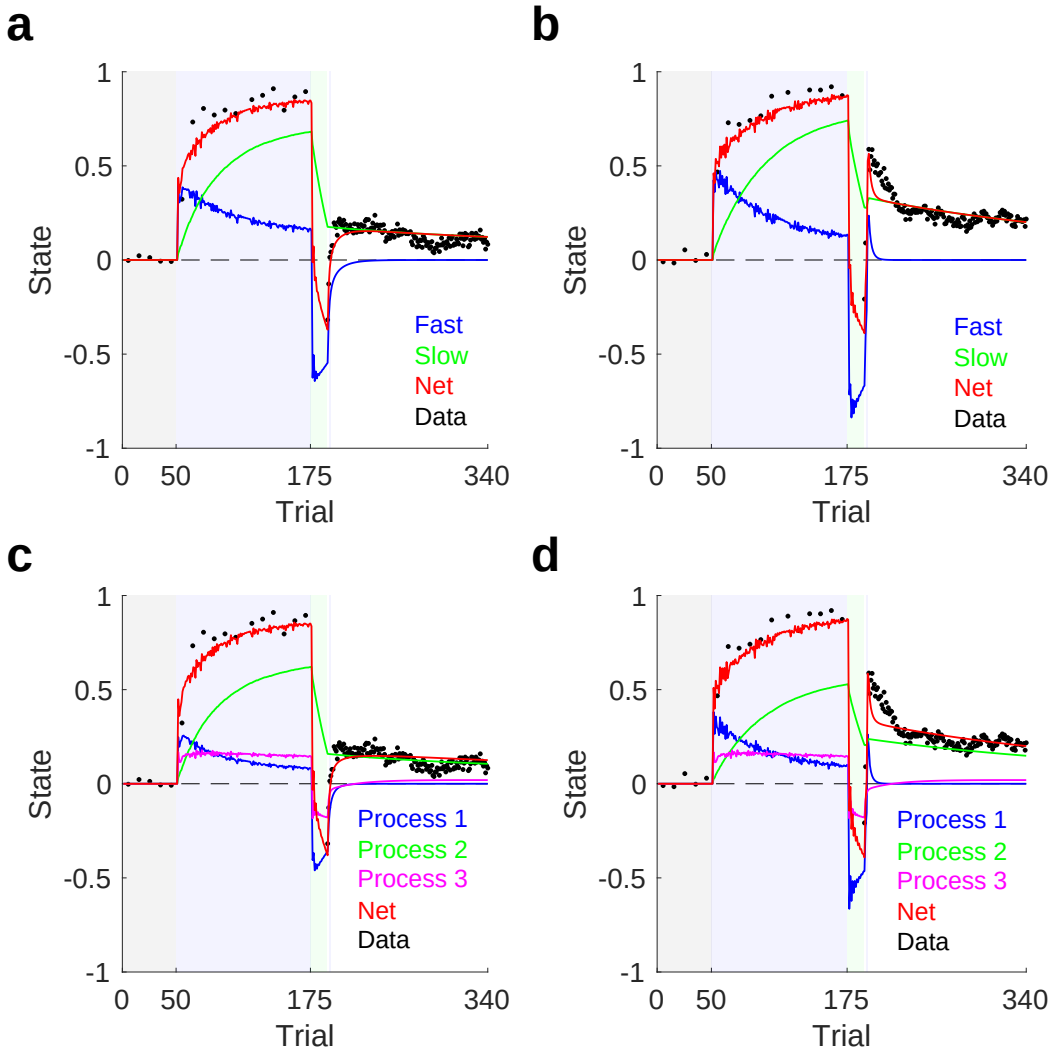


Figure 5.3. Model fits to adaptation data. **a**, The NSSSM model fit to the spontaneous and evoked recovery data sets. Filled data points show the mean adaptation across participants. Solid lines show the average of the fits to individual participants. For each participant, we used the NSSSM variant that has the highest posterior probability. **b**, Same as **a** but for the 2-rate model.

recovery condition (Fig. 5.2) looks remarkably similar to the P^+ memory trace in the spontaneous recovery simulation in Chapter 4 (Fig. 4.3f), suggesting that the previous memory of P^+ may have been re-expressed.

The NSSSM was able to fit the first 2 trials of the channel phase, the abrupt step in adaptation following the two P^+ trials and the time course of subsequent decay (Fig. 5.3a solid red line). In contrast, the 2-rate model provided a poor fit to the first 2 trials of the channel phase and failed to reproduce the time course of decay (Fig. 5.3b solid blue line). Following the two P^+ trials, adaptation in the 2-rate model decayed



[t!]

Figure 5.4. Hidden states of the 2-rate and 3-rate models. **a**, The 2-rate model fit to the spontaneous recovery data set. Data points in black show the mean adaptation across participants. Solid lines show the average of the fits to individual participants. The fast and slow states are shown in blue and green, respectively, and the net motor output is shown in red. **b**, Same as **a** but for the evoked recovery data set. **c,d**, Same as **a** and **b**, respectively, but for the 3-rate model. The hidden states have not been labelled as fast, slow and ultraslow processes, as we did not impose inequality constraints on the parameters when fitting the model to data.

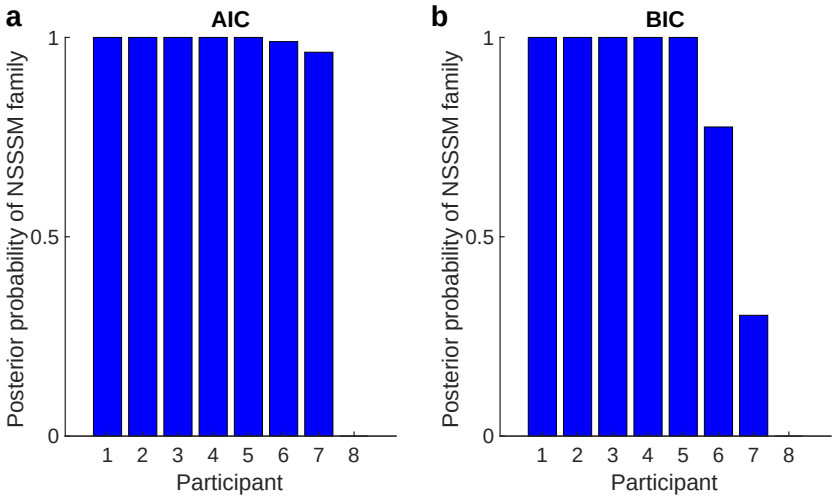


Figure 5.5. Bayesian model comparison in the evoked recovery condition. **a**, The posterior probability that a model in the NSSSM family generated the data of each participant. Model comparison was performed using the Akaike information criterion (AIC) to approximate the log model evidence. **b**, Same as **a** but using the Bayesian information criterion (BIC) to approximate the log model evidence.

rapidly over ∼5 trials before exhibiting a slower decline. The reason for this is clear. In the 2-rate model, the abrupt step in adaptation following the 2 P^+ trials is driven almost entirely by the fast process, which decays quickly (Fig. 5.4b). The addition of a third process was unable to fix this problem (Fig. 5.4d).

Bayesian model comparison revealed that the NSSSM family was more likely than the multi-rate family for the majority of participants (6/8) in the evoked recovery condition (Fig. 5.5). The superiority of the NSSSM was also supported at the group level—the exceedance probability that the NSSSM family was more likely than the multi-rate family (given the data of all participants in both groups) was greater than 0.9999 based on AIC and greater than 0.9995 based on BIC.

Taken together, these experimental and modelling findings provide compelling evidence that the NSSSM is a better model of spontaneous recovery than typical multi-rate state-space models. This suggests that spontaneous recovery is better interpreted as the re-expression of a memory acquired in a previous context due to dynamic Bayesian context estimation.

5.4 Discussion

In the previous chapter we presented a novel interpretation of spontaneous recovery based on the NSSSM. Here we tested a prediction of this interpretation, namely, that it should be possible to re-express the memory of P^+ if it has not been overwritten by P^-. This prediction was borne out. The two P^+ trials in the channel trial phase induced a qualitative switch in behaviour. That is, rather than observing a non-monotonic spontaneous recovery curve, we observed a monotonically decreasing adaptation trace that started from a level close to P^+. Moreover, model comparison revealed that the NSSSM provided a better description of the data than the multi-rate class of state-space models, which includes the dual-rate model that is typically used to model spontaneous recovery. Importantly, the multi-rate class of models showed qualitative mismatches on the evoked recovery data set, as they failed to capture the time course of decay following the two P^+ trials. In contrast, the NSSSM was able to account for both the spontaneous and evoked recovery data sets.

A previous study by Pekny et al. (2011) explored the question of whether the memory of a perturbation is destroyed during counter-adaptation to an opposite perturbation. To address this question, the authors had participants first adapt to P^- and then perform the standard spontaneous recovery paradigm. Compared with a control group that did not adapt to P^- before performing the spontaneous recovery paradigm, the amount of rebound in the channel trial phase was reduced (i.e., adaptation was biased towards P^-). The authors concluded that the memory of P^- had been partially protected from destruction during adaptation to P^+. This between-groups effect provided indirect evidence that two separate context-specific memories had been created and existed at the same time. In the current experiment, we show this definitively by evoking the memory of P^+ in individual participants. That is, we provide direct evidence that—even in the absence of conventional contextual cues—multiple context-specific memories are created, exist simultaneously and can be switched between. Interestingly,

in the same paper, Pekny et al. (2011) also showed that spontaneous recovery can be induced by omitting reward feedback after a participant has adapted and de-adapted to a perturbation. This finding is outside the remit of the NSSSM, which does not include explicit rewards. To explain this finding, the NSSSM would need to be incorporated into a wider reinforcement learning framework.

In visuomotor rotation learning, attempts have been made to map the slow and fast states of the dual-rate model onto implicit and explicit learning systems (McDougle et al., 2015). A recent force-field adaptation study by Keisler and Shadmehr (2010) also claimed to show that the fast process of the dual-rate model can be interrupted by a declarative memory task. In this study, the authors had participants perform a 3-minute declarative memory task in between the counter-exposure phase and the channel trial phase of a spontaneous recovery paradigm. Rather than observing spontaneous recovery in the channel trial phase, the authors observed a monotonically decaying adaptation trace that resembled the slow process of the dual-rate model. Therefore, they concluded that the fast process had been disrupted by the memory task and that this disruption implies that the fast process shares resources with the declarative memory system. Interestingly, the monotonically decaying adaptation trace that they observed looks very similar to the trace that we observe following evoked recovery. Hence, the NSSSM offers an alternative interpretation of their data, namely, that the declarative memory task, like the P^+ trials in the evoked recovery paradigm, altered or disrupted the context estimate. This implies that the declarative memory system is involved in context estimation. In the NSSSM, the context estimate is made up of multiple terms including one that depends on the previous context (how common each context transition is) and one that is independent of the previous context (how common each context is). One could speculate that the declarative memory task erases the memory of the previous context, P^-, causing the context estimate to depend only on the most frequently encountered context, P^+. Indeed, recent work has suggested that the prefrontal cortex can be viewed a recurrent neural network that learns the statistical structure of tasks (Wang et al., 2018), and so it is conceivable that this

brain region plays an important role in learning context transitions in motor tasks and performing Bayesian context estimation.

In the next chapter, we test another prediction of the NSSSM, that the partitioning of learning between memories and the expression of learning are determined by Bayesian context estimation.

Chapter 6

Bayesian context estimation underlies single-trial learning

6.1 Introduction

The NSSSM introduced in Chapter 4 suggests that the belief about the context after the movement should influence which memories are updated in response to errors. It also predicts that the motor output (that is, the expression of learning) should depend on the prior belief about the context before the movement. In this chapter, we test these predictions by investigating how the posterior probability of each context influences single-trial learning and how the prior probability of each context affects the subsequent expression of learning.

According to the NSSSM, the posterior probabilities determine how much each memory is updated when a perturbation is experienced (i.e., how learning is partitioned across memories). The NSSSM assumes that only one perturbation—the perturbation associated with the current context—influences the state feedback. Therefore, after observing the state feedback, if the current context is known, only the estimate of the

perturbation associated with this context should be updated:

$$\hat{x}^j_{t|t}(c_t) = \begin{cases} \hat{x}^j_{t|t-1} + k^j_t(y_t - \hat{x}^j_{t|t-1}) & \text{if } c_t = j \\ \hat{x}^j_{t|t-1} & \text{if } c_t \neq j, \end{cases} \tag{6.1}$$

where $\hat{x}^j_{t|t-1}$ and $\hat{x}^j_{t|t}$ are the predicted and updated estimates of the perturbation associated with context j; k^j_t is the Kalman gain for context j; y_t is the state feedback; and c_t is the current context.

However, in general, the context is not known and must be inferred. This uncertainty can be dealt with in a Bayesian manner by calculating a weighted average of updates across contexts:

$$\begin{aligned} \hat{x}^j_{t|t} &= \sum_j p(c_t = j|y_{1:t}, q_{1:t})\hat{x}^j_{t|t}(c_t) \\ &= \hat{x}^j_{t|t-1} + \underbrace{p(c_t = j|y_{1:t}, q_{1:t})k^j_t}_{\text{effective learning rate}}(y_t - \hat{x}^j_{t|t-1}), \end{aligned} \tag{6.2}$$

where the weights are given by the posterior probabilities of the contexts. The update rule in equation 6.2 is governed by three factors: the discrepancy between the predicted and actual state feedback (prediction error), the Kalman gain (learning rate), and the posterior probability that the perturbation generated the state feedback (responsibility). The posterior probability effectively scales the Kalman gain for each context, giving rise to an *effective learning rate* that lies between k^j_t (certain that $c_t = j$) and zero (certain that $c_t \neq j$). The posterior probability of each context is calculated by combining the prior probability of each context with the likelihood via Bayes' rule:

$$\underbrace{p(c_t = j|y_{1:t}, q_{1:t})}_{posterior} \propto \underbrace{p(c_t = j|y_{1:t-1}, q_{1:t})}_{prior} \underbrace{p(y_t|c_t = j, y_{1:t-1}, q_{1:t})}_{likelihood}. \tag{6.3}$$

The prior probability of each context is calculated before the movement and depends on the contextual cue (q_t), whereas the likelihood is calculated after the movement and

depends on the state feedback (y_t). Therefore, different combinations of contextual cues and state feedback will produce different posterior probabilities and hence different effective learning rates.

Here we test this prediction by examining single-trial learning under different combinations of contextual cues and force fields. We build on Chapter 3 by associating left and right control points on an object with CW and CCW force fields. We present control points and force fields that are either consistent or inconsistent with these trained associations and examine single-trial learning.

6.2 Methods

6.2.1 Participants

A total of 24 neurologically intact participants (10 males and 14 females; age 26.4 ± 4.2, mean ± s.d.) were recruited to participate in the experiment, which had been approved by the Cambridge Psychology Research Ethics Committee and the Columbia University IRB. All participants provided written informed consent and were right-handed according to the Edinburgh handedness inventory (Oldfield, 1971).

6.2.2 Apparatus

All experiments were performed using a vBOT planar robotic manipulandum with a virtual-reality system and air table, which have been described in Chapters 2 and 3.

6.2.3 Experiment

The aim of this experiment was to elucidate the relationship between context estimation and single-trial learning. We investigated whether context estimation depends on both the contextual cue (q^1 or q^2) and the perturbation (P^+ or P^-) and whether the context estimate modulates single-trial learning.

Paradigm

In all experiments, participants first performed a familiarisation phase of 80 trials consisting of null-field trials and channel trials for each contextual cue in a pseudorandomised order.

Participants began the experiment with a pre-training phase (Fig. 6.1). There were 4 blocks in the pre-training phase and each block contained 4 mini-blocks. Each mini-block was composed of a series of null-field trials followed by 1 triplet (Fig. 6.2). The null-field trials in each mini-block contained an equal number of both contextual cues presented in a pseudorandom order. The number of null-field trials in each mini-block was uniformly sampled without replacement from $\{8, 10, 12\}$; whenever the sample space emptied it was replenished. Each triplet type was presented once per block in a pseudorandom order. Note that there was no consistent relationship between contextual cues and force fields in the pre-training phase, as each triplet type was presented an equal number of times.

The pre-training training phase was followed by the training phase. There were 24 blocks in the training phase. Each block included the following sequence of trials: 2 channel trials; 32 force-field trials; 2 channel trials; 14, 16 or 18 null-field trials; 1 triplet; 6, 8 or 10 null-field trials; and 1 triplet. Each pair of channel trials presented each contextual cue once in an order that was counterbalanced every block. The 32 force-field trials were composed of 4 mini-blocks of 8 trials. Each mini-block contained

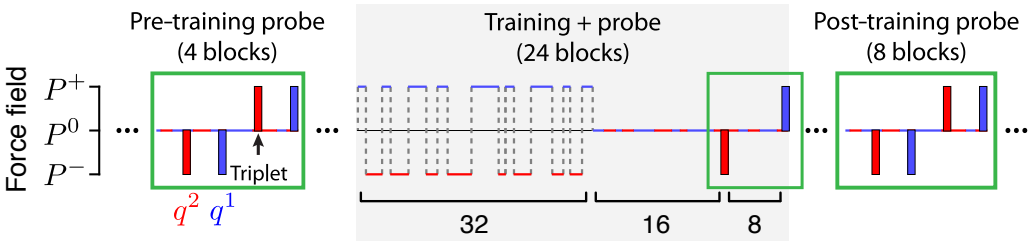


Figure 6.1. Experimental paradigm. In the pre-training phase, all 4 triplet types (filled bars) were presented in between null-field trials. The control point and the force field of the exposure trial of each triplet is denoted by the colour and direction of the filled bars, respectively. The red and blue colours also indicate the identity of the contextual cue on null-field and force-field trials outside of the triplets. During the force-field trials of the training phase, 2 cue-field associations were presented in a pseudorandom order. These trials were followed by null-field trials and then 2 triplets. The exposure trials of the triplets used the same cue-field pairs that were presented in the force-field trials. The post-training phase repeated the pre-training triplet paradigm.

an equal number of both force fields (P^+ and P^-) presented in a pseudorandom order. There was a one-to-one mapping between the force fields and the contextual cues. This mapping was preserved across blocks. The force field associated with each contextual cue was counterbalanced across participants. The mini-blocks of null-field trials that preceded each triplet contained an equal number of both contextual cues presented in a pseudorandom order. The number of null-field trials was uniformly sampled without replacement from the sample spaces described above, which were replenished whenever they emptied. The exposure trials of the triplets employed the same cue-field relationships that participants were trained on in the force-field trials. Thus, only 2 of the 4 triplet types were presented in the training phase. The order in which each triplet type was presented in a block was counterbalanced every block. A 60 s rest break was given after every 3 blocks during the training phase. After each rest break, 8 null-field trials were performed in which the contextual cues were presented in a pseudorandom order.

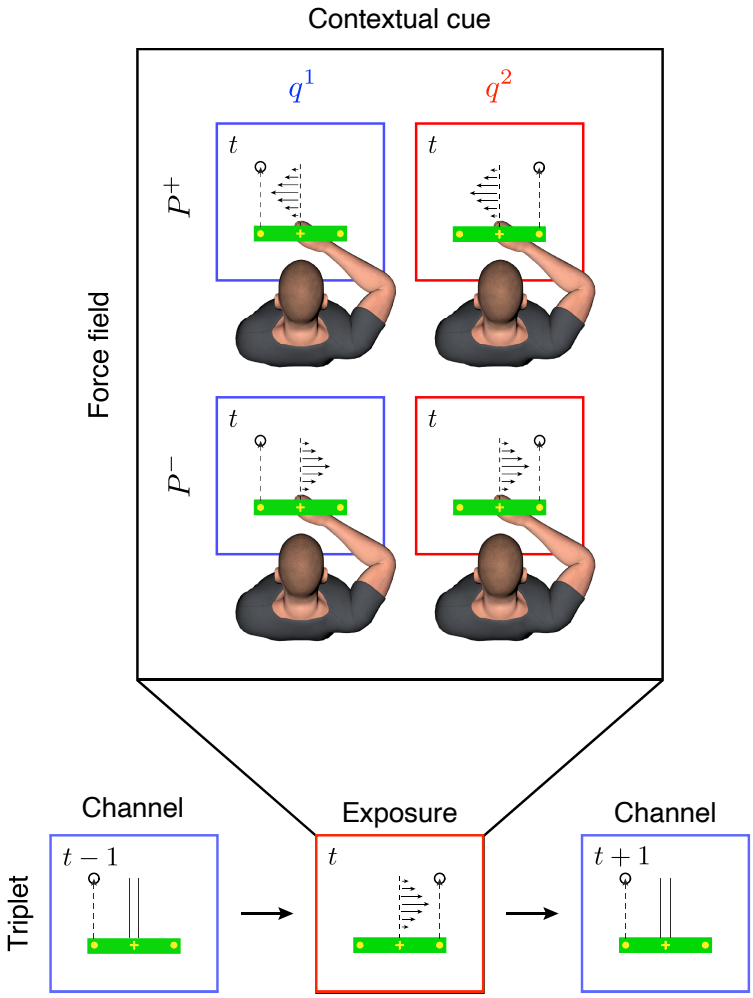


Figure 6.2. Triplets. On each trial, a target appeared either on the left or on the right and participants were instructed to move the corresponding left (q^1) or right (q^2) control point on a virtual object to the target. Both instructions required the same movement of the arm. Two possible force fields (P^+ or P^-) could be applied on each trial. To examine single-trial learning, we used a triplet probe in which an exposure (force-field) trial was performed in between two channel trials (both associated with q^1). The exposure trial could be one of 4 types based on all combinations of force fields (P^+ or P^-) and contextual cues (q^1 or q^2). Single-trial learning was calculated as the change in adaptation across the channel trials, with positive single-trial learning indicating increased adaptation to the force field presented on the exposure trial.

The training phase was followed by the post-training phase. The post-training phase started with 2 channel trials, 32 force-field trials, and 2 channel trials (just as in the training phase). Thereafter, the post-training phase was the same as the pre-training

phase, bar the following two exceptions. First, there were 2, 4 or 6 null-field trials between each triplet. Second, 8 blocks were performed instead of 4.

6.2.4 Data analysis

Data analysis was performed using MATLAB R2018b. Adaptation was measured on channel trials as the percentage of the force field that was compensated for, as described in Chapter 2.

To measure single-trial learning, we used a triplet probe consisting of a channel trial, an exposure (force-field) trial and a channel trial, in that order (Castro et al., 2014; Herzfeld et al., 2014). Single-trial learning (STL) was calculated as the change in adaptation from the first channel trial to the second second channel, where adaptation was calculated with respect to the force field presented on the exposure trial.

The exposure trial of the triplet could be one of 4 types based on different combinations of the force-field direction and the contextual cue. Thus, there were 4 triplet types in total. We refer to the field-cue pairings on the exposure trials of these triplets as $P^{+}q^{1}, P^{+}q^{2}, P^{-}q^{1}, P^{-}q^{2}$. The same contextual cue (q^{1}) was presented on both channel trials of the triples. The control point associated with q^{1} was counterbalanced across participants.

To identify changes in adaptation between triplets, repeated-measures ANOVAs and paired t-tests were performed. All statistical tests were two-sided with significance set to $P < 0.05$. Where values are reported, they represent the mean $\pm$ s.e.m.

6.3 Results

Throughout the experiment we used triplets to measure single-trial learning. Although the middle exposure trial of each triplet varied across triplets, the bracketing channel trials were always performed with the same contextual cue q^1 (the control point associated with q^1 was counterbalanced across participants).

In the pre-training phase, participants performed multiple repetitions of all triplet types, which were interspersed between null-field trials. The field-cue pairings presented on the exposure trials of these triplets could be P^+q^1, P^+q^2, P^-q^1 or P^-q^2. Fig 6.3b shows single-trial learning in the pre-training phase. There was no significant different in single-trial learning across the 4 triplets (repeated-measures ANOVA, $F_{3,69} = 1.87, P = 0.142$). If participants believed *a priori* that control points were informative of the perturbation, we would expect them to express more learning for the triplets with cue q^1 on the exposure trial, as cue q^1 was also presented on the neighbouring channel trials. The lack of such an effect suggests that participants did not hold such a belief at the start of the experiment.

In the training phase, participants experienced a consistent pairing of field and cue, that is P^+q^1 and P^-q^2 trials, and we used triplets with these same two combinations to assess single-trial learning. Although, in the pre-training phase there was no significant difference between the these two triplets (paired t-test, $t_{23} = 1.71, P = 0.100$), the two triplets diverged over the training phase. Fig 6.4 shows single-trial learning for the P^+q^1 (red) and P^-q^2 (blue) triplets over the course of the experiment (note that positive values of single-trial learning reflect adaptation to the field experienced on the exposure trial of the triplet). Single-trial learning increased for the P^+q^1 triplet and decreased for the P^-q^2 triplet. This suggests that participants learned that the control points were informative of the context, as they became able to differentially express learning based on whether the cue on the exposure trial of the triplet was the same as

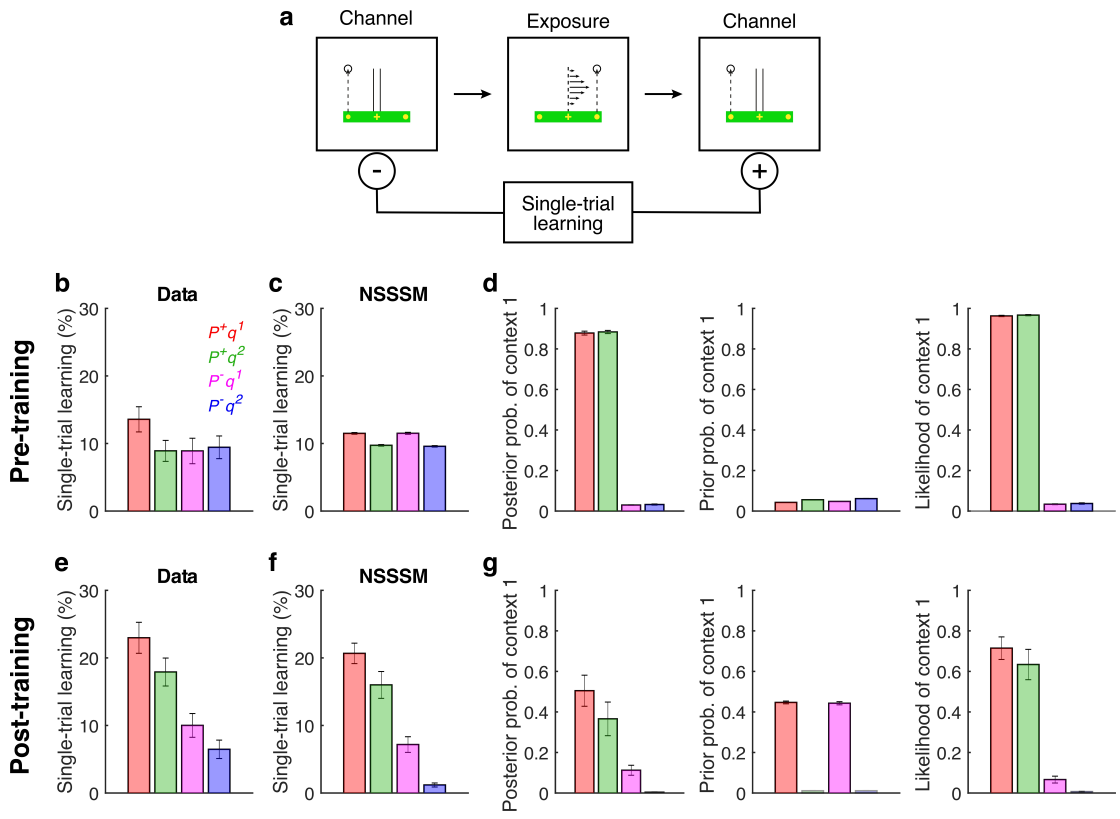


Figure 6.3. Bayesian context estimation modulates single-trial learning. **a**, Single-trial learning was calculated as the change in adaptation across the channel trials of a triplet, with positive single-trial learning indicating increased adaptation to the force field presented on the exposure trial. **b**, Single-trial learning for all combinations of cues and fields in the pre-training phase. Note that cue q^1 was presented on the channel trials of all triplets, and that this is crucial to the interpretation of the data. Positive values of single-trial learning reflects adaptation to the field experienced on the exposure trial of the triplet. Data averaged over all repetitions of each triplet and plotted as mean ± s.e.m. **c**, Same as **b** but for the fit of the NSSSM. **d**, The posterior probability, prior probability and normalised likelihood of context 1 (see text) on the exposure trial of each triplet in the NSSSM. The posterior probability is obtained by combining the prior and the likelihood via Bayes' rule. Probabilities were averaged over blocks 2-4 of the pre-training phase and plotted as mean ± s.e.m. Block 1 was excluded as P^+ (context 1) may not have been experienced on some of the triplets. Note that although we only show the probabilities for context 1, the probabilities for context 2 are implied by symmetry. For example, the probabilities for context 1 on the exposure trial of the P^+q^1 triplet are approximately equal to the probabilities for context 2 on the exposure trial of the P^-q^2 triplet, and the probabilities for context 1 on the exposure trial of the P^+q^2 triplet are approximately equal to the probabilities for context 2 on the exposure trial of the P^-q^1 triplet etc. The probabilities for the null-field context are equal to 1 minus the probabilities of contexts 1 and 2 summed together. **e**–**g** Same as **b**–**d** but for the post-training phase.

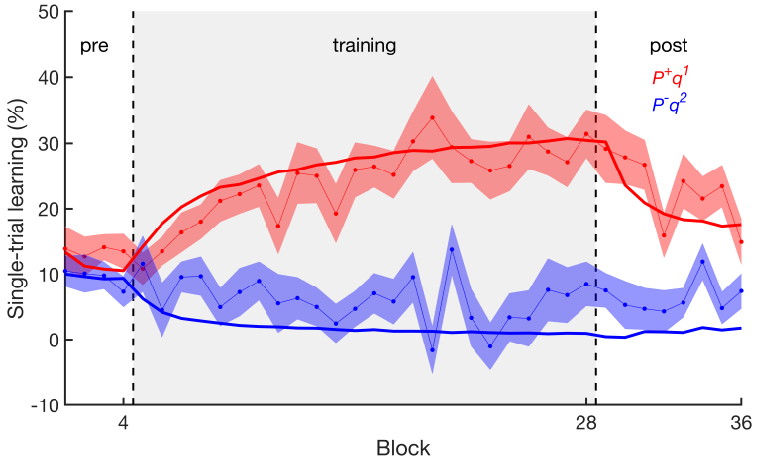


Figure 6.4. Single-trial learning for the training cue-field pairs throughout the experiment. Single-trial learning for the P^+q^1 triplet (red) and the P^-q^2 (blue) triplet in the pre-training, training (grey background) and post-training phases. Cue q^1 was presented on the channel trials of both triplets. Data are plotted as mean ± s.e.m. The solid lines show the mean NSSSM fit across participants. The model explains 78.0% of the variance of the data.

the cue on the neighbouring channel trials of the triplet. By the last two blocks of the training phase, single-trial learning was significantly different between the two triplet types (paired t-test, $t_{23} = 4.72, P = 9 \times 10^{-5}$).

Finally, and key to the study, the post-training phase examined all four triplet types again (as in the pre-training phase). How might single-trial learning now differ across the triplet types? The training phase has repeatedly presented two contexts of central importance: P^+q^1 (context 1) and P^-q^2 (context 2). On the exposure trial of a triplet, the learner must infer the context, and then update the memories for context 1 and context 2 based on this inference. Both memories may be updated on the exposure trial. However, the memory for context 1 will be preferentially expressed on the channel trials of the triplets, as cue q^1 is presented on these trials. Therefore, single-trial learning should be greatest when the posterior probability of context 1 on the exposure trial is greatest. This means that single-trial learning should be greatest for P^+q^1, as both the field and the cue are more probable in context 1; smallest for P^-q^2, as both the field and the cue are more probable in context 2; and intermediate for P^+q^2 and

P^-q^1, as the field and cue are more probable in different contexts (i.e., the contextual information is conflicting).

Fig 6.3e shows single-trial learning in the post-training phase for all four triplets. A gradation in the levels was observed, suggesting that both the force field and the control point contributed to context estimation (repeated-measures ANOVA, $F_{3,66} = 15.30, P = 1 \times 10^{-7}$). As predicted, single-trial learning was greatest for P^+q^1, smallest for P^-q^2, and intermediate for P^+q^2 and P^-q^1. Interestingly, single-trial learning was greater for P^+q^2 compared with P^-q^1, implying that the force field may have been more informative of the context than the contextual cue.

We fit the NSSSM model to the average data across participants. The model correctly instantiated three contexts—one for the null field and one each of the two force fields. In the pre-training phase, there is little difference in single-trial learning across the triplets (Fig 6.3c). On the exposure trial of each triplet, the prior probability of context 1 (and by symmetry context 2) is independent of the cue presented (Fig 6.3d prior), and hence the contextual cue is uninformative of the context. The uninformativeness of the cue arises because the cue emission distributions are the same across all contexts *a priori*—in the NSSSM, an uninformative cue is a cue that is equally probable in all contexts. The likelihood of context 1 depends on the force field presented (Fig 6.3d likelihood), and hence, in contrast to the cue, the state feedback is highly informative of the context. This results in a posterior distribution over contexts that is concentrated on context 1 or context 2 depending on the force field (Fig 6.3d posterior), and hence learning is predominantly assigned to one of these contexts. Crucially, the amount of learning that is expressed on the subsequent channel trial depends on the prior probability of each context, which determines how much each memory contributes to the motor output. Because the contextual cue is uninformative, the prior probabilities of contexts 1 and 2 are approximately the same on the channel trial, and so the memories of contexts 1 and 2 contribute equally to the motor output. Therefore, the expression of learning on the channel trial is independent of the partitioning of learning

between contexts 1 and 2 on the previous exposure trial, resulting in similar amounts of single-trial learning across the triplets.

The small prior probability of context 1 (and context 2) (Fig 6.3d prior) implies that the null-field context— which has been encountered most often—is the most probable context prior to movement. The high prior probability of the null-field context at this early stage of the experiment produces a relatively low amount of single-trial learning (Fig 6.3c).

The NSSSM is able to reproduce the divergence in single-trial learning between the $P^{+}q^{1}$ and $P^{-}q^{2}$ triplets in the training phase (Fig. 6.4 solid lines). This occurs as the context-specific cue emission distributions are updated during training. Because the model learns that q^{1} is more probable in context 1, it expresses more of the memory of context 1 and less of the memory of context 2 on the channel trials of the triplets (in which q^{1} is presented). This results in an increase in single-trial learning for the $P^{+}q^{1}$ triplet (where learning is predominantly assigned to context 1) and a decrease in single-trial learning for the $P^{-}q^{2}$ triplet (where learning is predominantly assigned to context 2). The NSSSM accounts for 78.0% of the variance for single-trial learning in these two triplets throughout the experiment.

In the post-training phase, single-trial learning in the NSSSM (Fig. 6.3f) shows the same gradation as the empirical data. Because the cue emission probabilities are updated during training, the prior probability of context 1 on the exposure trial of the triplet now depends on the contextual cue presented, and so the prior exhibits two levels (one for each cue) (Fig. 6.3g prior). As in the pre-training phase, the likelihood of context 1 also exhibits two levels (one for each force field) (Fig. 6.3g likelihood). When the two levels of the prior and the likelihood are combined via Bayes' rule, a posterior with four levels is produced (Fig. 6.3g posterior). Because the cue is informative of the context, the prior probability of context 1 on the channel trial (in which q^{1} is presented) is greater than the prior probability of context 2, and so the memory of

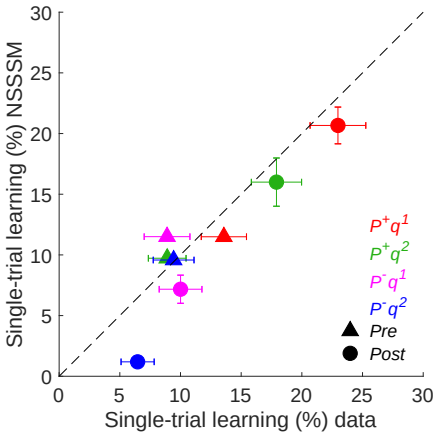


Figure 6.5. Comparison of single-trial learning between the empirical data and the NSSSM. The plot shows NSSSM fits vs. empirical single-trial learning for all four triplets of the pre-training (triangles) and post-training (circles) phases. Data averaged over all repetitions of each triplet and plotted as mean ± s.e.m. Some error bars are contained within the symbols. The model explains 74.0% of the variance of the data.

context 1 contributes more to the motor output than the memory of context 2. Hence, the expression of learning on the channel trial now depends on the partitioning of learning between contexts 1 and 2 on the previous exposure trial. This allows the four levels of the internal memory updates induced by the posterior to manifest as four levels of single-trial learning.

The NSSSM explains 74.0% of the variance for single-trial learning across all four triplets during the pre- and post-training phases (Fig 6.5).

6.4 Discussion

In this experiment, we used single-trial learning as a window into context estimation. This allowed us to dissect the contributions of cues and state feedback to context estimation and validate a number of key predictions of the NSSSM. First, we showed that learning is partitioned between context-specific memories in a Bayes-optimal manner,

that is, single-trial learning was proportional to the posterior context probability. Second, we showed that both the force field and the contextual cue are used to infer the context. Third, we showed that the effectiveness of contextual cues is not fixed but can adapt to the environment.

The idea that learning within internal models should be gated by context probabilities has been proposed before. In the MOSAIC model (Haruno et al., 2001; Wolpert and Kawato, 1998), internal models are trained via error signals that are scaled by their posterior probabilities. These posterior probabilities are referred to as responsibility signals. Responsibility signals are not unique to the MOSAIC model. They are a feature of all models that switch between multiple discrete latent states (e.g., Gaussian mixture models, the hidden Markov model, the switching state-space model), where they represent the posterior probability that a particular state generated the current observation. Importantly, the MOSAIC model did not provide experimental evidence of responsibility signals in human motor learning. Here we provide experimental evidence that responsibility signals do indeed gate human motor learning. Progress on this question was made possible by the recently developed triplet assay of single-trial learning (Castro et al., 2014; Herzfeld et al., 2014).

Note that the NSSSM is richer than the MOSAIC model in two key respects. First, the NSSSM contains an explicit model of how perturbations evolve over time, whereas the MOSAIC model does not. Therefore, memories in the NSSSM can be forgotten or drift over time, even in the absence of errors. In contrast, memories in the MOSAIC model do not change in the absence of errors. Second, the number of contexts and cues in the NSSSM is learned online, whereas the number of modules in the MOSAIC model is specified in advance and fixed. Therefore, while the NSSSM can adapt its complexity to the environment, the MOSAIC model cannot.

We found that participants were able to learn to use contextual cues to infer the current context. Initially, control points had no effect on single-trial learning. However, once

each control point had been paired with a different force field in the training phase, differences in single-trial learning emerged between the triplets. This suggests that the control points became able to preferentially weight some memories over others during the expression of learning. This raises the question of why some contextual cues (e.g., control points) can become effective with training and others (e.g., colours) can't (Gandolfo et al., 1996; Howard et al., 2013; Krakauer, 2009). One possibility is that participants enter the laboratory with strong prior beliefs about particular cues and their informativeness. For example, a participant may have experienced the color red with both heavy and light objects before, rendering it an uninformative cue, and a typical experiment may be too short to overcome this prior experience. However, it should be possible to overwhelm prior experience with evidence. In support of this view, color cues can facilitate adaptation to multiple dynamic contexts following extensive training over several days (Addou et al., 2011; Krouchev and Kalaska, 2003; Wada et al., 2003).

The post-exposure phase demonstrated that single-trial learning was modulated by both the force field direction and the contextual cue. This finding cannot be explained by models that assume that the contextual cue alone determines how much to update each memory (Kim et al., 2015; Lee and Schweighofer, 2009). Recently, Ingram et al. (2017) developed an 'error-tuned model' for sensorimotor learning that consisted of a set of primitives that were each able to compensate for errors in a single direction. When a kinematic error was experienced, each primitive updated its adaptive state based on two factors: i) the contextual tuning of the primitive to a visual cue; and ii) the ability of the primitive to reduce the kinematic error. While this superficially resembles the integration of contextual cues and state feedback in the NSSSM, a number of important differences should be noted. First, the error-tuned model is a mechanistic model, whereas the NSSSM is a normative probabilistic model. Second, the contextual tuning and preferred direction of each primitive in the error-tuned model are fixed from the start and do not change with experience. In contrast, in the NSSSM, the relationship between contexts and cues is learned online by estimating the cue emission

counts. Moreover, the relationship between contexts and perturbations is also learned online as the likelihood of each context depends on the time-varying estimate of the perturbation. These differences mean that the NSSSM has greater learning capabilities and flexibility compared to the error-tuned model, which has several hand-crafted features.

Chapter 7

Conclusion

The aim of this thesis was to investigate sensorimotor learning under switching dynamics, including switching from unperturbed to perturbed and switching between different perturbations. We first examined the role that muscle co-contraction plays in learning dynamics. When the dynamics of the motor plant undergo an abrupt switch (e.g., following the introduction of a force field), motor performance worsens, causing muscle co-contraction to increase. It is well known that muscle co-contraction leads to an immediate improvement in task accuracy, as the motor plant is more robust to external perturbations and intrinsic noise. In Chapter 2, we demonstrated an additional benefit, namely, that muscle co-contraction increases the rate of dynamic learning. We speculated that this increase in learning might arise as muscle co-contraction concentrates learning in the region of state-space close to the final solution.

The remaining chapters focused on learning multiple perturbations. In Chapter 3 we explored how controlling different locations on a tool affected dynamic learning. Flexible and dexterous tool use relies on the ability to learn the dynamics of objects. It has been assumed that objects dynamics are learned in a holistic way, that is, once the dynamics of an object are learned, one can manipulate the object in a variety of different scenarios (e.g., hitting, balancing, cleaning). Contrary to this view, we found

that learning to manipulate an object depends in a fundamental way on the point on the object that a person controls. For example, when a person strikes a ball with the midpoint vs. the tip of a bat, different motor memories are engaged, even if the motion of the hand and the bat are identical for either control point. This suggests that objects are not represented in the sensorimotor system as holistic entities, but rather in terms of the specific tasks and control problems that they are being used to perform.

We modeled this learning using a switching state space model in which the cues (control points) and sensory feedback were used to estimate which of the two context (dynamic perturbations) was likely to experience. Although this model fit the data well, it had the weakness that we specified the number of contexts and cues in the model to be the same as in the experiment. In Chapter 4, we relaxed this assumption by developing a nonparametric version of the model which does not pre-specify the number of contexts or cues but has to estimate these as part of the learning process.

This novel normative model of sensorimotor learning departed substantially from previous models. The model has five key innovations. First, we generalised the Kalman filter to environments with multiple switching contexts that are observed indirectly via state feedback and contextual cues. Second, we assumed that participants do not know the number of contexts or cues in the environment and must learn these online from experience. Third, we treated the parameters of the model as latent variables and model motor adaptation as combined Bayesian state and parameter estimation. This produces rich temporal behaviour on multiple timescales and increases the expressiveness of the model. Fourth, we introduced a new parameter called drift that can capture perturbations that increase over time (e.g., when muscles become fatigued with exertion). Fifth, we incorporated fictitious observations into channel trials, which support learning. The NSSSM provided a unifying explanation of phenomena such as spontaneous recovery and the effect of environmental consistency and context transition frequencies on single-trial learning.

In Chapter 5, we directly tested the NSSSM interpretation of spontaneous recovery. The NSSSM assumes that spontaneous recovery reflects the re-expression of a memory acquired in a previous context. This re-expression relies on Bayesian context estimation. Because there is no strong contextual information in channel trials, the re-expression is gradual. By briefly presenting two force-field trials early in the channel trial phase, we were able to produce rapid re-expression of the memory, which we refer to as 'evoked recovery'. Model comparison showed that the NSSSM was a superior fit to the data than the the multi-rate state-space model, which is typically used to explain spontaneous recovery.

The NSSSM specifies how sensory cues and sensory feedback should affect context estimation and how this should change with experience. In Chapter 6, we investigated these predictions of the model. Knowledge of the context allows the learner to i) engage the appropriate context-specific memory during movement and ii) assign responsibilities to context-specific memories when a movement is perturbed. We used single-trial learning as a window into context estimation. The NSSSM predicted that single-trial learning would be modulated by both contextual cues and state feedback as both influence context estimation. We confirmed this prediction. We also confirmed that control points only become effective as contextual cues once they have been paired with different perturbations.

7.1 Future directions

In a non-switching state-space model, there is only one state, and so the origin of each observation is known. This allows exact inference to be performed in closed-form using a Kalman filter. In contrast, in a switching state-space model, multiple states co-exist, and so the origin of each observation is unknown. Hence, exact inference in the switching state-space model—even when the number of contexts is known—is intractable. This leads to a well-known problem called the data association problem

(Bar-Shalom et al., 2009). Knowing the origin of each observation is crucial, as it determines which state to update following an observation. In the absence of such knowledge, one must consider all possible assignments of observations to states, which is computationally infeasible. For example, the number of possible associations of T observations to C states is C^T. Therefore, approximate inference schemes are required.

An interesting line of future research would be to investigate what kinds of approximations the brain makes to tackle intractable inference problems. Approximate Bayesian inference can be performed using either deterministic (e.g., variational Bayes) or sampling-based (e.g., Sequential Monte Carlo) algorithms (Bishop, 2006). In this thesis, we provided examples of both: the parametric SSSM in Chapter 3 used a deterministic algorithm known as the generalized pseudo-Bayesian estimator of order 1 (GPB1), and the nonparametric SSSM in Chapter 4 used a sampling-based algorithm known as particle learning. However, we make no claims at the algorithmic level. Instead, we draw conclusions at the computational level concerning the types of generative models that people build and the kinds of inferences that they perform in those generative models. It remains an open question whether studies of behaviour can shed light on the algorithms used by the brain, as different algorithms that realise or approximate the same computation may produce indistinguishable behaviour. Hence, progress at the algorithmic level may require converging evidence from behavioural experiments, imaging studies and neural recordings.

The NSSSM has a number of limitations that should be addressed in future research. For example, although the parameters of the NSSSM can be learned online, we assumed that the true parameters being learned are static (i.e., time-invariant). This simplifying assumption is unlikely to be true, particularly over long timescales. An obvious extension of the model would be to assume that the parameters—like the perturbations—vary over time. The time evolution of the parameters could be described by a transition function that itself has parameters. In principle, these higher-level parameters could also evolve over time, perhaps on a slower timescale. Thus, it is

possible to construct deep hierarchical state-space models, in which the parameters at each level determine the time evolution of the parameters at the level below (Mathys et al., 2011, 2014). Additional temporal richness could also be added by modelling each perturbation as a vector that is composed of fast and slow (and ultra-slow, hyper-slow etc.) states. In a nonparametric setting, the dimensionality of each perturbation could also be learned online (Fox et al., 2011a).

Once a memory has been instantiated in the NSSSM, it persists forever. This is because the NSSSM assumes that new contexts can be added but old contexts can never be removed. Hence, the number of contexts is a monotonically increasing function of time. This may not be an accurate description of the real world, in which contexts both come and go (i.e. a perturbation you will never experience again such as the last time you snowboard). An alternative model that addresses this shortcoming is the multiple target tracking model (Vo and Ma, 2006). In the multiple target tracking model, there are an unknown and time-varying number of targets (e.g., perturbations), whose latent states and parameters need be estimated from observations of unknown origin (Yıldırım et al., 2015). At each point in time, new targets can be born, and existing targets can die. When a target disappears from the environment, the memory associated with that target is no longer needed and can be removed. The removal of superfluous memories reduces the complexity of inference and frees up memory resources. However, the multiple target tracking model has its own shortcomings. For example, there is no notion of context transition probabilities, that is, the probability of observing a target on one trial does not depend on the target that was observed on the previous trial. An interesting direction for future research would be to merge the NSSSM and the multiple target tracking model into a single framework that incorporates desirable features of both models, such as the death of targets and Markovian transition probabilities.

Many tasks involve sequences of subtasks that are coordinated to achieve higher-level goals. Consider making a cup of tea, which involves boiling water, which in turn involves turning on a stove, filling a kettle, etc. Hierarchical models can support

efficient representations of such tasks and facilitate the discovery of new adaptive behaviors (Solway et al., 2014). The switching state-space model could be extended to environments with hierarchically-organised contexts. In a two-level hierarchy, the high-level context could determine the parameters of the hidden Markov model (e.g., the context transition probabilities and cue emission probabilities) at the level below, and the low-level context could determine which perturbation is observed. Thus, both the parameters of the hidden Markov model at the lower level and the observed perturbation can switch. Such a model could account for paradigms when the context transitions probabilities themselves switch over time (e.g., Experiment 2 in Herzfeld et al., 2014).

The state-space framework reduces the problem of learning a control policy in a high-dimensional state and action space to the abstract problem of estimating a low-dimensional perturbation. While this simplification helps to distill certain high-level processes, it also discards many details that make motor learning hard. For example, it does not address the curse of dimensionality, the exploration-exploitation dilemma and the temporal credit assignment problem. An interesting future direction would be to incorporate the NSSSM into a reinforcement learning algorithm that learns to control an anthropomorphic limb in simulation (e.g., using the MuJoCo physics engine). In the nonparametric setting, the dynamics of the limb could switch between an unknown number of probabilistic forward models. The agent could learn multiple context-specific dynamics models, value functions and control policies and perform state estimation, planning and real-time control to maximise a task-relevant reward function, such as the negative cost function of OFC. The richer state and action space as well as the ability to plan would allow new questions to be investigated with respect to sensorimotor learning under switching dynamics.

In conclusion, this thesis has provided novel experimental results which we have interpreted within a new framework, the NSSSM. This has led to predictions that we have experimentally confirmed.

References

Acerbi, L. and Ma, W. J. (2017). Practical Bayesian optimization for model fitting with Bayesian adaptive direct search. In *Advances in Neural Information Processing Systems*, pages 1836–1846.

Acerbi, L., Vijayakumar, S., and Wolpert, D. M. (2014). On the origins of suboptimality in human probabilistic inference. *PLoS Computational Biology*, 10(6):e1003661.

Acerbi, L., Wolpert, D. M., and Vijayakumar, S. (2012). Internal representations of temporal statistics and feedback calibrate motor-sensory interval timing. *PLoS Computational Biology*, 8(11):e1002771.

Ackerson, G. and Fu, K. (1970). On state estimation in switching environments. *IEEE Transactions on Automatic Control*, 15(1):10–17.

Addou, T., Krouchev, N., and Kalaska, J. F. (2011). Colored context cues can facilitate the ability to learn and to switch between multiple dynamical force fields. *Journal of Neurophysiology*, 106(1):163–183.

Ahmed, A. A., Wolpert, D. M., and Flanagan, J. R. (2008). Flexible representations of dynamics are used in object manipulation. *Current Biology*, 18(10):763–768.

Akaike, H. (1974). A new look at the statistical model identification. In *Selected Papers of Hirotugu Akaike*, pages 215–222. Springer.

Akazawa, K., Milner, T. E., and Stein, R. B. (1983). Modulation of reflex EMG and stiffness in response to stretch of human finger muscle. *Journal of Neurophysiology*, 49(1):16–27.

Albert, S. T. and Shadmehr, R. (2016). The neural feedback response to error as a teaching signal for the motor learning system. *Journal of Neuroscience*, 36(17):4832–4845.

Albert, S. T. and Shadmehr, R. (2017). Estimating properties of the fast and slow adaptive processes during sensorimotor adaptation. *Journal of Neurophysiology*, 119(4):1367–1393.

Albus, J. S. (1971). A theory of cerebellar function. *Mathematical Biosciences*, 10(1-2):25–61.

Baddeley, R., Ingram, H., and Miall, R. (2003). System identification applied to a visuomotor task: near-optimal human performance in a noisy changing task. *Journal of Neuroscience*, 23(7):3066–3075.

Bar-Shalom, Y., Daum, F., and Huang, J. (2009). The probabilistic data association filter. *IEEE Control Systems Magazine*, 29(6):82–100.

Bar-Shalom, Y., Li, X. R., and Kirubarajan, T. (2004). *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons.

Baugh, L. A., Kao, M., Johansson, R. S., and Flanagan, J. R. (2012). Material evidence: Interaction of well-learned priors and sensorimotor memory when lifting objects. *Journal of Neurophysiology*, 108(5):1262–1269.

Bays, P. M., Flanagan, J. R., and Wolpert, D. M. (2006). Attenuation of self-generated tactile sensations is predictive, not postdictive. *PLoS Biology*, 4(2):e28.

Bays, P. M., Wolpert, D. M., and Flanagan, J. R. (2005). Perception of the consequences of self-action is temporally tuned and event driven. *Current Biology*, 15(12):1125–1128.

Bell, C. C., Han, V., and Sawtell, N. B. (2008). Cerebellum-like structures and their implications for cerebellar function. *Annual Review of Neuroscience*, 31:1–24.

Bennett, D. J. (1993). Torques generated at the human elbow joint in response to constant position errors imposed during voluntary movements. *Experimental Brain Research*, 95(3):488–498.

Benton, A. L. (1959). Right-left discrimination and finger localization: Development and pathology.

Berniker, M., Franklin, D. W., Flanagan, J. R., Wolpert, D. M., and Kording, K. (2013). Motor learning of novel dynamics is not represented in a single global coordinate system: evaluation of mixed coordinate representations and local learning. *Journal of Neurophysiology*, 111(6):1165–1182.

Berniker, M. and Kording, K. (2008). Estimating the sources of motor errors for adaptation and generalization. *Nature Neuroscience*, 11(12):1454.

Bhanpuri, N. H., Okamura, A. M., and Bastian, A. J. (2013). Predictive modeling by the cerebellum improves proprioception. *Journal of Neuroscience*, 33(36):14301–14306.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Blakemore, S. J., Wolpert, D., and Frith, C. (2000). Why can't you tickle yourself? *Neuroreport*, 11(11):R11–16.

Blakemore, S.-J., Wolpert, D. M., and Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, 1(7):635.

Brashers-Krug, T., Shadmehr, R., and Bizzi, E. (1996). Consolidation in human motor memory. *Nature*, 382(6588):252.

Brayanov, J. B., Press, D. Z., and Smith, M. A. (2012). Motor memory is encoded as a gain-field combination of intrinsic and extrinsic action representations. *Journal of Neuroscience*, 32(43):14951–14965.

Brennan, A. E. and Smith, M. A. (2015). The decay of motor memories is independent of context change detection. *PLoS Computational Biology*, 11(6):e1004278.

Brooks, J. X., Carriot, J., and Cullen, K. E. (2015). Learning to expect the unexpected: rapid updating in primate cerebellum during voluntary self-motion. *Nature Neuroscience*, 18(9):1310.

Burdet, E., Osu, R., Franklin, D. W., Milner, T. E., and Kawato, M. (2001). The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature*, 414(6862):446.

Burdet, E., Tee, K. P., Mareels, I., Milner, T. E., Chew, C.-M., Franklin, D. W., Osu, R., and Kawato, M. (2006). Stability and motor adaptation in human arm movements. *Biological Cybernetics*, 94(1):20–32.

Burge, J., Ernst, M. O., and Banks, M. S. (2008). The statistical determinants of adaptation rate in human reaching. *Journal of Vision*, 8(4):20–20.

Bursztyn, L. L. and Flanagan, J. R. (2008). Sensorimotor memory of weight asymmetry in object manipulation. *Experimental Brain Research*, 184(1):127–133.

Caithness, G., Osu, R., Bays, P., Chase, H., Klassen, J., Kawato, M., Wolpert, D. M., and Flanagan, J. R. (2004). Failure to consolidate the consolidation theory of learning for sensorimotor adaptation tasks. *Journal of Neuroscience*, 24(40):8662–8671.

Cappé, O. (2011). Online EM algorithm for hidden Markov models. *Journal of Computational and Graphical Statistics*, 20(3):728–749.

Cappé, O. and Moulines, E. (2009). On-line expectation–maximization algorithm for latent data models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(3):593–613.

Carvalho, C. M., Johannes, M. S., Lopes, H. F., and Polson, N. G. (2010). Particle learning and smoothing. *Statistical Science*, 25(1):88–106.

Casadio, M., Pressman, A., and Mussa-Ivaldi, F. A. (2015). Learning to push and learning to move: the adaptive control of contact forces. *Frontiers in Computational Neuroscience*, 9:118.

Castro, L. N. G., Hadjiosif, A. M., Hemphill, M. A., and Smith, M. A. (2014). Environmental consistency determines the rate of motor adaptation. *Current Biology*, 24(10):1050–1061.

Castro, L. N. G., Monsen, C. B., and Smith, M. A. (2011). The binding of learning to action in motor adaptation. *PLoS Computational Biology*, 7(6):e1002052.

Chang, C.-B. and Athans, M. (1978). State estimation for discrete systems with switching parameters. *IEEE Transactions on Aerospace and Electronic Systems*, (3):418–425.

Charlesworth, J. D., Warren, T. L., and Brainard, M. S. (2012). Covert skill learning in a cortical-basal ganglia circuit. *Nature*, 486(7402):251.

Cheng, S. and Sabes, P. N. (2006). Modeling sensorimotor learning with linear dynamical systems. *Neural Computation*, 18(4):760–793.

Chib, V. S., Krutky, M. A., Lynch, K. M., and Mussa-Ivaldi, F. A. (2009). The separate neural control of hand movements and contact forces. *Journal of Neuroscience*, 29(12):3939–3947.

Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Foster, J. D., Nuyujukian, P., Ryu, S. I., and Shenoy, K. V. (2012). Neural population dynamics during reaching. *Nature*, 487(7405):51.

Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Ryu, S. I., and Shenoy, K. V. (2010). Cortical preparatory activity: representation of movement or first cog in a dynamical machine? *Neuron*, 68(3):387–400.

Churchland, M. M. and Shenoy, K. V. (2007). Temporal complexity and heterogeneity of single-neuron activity in premotor and motor cortex. *Journal of Neurophysiology*, 97(6):4235–4257.

Colebatch, J. and McCloskey, D. (1987). Maintenance of constant arm position or force: reflex and volitional components in man. *The Journal of physiology*, 386(1):247–261.

Conditt, M. A., Gandolfo, F., and Mussa-Ivaldi, F. A. (1997). The motor system does not learn the dynamics of the arm by rote memorization of past experience. *Journal of Neurophysiology*, 78(1):554–560.

Cothros, N., Wong, J., and Gribble, P. (2006). Are there distinct neural representations of object and limb dynamics? *Experimental Brain Research*, 173(4):689–697.

Cothros, N., Wong, J., and Gribble, P. L. (2009). Visual cues signaling object grasp reduce interference in motor learning. *Journal of Neurophysiology*, 102(4):2112–2120.

Crapse, T. B. and Sommer, M. A. (2008). Corollary discharge across the animal kingdom. *Nature Reviews Neuroscience*, 9(8):587.

Criscimagna-Hemminger, S. E. and Shadmehr, R. (2008). Consolidation patterns of human motor memory. *Journal of Neuroscience*, 28(39):9610–9618.

Danion, F., Diamond, J. S., and Flanagan, J. R. (2011). The role of haptic feedback when manipulating nonrigid objects. *Journal of Neurophysiology*, 107(1):433–441.

Darainy, M. and Ostry, D. J. (2008). Muscle cocontraction following dynamics learning. *Experimental Brain Research*, 190(2):153–163.

Day, K. A., Leech, K. A., Roemmich, R. T., and Bastian, A. J. (2018). Accelerating locomotor savings in learning: compressing four training days to one. *Journal of Neurophysiology*, 119(6):2100–2113.

Deisenroth, M. and Rasmussen, C. E. (2011). PILCO: A model-based and data-efficient approach to policy search. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, pages 465–472.

Del Moral, P., Doucet, A., and Singh, S. (2010). Forward smoothing using sequential Monte Carlo. *arXiv preprint arXiv:1012.5390*.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22.

Dideriksen, J. L., Negro, F., and Farina, D. (2015). The optimal neural strategy for a stable motor task requires a compromise between level of muscle cocontraction and synaptic gain of afferent feedback. *Journal of neurophysiology*, 114(3):1895–1911.

Diedrichsen, J. (2007). Optimal task-dependent changes of bimanual feedback control and adaptation. *Current Biology*, 17(19):1675–1679.

Diedrichsen, J., Shadmehr, R., and Ivry, R. B. (2010a). The coordination of movement: optimal feedback control and beyond. *Trends in Cognitive Sciences*, 14(1):31–39.

Diedrichsen, J., White, O., Newman, D., and Lally, N. (2010b). Use-dependent and error-based learning of motor behaviors. *Journal of Neuroscience*, 30(15):5159–5166.

Dimitriou, M. (2014). Human muscle spindle sensitivity reflects the balance of activity between antagonistic muscles. *Journal of Neuroscience*, 34(41):13644–13655.

Dingwell, J. B., Mah, C. D., and Mussa-Ivaldi, F. A. (2002). Manipulating objects with internal degrees of freedom: evidence for model-based control. *Journal of Neurophysiology*, 88(1):222–235.

Dingwell, J. B., Mah, C. D., and Mussa-Ivaldi, F. A. (2004). Experimentally confirmed mathematical model for human control of a non-rigid object. *Journal of Neurophysiology*, 91(3):1158–1170.

Doemges, F. and Rack, P. (1992). Task-dependent changes in the response of human wrist joints to mechanical disturbance. *The Journal of Physiology*, 447(1):575–585.

Domkin, D., Laczko, J., Jaric, S., Johansson, H., and Latash, M. L. (2002). Structure of joint variability in bimanual pointing tasks. *Experimental Brain Research*, 143(1):11–23.

Donchin, O., Francis, J. T., and Shadmehr, R. (2003). Quantifying generalization from trial-by-trial behavior of adaptive systems that learn with basis functions: theory and experiments in human motor control. *Journal of Neuroscience*, 23(27):9032–9045.

Doucet, A., Gordon, N. J., and Krishnamurthy, V. (2001). Particle filters for state estimation of jump markov linear systems. *IEEE Transactions on Signal Processing*, 49(3):613–624.

Ernst, M. O. and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870):429.

Faisal, A. A., Selen, L. P., and Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4):292.

Farnè, A., Iriki, A., and Làdavas, E. (2005). Shaping multisensory action–space with tools: evidence from patients with cross-modal extinction. *Neuropsychologia*, 43(2):238–248.

Fine, M. S. and Thoroughman, K. A. (2007). Trial-by-trial transformation of error into sensorimotor adaptation changes with environmental dynamics. *Journal of Neurophysiology*, 98(3):1392–1404.

Fiser, J., Berkes, P., Orbán, G., and Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. *Trends in Cognitive Sciences*, 14(3):119–130.

Flanagan, J. R. and Beltzner, M. A. (2000). Independence of perceptual and sensorimotor predictions in the size–weight illusion. *Nature Neuroscience*, 3(7):737.

Flanagan, J. R., Bowman, M. C., and Johansson, R. S. (2006). Control strategies in object manipulation tasks. *Current Opinion in Neurobiology*, 16(6):650–659.

Flanagan, J. R. and Wing, A. M. (1997). The role of internal models in motion planning and control: evidence from grip force adjustments during movements of hand-held loads. *Journal of Neuroscience*, 17(4):1519–1528.

Flash, T. and Hogan, N. (1985). The coordination of arm movements: an experimentally confirmed mathematical model. *Journal of Neuroscience*, 5(7):1688–1703.

Flash, T. and Mussa-Ivaldi, F. (1990). Human arm stiffness characteristics during the maintenance of posture. *Experimental Brain Research*, 82(2):315–326.

Fox, E., Sudderth, E. B., Jordan, M. I., and Willsky, A. S. (2009). Nonparametric Bayesian learning of switching linear dynamical systems. In *Advances in Neural Information Processing Systems*, pages 457–464.

Fox, E., Sudderth, E. B., Jordan, M. I., and Willsky, A. S. (2011a). Bayesian nonparametric inference of switching dynamic linear models. *IEEE Transactions on Signal Processing*, 59(4):1569–1585.

Fox, E. B., Sudderth, E. B., Jordan, M. I., Willsky, A. S., et al. (2011b). A sticky HDP-HMM with application to speaker diarization. *The Annals of Applied Statistics*, 5(2A):1020–1056.

Franklin, D. W., Burdet, E., Tee, K. P., Osu, R., Chew, C.-M., Milner, T. E., and Kawato, M. (2008). CNS learns stable, accurate, and efficient movements using a simple algorithm. *Journal of Neuroscience*, 28(44):11165–11173.

Franklin, D. W., Liaw, G., Milner, T. E., Osu, R., Burdet, E., and Kawato, M. (2007). Endpoint stiffness of the arm is directionally tuned to instability in the environment. *Journal of Neuroscience*, 27(29):7705–7716.

Franklin, D. W., Osu, R., Burdet, E., Kawato, M., and Milner, T. E. (2003). Adaptation to stable and unstable dynamics achieved by combined impedance control and inverse dynamics model. *Journal of Neurophysiology*, 90(5):3270–3282.

Franklin, D. W., So, U., Kawato, M., and Milner, T. E. (2004). Impedance control balances stability with metabolically costly muscle activation. *Journal of Neurophysiology*, 92(5):3097–3105.

Franklin, S., Wolpert, D. M., and Franklin, D. W. (2012). Visuomotor feedback gains upregulate during the learning of novel dynamics. *Journal of Neurophysiology*, 108(2):467–478.

Fu, Q. and Santello, M. (2012). Context-dependent learning interferes with visuomotor transformations for manipulation planning. *Journal of Neuroscience*, 32(43):15086–15092.

Fu, Q. and Santello, M. (2014). Retention and interference of learned dexterous manipulation: interaction between multiple sensorimotor processes. *Journal of Neurophysiology*, 113(1):144–155.

Galea, J. M., Mallia, E., Rothwell, J., and Diedrichsen, J. (2015). The dissociable effects of punishment and reward on motor learning. *Nature Neuroscience*, 18(4):597.

Gandevia, S., McCloskey, D., and Burke, D. (1992). Kinaesthetic signals and muscle contraction. *Trends in neurosciences*, 15(2):62–65.

Gandolfo, F., Mussa-Ivaldi, F., and Bizzi, E. (1996). Motor learning by field approximation. *Proceedings of the National Academy of Sciences*, 93(9):3843–3846.

George, A. P. and Powell, W. B. (2006). Adaptive stepsizes for recursive estimation with applications in approximate dynamic programming. *Machine Learning*, 65(1):167–198.

Gershman, S. J. and Blei, D. M. (2012). A tutorial on Bayesian nonparametric models. *Journal of Mathematical Psychology*, 56(1):1–12.

Ghahramani, Z. and Hinton, G. E. (2000). Variational learning for switching state-space models. *Neural Computation*, 12(4):831–864.

Gomi, H. and Kawato, M. (1996). Equilibrium-point control hypothesis examined by measured arm stiffness during multijoint movement. *Science*, 272(5258):117–120.

Goodbody, S. J. and Wolpert, D. M. (1998). Temporal and amplitude generalization in motor learning. *Journal of Neurophysiology*, 79(4):1825–1838.

Gordon, A. M., Westling, G., Cole, K. J., and Johansson, R. S. (1993). Memory representations underlying motor commands used during manipulation of common and novel objects. *Journal of Neurophysiology*, 69(6):1789–1796.

Gribble, P. L., Mullin, L. I., Cothros, N., and Mattar, A. (2003). Role of cocontraction in arm movement accuracy. *Journal of Neurophysiology*, 89(5):2396–2405.

Haggard, P. and Wolpert, D. M. (2005). Disorders of body scheme. In *In Freund, HJ, Jeannerod, M., Hallett, M., Leiguarda R.,(Eds.), Higher-Order Motor Disorders*.

Harris, C. M. and Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, 394(6695):780.

Haruno, M., Wolpert, D. M., and Kawato, M. (2001). MOSAIC model for sensorimotor learning and control. *Neural Computation*, 13(10):2201–2220.

Hastie, T., Tibshirani, R., Friedman, J., and Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*, 27(2):83–85.

Herzfeld, D. J. and Shadmehr, R. (2014). Motor variability is not noise, but grist for the learning mill. *Nature Neuroscience*, 17(2):149.

Herzfeld, D. J., Vaswani, P. A., Marko, M. K., and Shadmehr, R. (2014). A memory of errors in sensorimotor learning. *Science*, 345(6202):1349–1353.

Hewitt, A. L., Popa, L. S., Pasalar, S., Hendrix, C. M., and Ebner, T. J. (2011). Representation of limb kinematics in Purkinje cell simple spike discharge is conserved across multiple tasks. *Journal of Neurophysiology*, 106(5):2232–2247.

Hirashima, M. and Nozaki, D. (2012). Distinct motor plans form and retrieve distinct motor memories for physically identical movements. *Current Biology*, 22(5):432–436.

Hogan, N. (1985). The mechanics of multi-joint posture and movement control. *Biological Cybernetics*, 52(5):315–331.

Howard, I. S., Ingram, J. N., Franklin, D. W., and Wolpert, D. M. (2012). Gone in 0.6 seconds: the encoding of motor memories depends on recent sensorimotor states. *Journal of Neuroscience*, 32(37):12756–12768.

Howard, I. S., Ingram, J. N., and Wolpert, D. M. (2009). A modular planar robotic manipulandum with end-point torque control. *Journal of Neuroscience Methods*, 181(2):199–211.

Howard, I. S., Wolpert, D. M., and Franklin, D. W. (2013). The effect of contextual cues on the encoding of motor memories. *Journal of Neurophysiology*, 109(10):2632–2644.

Howard, I. S., Wolpert, D. M., and Franklin, D. W. (2015). The value of the follow-through derives from motor learning depending on future actions. *Current Biology*, 25(3):397–401.

Huang, H. J., Kram, R., and Ahmed, A. A. (2012). Reduction of metabolic cost during motor learning of arm reaching dynamics. *Journal of Neuroscience*, 32(6):2182–2190.

Huang, V. S., Haith, A., Mazzoni, P., and Krakauer, J. W. (2011). Rethinking motor learning and savings in adaptation paradigms: model-free memory for successful actions combines with internal models. *Neuron*, 70(4):787–801.

Huang, V. S. and Shadmehr, R. (2009). Persistence of motor memories reflects statistics of the learning event. *Journal of Neurophysiology*, 102(2):931–940.

Huberdeau, D. M., Haith, A. M., and Krakauer, J. W. (2015). Formation of a long-term memory for visuomotor adaptation following only a few trials of practice. *Journal of Neurophysiology*, 114(2):969–977.

Hwang, E. J., Smith, M. A., and Shadmehr, R. (2006). Dissociable effects of the implicit and explicit memory systems on learning control of reaching. *Experimental Brain Research*, 173(3):425–437.

Ingram, J. N., Flanagan, J. R., and Wolpert, D. M. (2013). Context-dependent decay of motor memories during skill acquisition. *Current Biology*, 23(12):1107–1112.

Ingram, J. N., Howard, I. S., Flanagan, J. R., and Wolpert, D. M. (2010). Multiple grasp-specific representations of tool dynamics mediate skillful manipulation. *Current Biology*, 20(7):618–623.

Ingram, J. N., Howard, I. S., Flanagan, J. R., and Wolpert, D. M. (2011). A single-rate context-dependent learning process underlies rapid adaptation to familiar object dynamics. *PLoS Computational Biology*, 7(9):e1002196.

Ingram, J. N., Sadeghi, M., Flanagan, J. R., and Wolpert, D. M. (2017). An error-tuned model for sensorimotor learning. *PLoS Computational Biology*, 13(12):e1005883.

Ingram, J. N. and Wolpert, D. M. (2011). Naturalistic approaches to sensorimotor control. In *Progress in Brain Research*, volume 191, pages 3–29. Elsevier.

Inoue, M., Uchimura, M., Karibe, A., O'Shea, J., Rossetti, Y., and Kitazawa, S. (2014). Three timescales in prism adaptation. *Journal of Neurophysiology*, 113(1):328–338.

Iriki, A., Tanaka, M., and Iwamura, Y. (1996). Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport*, 7(14):2325–2330.

Izawa, J., Criscimagna-Hemminger, S. E., and Shadmehr, R. (2012). Cerebellar contributions to reach adaptation and learning sensory consequences of action. *Journal of Neuroscience*, 32(12):4230–4239.

Izawa, J., Rane, T., Donchin, O., and Shadmehr, R. (2008). Motor adaptation as a process of reoptimization. *Journal of Neuroscience*, 28(11):2883–2891.

Izawa, J. and Shadmehr, R. (2011). Learning from sensory and reward prediction errors during motor adaptation. *PLoS Computational Biology*, 7(3):e1002012.

Jacobs, R. (1999). Optimal integration of texture and motion cues to depth. *Vision Research*, 39(21):3621–3629.

Jacobs, R. A., Jordan, M. I., Nowlan, S. J., Hinton, G. E., et al. (1991). Adaptive mixtures of local experts. *Neural Computation*, 3(1):79–87.

Jazayeri, M. and Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, 13(8):1020.

Jeffreys, H. (1998). *The theory of probability*. OUP Oxford.

Johansson, R. and Westling, G. (1988). Coordinated isometric muscle commands adequately and erroneously programmed for the weight during lifting task with precision grip. *Experimental Brain Research*, 71(1):59–71.

Johansson, R. S. and Flanagan, J. R. (2009). Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nature Reviews Neuroscience*, 10(5):345.

Johansson, R. S. and Westling, G. (1984). Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects. *Experimental Brain Research*, 56(3):550–564.

Johansson, R. S., Westling, G., Bäckström, A., and Flanagan, J. R. (2001). Eye–hand coordination in object manipulation. *Journal of Neuroscience*, 21(17):6917–6932.

Jones, K. E., Hamilton, A. F. d. C., and Wolpert, D. M. (2002). Sources of signal-dependent noise during isometric force production. *Journal of Neurophysiology*, 88(3):1533–1544.

Jordan, M. I. and Rumelhart, D. E. (1992). Forward models: supervised learning with a distal teacher. *Cognitive Science*, 16(3):307–354.

Kagerer, F. A., Contreras-Vidal, J. L., and Stelmach, G. E. (1997). Adaptation to gradual as compared with sudden visuo-motor distortions. *Experimental Brain Research*, 115(3):557–561.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45.

Kalman, R. E. and Bucy, R. S. (1961). New results in linear filtering and prediction theory. *Journal of Basic Engineering*, 83(1):95–108.

Kami, A., Meyer, G., Jezzard, P., Adams, M. M., Turner, R., and Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature*, 377(6545):155.

Kao, M. H., Doupe, A. J., and Brainard, M. S. (2005). Contributions of an avian basal ganglia–forebrain circuit to real-time modulation of song. *Nature*, 433(7026):638.

Karniel, A. and Mussa-Ivaldi, F. A. (2002). Does the motor control system use multiple models and context switching to cope with a variable environment? *Experimental Brain Research*, 143(4):520–524.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795.

Kaufman, M. T., Churchland, M. M., Ryu, S. I., and Shenoy, K. V. (2014). Cortical activity in the null space: permitting preparation without movement. *Nature Neuroscience*, 17(3):440.

Kawato, M. (1990). Feedback-error-learning neural network for supervised motor learning. In *Advanced Neural Computers*, pages 365–372. Elsevier.

Kawato, M., Furukawa, K., and Suzuki, R. (1987). A hierarchical neural-network model for control and learning of voluntary movement. *Biological Cybernetics*, 57(3):169–185.

Keisler, A. and Shadmehr, R. (2010). A shared resource between declarative memory and motor memory. *Journal of Neuroscience*, 30(44):14817–14823.

Kim, S., Oh, Y., and Schweighofer, N. (2015). Between-trial forgetting due to interference and time in motor adaptation. *PLoS One*, 10(11):e0142963.

Kitago, T., Ryan, S. L., Mazzoni, P., Krakauer, J. W., and Haith, A. M. (2013). Unlearning versus savings in visuomotor adaptation: comparing effects of washout, passage of time, and removal of errors on motor memory. *Frontiers in Human Neuroscience*, 7:307.

Kitazawa, S., Kimura, T., and Yin, P.-B. (1998). Cerebellar complex spikes encode both destinations and errors in arm movements. *Nature*, 392(6675):494.

Klassen, J., Tong, C., and Flanagan, J. R. (2005). Learning and recall of incremental kinematic and dynamic sensorimotor transformations. *Experimental Brain Research*, 164(2):250–259.

Kluzik, J., Diedrichsen, J., Shadmehr, R., and Bastian, A. J. (2008). Reach adaptation: what determines whether we learn an internal model of the tool or adapt the model of our arm? *Journal of Neurophysiology*, 100(3):1455–1464.

Kobayashi, Y., Kawano, K., Takemura, A., Inoue, Y., Kitama, T., Gomi, H., and Kawato, M. (1998). Temporal firing patterns of Purkinje cells in the cerebellar ventral paraflocculus during ocular following responses in monkeys II. Complex spikes. *Journal of Neurophysiology*, 80(2):832–848.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., and Shams, L. (2007a). Causal inference in multisensory perception. *PLoS One*, 2(9):e943.

Körding, K. P., Ku, S.-p., and Wolpert, D. M. (2004). Bayesian integration in force estimation. *Journal of Neurophysiology*, 92(5):3161–3165.

Körding, K. P., Tenenbaum, J. B., and Shadmehr, R. (2007b). The dynamics of memory as a consequence of optimal adaptation to a changing body. *Nature Neuroscience*, 10(6):779.

Körding, K. P. and Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427(6971):244.

Korenberg, A. T. and Ghahramani, Z. (2002). A Bayesian view of motor adaptation. *Current Psychology of Cognition*, 21(4/5):537–564.

Krakauer, J. W. (2009). Motor learning and consolidation: the case of visuomotor rotation. In *Progress in Motor Control*, pages 405–421. Springer.

Krakauer, J. W., Ghez, C., and Ghilardi, M. F. (2005). Adaptation to visuomotor transformations: consolidation, interference, and forgetting. *Journal of Neuroscience*, 25(2):473–478.

Krakauer, J. W., Ghilardi, M.-F., and Ghez, C. (1999). Independent learning of internal models for kinematic and dynamic control of reaching. *Nature Neuroscience*, 2(11):1026.

Krakauer, J. W. and Mazzoni, P. (2011). Human sensorimotor learning: adaptation, skill, and beyond. *Current Opinion in Neurobiology*, 21(4):636–644.

Krakauer, J. W., Mazzoni, P., Ghazizadeh, A., Ravindran, R., and Shadmehr, R. (2006). Generalization of motor learning depends on the history of prior action. *PLoS Biology*, 4(10):e316.

Krakauer, J. W., Pine, Z. M., Ghilardi, M.-F., and Ghez, C. (2000). Learning of visuomotor transformations for vectorial planning of reaching trajectories. *Journal of Neuroscience*, 20(23):8916–8924.

Krouchev, N. I. and Kalaska, J. F. (2003). Context-dependent anticipation of different task dynamics: rapid recall of appropriate motor skills using visual cues. *Journal of Neurophysiology*, 89(2):1165–1175.

Krutky, M. A., Ravichandran, V. J., Trumbower, R. D., and Perreault, E. J. (2009). Interactions between limb and environmental mechanics influence stretch reflex sensitivity in the human arm. *Journal of Neurophysiology*, 103(1):429–440.

Kurtzer, I., DiZio, P. A., and Lackner, J. R. (2005). Adaptation to a novel multi-force environment. *Experimental Brain Research*, 164(1):120–132.

Lackner, J. R. and Dizio, P. (1994). Rapid adaptation to coriolis force perturbations of arm trajectory. *Journal of Neurophysiology*, 72(1):299–313.

Lackner, J. R. and DiZio, P. (1996). Motor function in microgravity: movement in weightlessness. *Current Opinion in Neurobiology*, 6(6):744–750.

Lametti, D. R., Houle, G., and Ostry, D. J. (2007). Control of movement variability and the regulation of limb impedance. *Journal of Neurophysiology*, 98(6):3516–3524.

Land, M. F. and Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 352(1358):1231–1239.

Land, M. F., Mennie, N., and Rusted, J. (1999). Eye movements and the roles of vision in activities of daily living: making a cup of tea. *Perception*, 28(4):1311–1328.

Laquitaine, S. and Gardner, J. L. (2018). A switching observer for human perceptual estimation. *Neuron*, 97(2):462–474.

Latash, M. L., Scholz, J. P., and Schöner, G. (2002). Motor control strategies revealed in the structure of motor variability. *Exercise and Sport Sciences Reviews*, 30(1):26–31.

Lee, J.-Y. and Schweighofer, N. (2009). Dual adaptation supports a parallel architecture of motor memory. *Journal of Neuroscience*, 29(33):10396–10404.

Lillicrap, T. P. and Scott, S. H. (2013). Preference distributions of primary motor cortex neurons reflect control solutions optimized for limb biomechanics. *Neuron*, 77(1):168–179.

Liu, D. and Todorov, E. (2007). Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. *Journal of Neuroscience*, 27(35):9354–9368.

Liu, J. and Reinkensmeyer, D. J. (2007). Motor adaptation to a small force field superimposed on a large background force. *Experimental Brain Research*, 178(3):402–414.

Lonini, L., Dipietro, L., Zollo, L., Guglielmelli, E., and Krebs, H. I. (2009). An internal model for acquisition and retention of motor learning during arm reaching. *Neural Computation*, 21(7):2009–2027.

Malfait, N., Shiller, D. M., and Ostry, D. J. (2002). Transfer of motor learning across arm configurations. *Journal of Neuroscience*, 22(22):9656–9660.

Manohar, S. G., Chong, T. T.-J., Apps, M. A., Batla, A., Stamelou, M., Jarman, P. R., Bhatia, K. P., and Husain, M. (2015). Reward pays the cost of noise reduction in motor and cognitive control. *Current Biology*, 25(13):1707–1716.

Marko, M. K., Haith, A. M., Harran, M. D., and Shadmehr, R. (2012). Sensitivity to prediction error in reach adaptation. *Journal of Neurophysiology*, 108(6):1752–1763.

Marr, D. and Thach, W. T. (1991). A theory of cerebellar cortex. In *From the Retina to the Neocortex*, pages 11–50. Springer.

Mathys, C., Daunizeau, J., Friston, K. J., and Stephan, K. E. (2011). A Bayesian foundation for individual learning under uncertainty. *Frontiers in Human Neuroscience*, 5:39.

Mathys, C. D., Lomakina, E. I., Daunizeau, J., Iglesias, S., Brodersen, K. H., Friston, K. J., and Stephan, K. E. (2014). Uncertainty in perception and the Hierarchical Gaussian Filter. *Frontiers in Human Neuroscience*, 8:825.

McDougle, S. D., Bond, K. M., and Taylor, J. A. (2015). Explicit and implicit processes constitute the fast and slow processes of sensorimotor learning. *Journal of Neuroscience*, 35(26):9568–9579.

Mehta, B. and Schaal, S. (2002). Forward models in visuomotor control. *Journal of Neurophysiology*, 88(2):942–953.

Miall, R. C., Christensen, L. O., Cain, O., and Stanley, J. (2007). Disruption of state estimation in the human lateral cerebellum. *PLoS Biology*, 5(11):e316.

Milner, T. E. and Franklin, D. W. (2005). Impedance control and internal model use during the initial stage of adaptation to novel dynamics in humans. *The Journal of Physiology*, 567(2):651–664.

Milner, T. E. and Hinder, M. R. (2006). Position information but not force information is used in adapting to changes in environmental dynamics. *Journal of Neurophysiology*, 96(2):526–534.

Mitrovic, D., Klanke, S., Osu, R., Kawato, M., and Vijayakumar, S. (2010). A computational model of limb impedance control based on principles of internal model uncertainty. *PLoS One*, 5(10):e13601.

Miyazaki, M., Nozaki, D., and Nakajima, Y. (2005). Testing Bayesian models of human coincidence timing. *Journal of Neurophysiology*, 94(1):395–399.

Modchalingam, S., Vachon, C. M., Marius't Hart, B., and Henriques, D. Y. (2019). The effects of awareness of the perturbation during motor adaptation on hand localization. *PLoS One*, 14(8):e0220884.

Morehead, J. R., Qasim, S. E., Crossley, M. J., and Ivry, R. (2015). Savings upon re-aiming in visuomotor adaptation. *Journal of Neuroscience*, 35(42):14386–14396.

Müller, H. and Sternad, D. (2004). Decomposition of variability in the execution of goal-oriented tasks: three components of skill improvement. *Journal of Experimental Psychology: Human Perception and Performance*, 30(1):212.

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

Mussa-Ivaldi, F. A., Hogan, N., and Bizzi, E. (1985). Neural, mechanical, and geometric factors subserving arm posture in humans. *Journal of Neuroscience*, 5(10):2732–2743.

Nagengast, A. J., Braun, D. A., and Wolpert, D. M. (2009). Optimal control predicts human performance on objects with internal degrees of freedom. *PLoS Computational Biology*, 5(6):e1000419.

Nashed, J. Y., Crevecoeur, F., and Scott, S. H. (2012). Influence of the behavioral goal and environmental obstacles on rapid feedback responses. *Journal of Neurophysiology*, 108(4):999–1009.

Nashed, J. Y., Crevecoeur, F., and Scott, S. H. (2014). Rapid online selection between multiple motor plans. *Journal of Neuroscience*, 34(5):1769–1780.

Nasseroleslami, B., Hasson, C. J., and Sternad, D. (2014). Rhythmic manipulation of objects with complex dynamics: predictability over chaos. *PLoS Computational Biology*, 10(10):e1003900.

Nielsen, J. and Kagamihara, Y. (1993). The regulation of presynaptic inhibition during co-contraction of antagonistic muscles in man. *The Journal of physiology*, 464(1):575–593.

Nikooyan, A. A. and Ahmed, A. A. (2014). Reward feedback accelerates motor learning. *Journal of Neurophysiology*, 113(2):633–646.

Nissen, M. J. and Bullemer, P. (1987). Attentional requirements of learning: evidence from performance measures. *Cognitive Psychology*, 19(1):1–32.

Nozaki, D., Kurtzer, I., and Scott, S. H. (2006). Limited transfer of learning between unimanual and bimanual skills within the same limb. *Nature Neuroscience*, 9(11):1364.

Nozaki, D. and Scott, S. H. (2009). Multi-compartment model can explain partial transfer of learning within the same limb between unimanual and bimanual reaching. *Experimental Brain Research*, 194(3):451–463.

Oh, Y. and Schweighofer, N. (2018). To overwrite or to recall? Individual differences in motor adaptation. *bioRxiv*, page 439406.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1):97–113.

Omrani, M., Kaufman, M. T., Hatsopoulos, N. G., and Cheney, P. D. (2017). Perspectives on classical controversies about the motor cortex. *Journal of Neurophysiology*, 118(3):1828–1848.

Osu, R., Burdet, E., Franklin, D. W., Milner, T. E., and Kawato, M. (2003). Different mechanisms involved in adaptation to stable and unstable dynamics. *Journal of Neurophysiology*, 90(5):3255–3269.

Osu, R., Hirai, S., Yoshioka, T., and Kawato, M. (2004a). Random presentation enables subjects to adapt to two opposing forces on the hand. *Nature Neuroscience*, 7(2):111.

Osu, R., Kamimura, N., Iwasaki, H., Nakano, E., Harris, C. M., Wada, Y., and Kawato, M. (2004b). Optimal impedance control for task achievement in the presence of signal-dependent noise. *Journal of Neurophysiology*, 92(2):1199–1215.

Özkan, E., Lindsten, F., Fritsche, C., and Gustafsson, F. (2014). Recursive maximum likelihood identification of jump Markov nonlinear systems. *IEEE Transactions on Signal Processing*, 63(3):754–765.

Patton, J. L., Wei, Y. J., Bajaj, P., and Scheidt, R. A. (2013). Visuomotor learning enhanced by augmenting instantaneous trajectory error feedback during reaching. *PLoS One*, 8(1):e46466.

Pekny, S. E., Criscimagna-Hemminger, S. E., and Shadmehr, R. (2011). Protection and expression of human motor memories. *Journal of Neuroscience*, 31(39):13829–13839.

Pélisson, D., Alahyane, N., Panouilleres, M., and Tilikete, C. (2010). Sensorimotor adaptation of saccadic eye movements. *Neuroscience & Biobehavioral Reviews*, 34(8):1103–1120.

Penny, W. D., Stephan, K. E., Daunizeau, J., Rosa, M. J., Friston, K. J., Schofield, T. M., and Leff, A. P. (2010). Comparing families of dynamic causal models. *PLoS Computational Biology*, 6(3):e1000709.

Penny, W. D., Stephan, K. E., Mechelli, A., and Friston, K. J. (2004). Comparing dynamic causal models. *Neuroimage*, 22(3):1157–1172.

Proske, U. and Gandevia, S. C. (2012). The proprioceptive senses: their roles in signaling body shape, body position and movement, and muscle force. *Physiological reviews*, 92(4):1651–1697.

Pruszynski, J. A., Kurtzer, I., Lillicrap, T. P., and Scott, S. H. (2009). Temporal evolution of "automatic gain-scaling". *Journal of Neurophysiology*, 102(2):992–1003.

Ribot-Ciscar, E. and Roll, J.-P. (1998). Ago-antagonist muscle spindle inputs contribute together to joint movement coding in man. *Brain research*, 791(1-2):167–176.

Richter, S., Jansen-Osmann, P., Konczak, J., and Kalveram, K.-T. (2004). Motor adaptation to different dynamic environments is facilitated by indicative context stimuli. *Psychological Research*, 68(4):245–251.

Rigoux, L., Stephan, K. E., Friston, K. J., and Daunizeau, J. (2014). Bayesian model selection for group studies—revisited. *Neuroimage*, 84:971–985.

Robbins, H. and Monro, S. (1951). A stochastic approximation method. *The Annals of Mathematical Statistics*, pages 400–407.

Robinson, F. R., Noto, C. T., and Bevans, S. E. (2003). Effect of visual error size on saccade adaptation in monkey. *Journal of Neurophysiology*, 90(2):1235–1244.

Roemmich, R. T. and Bastian, A. J. (2015). Two ways to save a newly learned motor pattern. *Journal of Neurophysiology*, 113(10):3519–3530.

Roemmich, R. T. and Bastian, A. J. (2018). Closing the loop: from motor neuroscience to neurorehabilitation. *Annual Review of Neuroscience*, 41:415–429.

Roitman, A. V., Pasalar, S., Johnson, M. T., and Ebner, T. J. (2005). Position, direction of movement, and speed tuning of cerebellar Purkinje cells during circular manual tracking in monkey. *Journal of Neuroscience*, 25(40):9244–9257.

Ronsse, R., Wei, K., and Sternad, D. (2010). Optimal control of a hybrid rhythmic-discrete task: the bouncing ball revisited. *Journal of Neurophysiology*, 103(5):2482–2493.

Rothwell, J., Traub, M., Day, B., Obeso, J., Thomas, P., and Marsden, C. (1982). Manual motor performance in a deafferented man. *Brain*, 105(Pt 3):515–542.

Salimi, I., Hollender, I., Frazier, W., and Gordon, A. M. (2000). Specificity of internal representations underlying grasping. *Journal of Neurophysiology*, 84(5):2390–2397.

Sanger, T. D. (1994). Neural network learning control of robot manipulators using gradually increasing task difficulty. *IEEE Transactions on Robotics and Automation*, 10(3):323–333.

Santello, M. and Soechting, J. F. (2000). Force synergies for multifingered grasping. *Experimental Brain Research*, 133(4):457–467.

Sarwary, A. M., Stegeman, D. F., Selen, L. P., and Medendorp, W. P. (2015). Generalization and transfer of contextual cues in motor learning. *Journal of Neurophysiology*, 114(3):1565–1576.

Scheidt, R. A., Reinkensmeyer, D. J., Conditt, M. A., Rymer, W. Z., and Mussa-Ivaldi, F. A. (2000). Persistence of motor adaptation during constrained, multi-joint, arm movements. *Journal of Neurophysiology*, 84(2):853–862.

Scholz, J. P. and Schöner, G. (1999). The uncontrolled manifold concept: identifying control variables for a functional task. *Experimental Brain Research*, 126(3):289–306.

Schwarz, G. et al. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

Schween, R., McDougle, S. D., Hegele, M., and Taylor, J. A. (2019). Explicit strategies in force field adaptation. *bioRxiv*.

Scott, S. H. (2004). Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews Neuroscience*, 5(7):532.

Scott, S. H. (2008). Inconvenient truths about neural processing in primary motor cortex. *The Journal of Physiology*, 586(5):1217–1224.

Scott, S. H., Cluff, T., Lowrey, C. R., and Takei, T. (2015). Feedback control during voluntary motor actions. *Current opinion in neurobiology*, 33:85–94.

Selen, L. P., Beek, P. J., and Van Dieën, J. H. (2005). Can co-activation reduce kinematic variability? A simulation study. *Biological Cybernetics*, 93(5):373–381.

Selen, L. P., Franklin, D. W., and Wolpert, D. M. (2009). Impedance control reduces instability that arises from motor noise. *Journal of Neuroscience*, 29(40):12606–12616.

Sethuraman, J. (1994). A constructive definition of Dirichlet priors. *Statistica Sinica*, pages 639–650.

Shadmehr, R. and Brashers-Krug, T. (1997). Functional stages in the formation of human long-term motor memory. *Journal of Neuroscience*, 17(1):409–419.

Shadmehr, R. and Mussa-Ivaldi, F. A. (1994). Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience*, 14(5):3208–3224.

Shadmehr, R., Smith, M. A., and Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual Review of Neuroscience*, 33:89–108.

Sheahan, H. R., Franklin, D. W., and Wolpert, D. M. (2016). Motor planning, not execution, separates motor memories. *Neuron*, 92(4):773–779.

Shergill, S. S., Bays, P. M., Frith, C. D., and Wolpert, D. M. (2003). Two eyes for an eye: the neuroscience of force escalation. *Science*, 301(5630):187–187.

Shmuelof, L., Huang, V. S., Haith, A. M., Delnicki, R. J., Mazzoni, P., and Krakauer, J. W. (2012). Overcoming motor "forgetting" through reinforcement of learned actions. *Journal of Neuroscience*, 32(42):14617–14621a.

Shumway, R. H. and Stoffer, D. S. (1991). Dynamic linear models with switching. *Journal of the American Statistical Association*, 86(415):763–769.

Sing, G. C., Joiner, W. M., Nanayakkara, T., Brayanov, J. B., and Smith, M. A. (2009). Primitives for motor adaptation reflect correlated neural tuning to position and velocity. *Neuron*, 64(4):575–589.

Smith, M. A., Ghazizadeh, A., and Shadmehr, R. (2006). Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biology*, 4(6):e179.

Solway, A., Diuk, C., Córdova, N., Yee, D., Barto, A. G., Niv, Y., and Botvinick, M. M. (2014). Optimal behavioral hierarchy. *PLoS Computational Biology*, 10(8):e1003779.

Sommer, M. A. and Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature*, 444(7117):374.

Stein, J. and Glickstein, M. (1992). Role of the cerebellum in visual guidance of movement. *Physiological Reviews*, 72(4):967–1017.

Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., and Friston, K. J. (2009). Bayesian model selection for group studies. *Neuroimage*, 46(4):1004–1017.

Sussillo, D., Churchland, M. M., Kaufman, M. T., and Shenoy, K. V. (2015). A neural network that finds a naturalistic solution for the production of muscle activity. *Nature Neuroscience*, 18(7):1025.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Suzuki, M., Shiller, D. M., Gribble, P. L., and Ostry, D. J. (2001). Relationship between cocontraction, movement kinematics and phasic muscle activity in single-joint arm movement. *Experimental Brain Research*, 140(2):171–181.

Takahashi, C. D., Nemet, D., Rose-Gottron, C. M., Larson, J. K., Cooper, D. M., and Reinkensmeyer, D. J. (2006). Effect of muscle fatigue on internal model formation and retention during reaching with the arm. *Journal of Applied Physiology*, 100(2):695–706.

Takiyama, K., Hirashima, M., and Nozaki, D. (2015). Prospective errors determine motor learning. *Nature Communications*, 6:5925.

Taylor, J. A., Krakauer, J. W., and Ivry, R. B. (2014). Explicit and implicit contributions to learning in a sensorimotor adaptation task. *Journal of Neuroscience*, 34(8):3023–3032.

Taylor, J. L. and McCloskey, D. (1992). Detection of slow movements imposed at the elbow during active flexion in man. *The Journal of physiology*, 457(1):503–513.

Teh, Y. W., Jordan, M. I., Beal, M. J., and Blei, D. M. (2006). Hierarchical dirichlet processes. *Journal of the American Statistical Association*, 101(476):1566–1581.

Thoroughman, K. A. and Shadmehr, R. (1999). Electromyographic correlates of learning an internal model of reaching movements. *Journal of Neuroscience*, 19(19):8573–8588.

Thoroughman, K. A. and Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. *Nature*, 407(6805):742.

Thoroughman, K. A. and Taylor, J. A. (2005). Rapid reshaping of human motor generalization. *Journal of Neuroscience*, 25(39):8948–8953.

Todorov, E. and Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, 5(11):1226.

Tong, C., Wolpert, D. M., and Flanagan, J. R. (2002). Kinematics and dynamics are not represented independently in motor working memory: evidence from an interference study. *Journal of Neuroscience*, 22(3):1108–1113.

Trewartha, K. M., Garcia, A., Wolpert, D. M., and Flanagan, J. R. (2014). Fast but fleeting: adaptive motor learning processes associated with aging and cognitive decline. *Journal of Neuroscience*, 34(40):13411–13421.

Tseng, Y.-w., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., and Bastian, A. J. (2007). Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *Journal of Neurophysiology*, 98(1):54–62.

Tumer, E. C. and Brainard, M. S. (2007). Performance variability enables adaptive plasticity of 'crystallized'adult birdsong. *Nature*, 450(7173):1240.

Turnham, E. J., Braun, D. A., and Wolpert, D. M. (2011). Facilitation of learning induced by both random and gradual visuomotor task variation. *Journal of Neurophysiology*, 107(4):1111–1122.

Uno, Y., Kawato, M., and Suzuki, R. (1989). Formation and control of optimal trajectory in human multijoint arm movement. *Biological Cybernetics*, 61(2):89–101.

Valero-Cuevas, F. J., Venkadesan, M., and Todorov, E. (2009). Structured variability of muscle activations supports the minimal intervention principle of motor control. *Journal of Neurophysiology*, 102(1):59–68.

Vallbo, Å. (1970). Discharge patterns in human muscle spindle afferents during isometric voluntary contractions. *Acta Physiologica Scandinavica*, 80(4):552–566.

Vallbo, A. (1974). Human muscle spindle discharge during isometric voluntary contractions. Amplitude relations between spindle frequency and torque. *Acta Physiologica Scandinavica*, 90(2):319–336.

van Beers, R., Sittig, A., and Gon, J. (1999). Integration of proprioceptive and visual position-information: An experimentally supported model. *Journal of Neurophysiology*, 81(3):1355–1364.

Vaswani, P. A. and Shadmehr, R. (2013). Decay of motor memories in the absence of error. *Journal of Neuroscience*, 33(18):7700–7709.

Verstynen, T. and Sabes, P. N. (2011). How each movement changes the next: an experimental and theoretical study of fast adaptive priors in reaching. *Journal of Neuroscience*, 31(27):10050–10059.

Vetter, P. and Wolpert, D. M. (2000). Context estimation for sensorimotor control. *Journal of Neurophysiology*, 84(2):1026–1034.

Vo, B.-N. and Ma, W.-K. (2006). The Gaussian mixture probability hypothesis density filter. *IEEE Transactions on Signal Processing*, 54(11):4091–4104.

Wada, Y., Kawabata, Y., Kotosaka, S., Yamamoto, K., Kitazawa, S., and Kawato, M. (2003). Acquisition and contextual switching of multiple internal models for different viscous force fields. *Neuroscience Research*, 46(3):319–331.

Wainscott, S. K., Donchin, O., and Shadmehr, R. (2005). Internal models and contextual cues: encoding serial order and direction of movement. *Journal of Neurophysiology*, 93(2):786–800.

Wang, J. X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., Hassabis, D., and Botvinick, M. (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nature Neuroscience*, 21(6):860.

Wei, K. and Kording, K. (2009). Relevance of error: what drives motor adaptation? *Journal of Neurophysiology*, 101(2):655–664.

Wei, K. and Körding, K. (2010). Uncertainty of feedback and state estimation determines the speed of motor adaptation. *Frontiers in Computational Neuroscience*, 4:11.

Westling, G. and Johansson, R. (1984). Factors influencing the force control during precision grip. *Experimental Brain Research*, 53(2):277–284.

White, O. and Diedrichsen, J. (2013). Flexible switching of feedback control mechanisms allows for learning of different task dynamics. *PLoS One*, 8(2):e54771.

Williams, C. K. and Rasmussen, C. E. (2006). *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA.

Willingham, D. B., Nissen, M. J., and Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(6):1047.

Wise, A. K. and Fallon, J. B. (2002). The effect of muscle contraction on kinaesthesia. In *Sensorimotor Control of Movement and Posture*, pages 87–94. Springer.

Wise, A. K., Gregory, J. E., and Proske, U. (1998). Detection of movements of the human forearm during and after co-contractions of muscles acting at the elbow joint. *The Journal of physiology*, 508(1):325–330.

Wise, A. K., Gregory, J. E., and Proske, U. (1999). The responses of muscle spindles to small, slow movements in passive muscle and during fusimotor activity. *Brain research*, 821(1):87–94.

Witt, J. K., Proffitt, D. R., and Epstein, W. (2005). Tool use affects perceived distance, but only when you intend to use it. *Journal of Experimental Psychology: Human Perception and Performance*, 31(5):880.

Wolpert, D. M., Diedrichsen, J., and Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience*, 12(12):739.

Wolpert, D. M. and Flanagan, J. R. (2010). Q&A: Robotics as a tool to understand the brain. *BMC Biology*, 8(1):92.

Wolpert, D. M. and Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, 3(11s):1212.

Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, 269(5232):1880–1882.

Wolpert, D. M. and Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, 11(7-8):1317–1329.

Wong, J., Wilson, E. T., Malfait, N., and Gribble, P. L. (2009). Limb stiffness is modulated with spatial accuracy requirements during movement in the absence of destabilizing forces. *Journal of Neurophysiology*, 101(3):1542–1549.

Wu, H. G., Miyamoto, Y. R., Castro, L. N. G., Ölveczky, B. P., and Smith, M. A. (2014). Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nature Neuroscience*, 17(2):312.

Yamamoto, S. and Kitazawa, S. (2001). Sensation at the tips of invisible tools. *Nature Neuroscience*, 4(10):979.

Yeo, S.-H., Wolpert, D. M., and Franklin, D. W. (2015). Coordinate representations for interference reduction in motor learning. *PLoS One*, 10(6):e0129388.

Yıldırım, S., Jiang, L., Singh, S. S., and Dean, T. A. (2015). Calibrating the Gaussian multi-target tracking model. *Statistics and Computing*, 25(3):595–608.

Yokoi, A., Hirashima, M., and Nozaki, D. (2011). Gain field encoding of the kinematics of both arms in the internal model enables flexible bimanual action. *Journal of Neuroscience*, 31(47):17058–17068.

Zarahn, E., Weston, G. D., Liang, J., Mazzoni, P., and Krakauer, J. W. (2008). Explaining savings for visuomotor adaptation: linear time-invariant state-space models are not sufficient. *Journal of Neurophysiology*, 100(5):2537–2548.

Zhang, Z., Guo, D., Huber, M. E., Park, S.-W., and Sternad, D. (2018). Exploiting the geometry of the solution space to reduce sensitivity to neuromotor noise. *PLoS Computational Biology*, 14(2):e1006013.

Appendix A

Instructions to participants in the muscle co-contraction experiment

Instructions to the stiff group (provided prior to commencement of the pulse phase)

I'm now going to describe the next phase of the experiment. Each trial will start when you place the green cursor in the center of the white home position. A brief force will then be applied to the handle of the robot in a randomly chosen direction. This will only last for half a second. Your task will be to keep the handle as still as possible. To keep the handle still, you will need to stiffen up your arm by tensing the muscles in your hand, arm and chest. If on one trial you get pushed outside of the red ring that surrounds the white home position, you should try to stiffen up more on the next trial to help you stay within the ring. Ideally, you should keep the handle as still and as close to the center of the white home position as possible. Once you have completed this phase of the experiment, all subsequent trials will be similar to the ones you've just done; that is, they will involve you making straight reaching movements to one of four targets. However, at some point your movements will be unexpectedly thrown off course by forces generated by the robot. Your task will be to keep your movements as straight as possible. In particular, you should try to stay as close as possible to the

red lines that lead from the home position to the targets. You may find that stiffening up your arm will help you to do this. If at any point you feel as though your arm is too sore or tired, just let me know and you can have a rest break immediately. You will automatically get a rest break every 5 minutes though and ideally you should wait until then in order to allow the experiment to run smoothly.

Instructions to the relaxed group (provided prior to commencement of the pulse phase)

I'm now going to describe the next phase of the experiment. Each trial will start when you place the green cursor in the center of the white home position. A brief force will then be applied to the handle of the robot in a randomly chosen direction. This will only last for half a second. You should let your hand move freely to wherever the handle takes it. To do this, you should relax the muscles in your arm as much as possible so that you don't provide any resistance to the handle as it moves. Once the handle has stopped the robot will automatically bring you back to the central home position. Once again, you should keep your arm as relaxed as possible when this happens. Throughout all the trials of this phase you will see a red ring surrounding the white home position. This ring has no relevance and you should ignore it. Once you have completed this phase of the experiment, all subsequent trials will be similar to the ones you've just done; that is, they will involve you making straight reaching movements to one of four targets. However, at some point your movements will be unexpectedly thrown off course by forces generated by the robot. Try not to stiffen up when this happens. If you relax your arm you will find that your movements will naturally become straighter as you adapt to the forces.

Instructions to the control group (provided prior to commencement of the experiment)

At some point during the experiment, your movements will be unexpectedly thrown off course by forces generated by the robot. When this happens, keep trying to reach the target, and you will find that you will naturally adapt to the forces over time.

Appendix B

The parametric switching state-space model

B.1 State inference

Several techniques have been developed to infer the joint state, $\{c_t, \boldsymbol{x}_t\}$, of an SSSM (Ackerson and Fu, 1970; Bar-Shalom et al., 2004; Chang and Athans, 1978; Doucet et al., 2001; Ghahramani and Hinton, 2000; Shumway and Stoffer, 1991). Here, we focus on a technique known as the generalized pseudo-Bayesian estimator of order 1 (GPB1) (Bar-Shalom et al., 2004), which we adapt for an SSSM in which the context, c_t, emits observable cues q_t. The GPB1 is an assumed-density filtering method that approximates the exact posterior of the state (a mixture of Gaussians with m^l components, where m is the number of modes and l is the number of trials) with a single Gaussian.

To maintain an estimate of the state, two Kalman filters operate in parallel (Fig. 3.2). Each filter has a different observation vector, $\boldsymbol{h}^{(j)}$ (defined in the text), and is thus specialized for a different context; all remaining parameters are the same for both filters. On trial t, both filters operate on the same Gaussian estimate of the continuous

state from the previous trial, with mean $\hat{\boldsymbol{x}}_{t-1|t-1}$ and covariance $\boldsymbol{V}_{t-1|t-1}$, to generate a Gaussian estimate of the continuous state on the current trial, with mean $\hat{\boldsymbol{x}}_{t|t}^{(j)}$ and covariance $\boldsymbol{V}_{t|t}^{(j)}$. This produces a mixture of Gaussians with mixing weights given by the probability of each context $\gamma_t^{(j)}$. Because the number of possible context sequences doubles with every trial, the number of components in this mixture will also double if no approximations are made, making inference intractable. Therefore, after each trial, the mixture of Gaussians is approximated by a single Gaussian via moment matching (Fig. 3.2), rendering inference tractable but approximate.

Upon observation of the contextual cue q_t, but prior to movement, the probability of each context (assuming equiprobable context transitions) is

$$\rho_t^{(j)} = \frac{\phi_{q_t j}}{\sum_{j'=1}^{2} \phi_{q_t j'}}. \tag{B.1}$$

The predicted observation used for control is

$$\hat{y}_t = \sum_{j=1}^{2} \rho_t^{(j)} \hat{y}_t^{(j)}, \tag{B.2}$$

where $\hat{y}_t^{(j)} = \boldsymbol{h}^{(j)} \hat{\boldsymbol{x}}_{t|t-1}^{(j)}$ is the observation predicted for each context. We assume that the predicted observation is expressed through a motor action that is corrupted by zero-mean Gaussian noise (see Model Fitting).

After the movement, the likelihood of the observation y_t given each context is

$$\lambda_t^{(j)} = \mathcal{N}(y_t | \hat{y}_t^{(j)}, \sigma_t^{2(j)}). \tag{B.3}$$

Here, we use $\mathcal{N}(y_t | \hat{y}_t^{(j)}, \sigma_t^{2(j)})$ to denote the probability density function for the normal distribution with mean $\hat{y}_t^{(j)}$ and variance $\sigma_t^{2(j)}$, where $\sigma_t^{2(j)}$ is the standard Kalman filter innovation covariance. The final probability of each context is given by Bayes'

rule:

$$\gamma_t^{(j)} = \frac{\lambda_t^{(j)} \rho_t^{(j)}}{\sum_{i=1}^{2} \lambda_t^{(i)} \rho_t^{(i)}}. \tag{B.4}$$

Finally, a single Gaussian estimate of the perturbations is obtained via moment matching:

$$\hat{\boldsymbol{x}}_{t|t} = \sum_{j=1}^{2} \gamma_t^{(j)} \hat{\boldsymbol{x}}_{t|t}^{(j)}, \tag{B.5}$$

$$\boldsymbol{V}_{t|t} = \sum_{j=1}^{2} \gamma_t^{(j)} \{ \boldsymbol{V}_{t|t}^{(j)} + [\hat{\boldsymbol{x}}_{t|t}^{(j)} - \hat{\boldsymbol{x}}_{t|t}][\hat{\boldsymbol{x}}_{t|t}^{(j)} - \hat{\boldsymbol{x}}_{t|t}]^{\mathrm{T}} \}. \tag{B.6}$$

B.2 Parameter learning

To learn the cue emission probabilities, we perform expectation maximization (EM)—an iterative method for obtaining maximum-likelihood estimates of unknown parameters in probabilistic models involving latent variables (Dempster et al., 1977). We utilize an online formulation of EM that can be interpreted as a stochastic approximation recursion on the expected complete-data sufficient statistics. A detailed description of this method can be found in refs (Cappé, 2011; Cappé and Moulines, 2009). Here we provide a brief outline as applied to our perturbation study.

The algorithm iterates between an E step, in which the joint probability of c_t and q_t is estimated, and an M step, in which the joint probability of c_t and q_t is normalized to obtain the probability of q_t given c_t.

B.2.1 Stochastic E step

The aim of the E step is to count the number of times that each value of c_t and q_t co-occur. If both c_t and q_t were known, this would involve summing (over trials) a

count of these co-occurrences:

$$\mathcal{S}_l = \sum_{t=1}^{l} \boldsymbol{S}_t(q_t, c_t), \tag{B.7}$$

where the matrix $\boldsymbol{S}_t(q_t, c_t)$ is the complete-data sufficient statistic for a multinomial distribution and has a single unity entry defined by the indicator function

$$s_t(q_t, c_t)_{kj} = \begin{cases} 1 & \text{if } q_t = k, c_t = j \\ 0 & \text{otherwise.} \end{cases} \tag{B.8}$$

To avoid storing the entire history of sufficient statistics, this sum can be calculated in a recursive manner:

$$\mathcal{S}_t = \mathcal{S}_{t-1} + \boldsymbol{S}_t(q_t, c_t), \tag{B.9}$$

with $\mathcal{S}_0$ initialized to a matrix of zeros. However, because c_t is a latent variable, this recursion must be modified in the following two ways. First, $\boldsymbol{S}_t(q_t, c_t)$ is estimated in the form of an expectation:

$$\mathbb{E}_{\hat{\theta}_{t-1}}[\boldsymbol{S}_t(q_t, c_t)|q_{1:t}, y_{1:t}] = \sum_{j=1}^{2} \gamma_t^{(j)} \boldsymbol{S}_t(q_t, c_t = j), \tag{B.10}$$

where the hat notation is used to indicate an estimated value. Second, because of observation noise, the expectations in equation B.10 will show trial-to-trial variability. Therefore, an adaptive stepsize, $\eta_t = \alpha/(\alpha + t - 1)$, is used to average the sequence of expectations and promote convergence (George and Powell, 2006). Because the stepsize decreases to zero according to the requirements of stochastic approximation theory, $\sum_{t=1}^{\infty} \eta_t = \infty$ and $\sum_{t=1}^{\infty} \eta_t^2 < \infty$, the effects of noise are eliminated in the long run (Robbins and Monro, 1951). Increasing α slows the rate at which the stepsize decreases, placing greater weight on more recent observations compared with older observations. We treat α as a free parameter of the model. Taken together, these two

modifications lead to the following recursion:

$$\mathcal{S}_t = (1 - \eta_t)\mathcal{S}_{t-1} + \eta_t \mathbb{E}_{\hat{\theta}_{t-1}}[\boldsymbol{S}_t(q_t, c_t)|q_{1:t}, y_{1:t}], \tag{B.11}$$

which computes a running average of the expected complete-data sufficient statistics. It is useful to note that the convex combination in equation B.11 ensures that all elements of $\mathcal{S}_t$ sum to 1. As such, $\mathcal{S}_t$ can be interpreted as an estimate of the joint probability of c_t and q_t, which is proportional to the expected number of times that each value of c_t and q_t co-occurred.

While it is possible to compute smoothed estimates of the complete-data sufficient statistics (Del Moral et al., 2010), exact computation of these smoothed estimates is intractable, requiring sequential Monte Carlo sampling techniques (Özkan et al., 2014). For simplicity, we only compute filtered estimates of the complete-data sufficient statistics, as in equation B.10.

B.2.2 M step

To calculate $\hat{\boldsymbol{\Phi}}_t$, we normalize $\mathcal{S}_t$ by dividing each element by its corresponding column sum to ensure that the probability of either cue being emitted in each context is 1. Note all elements of $\mathcal{S}_t$ must be nonzero for the solution of the EM algorithm to be well defined (otherwise some probabilities will always remain zero). Therefore, we omit the M step for the first 8 trials to ensure that both contextual cues have been observed.

Importantly, to infer the probability of each context given an observed cue, the model of how contexts emit cues (equation 3.7) is inverted (equation B.1).

B.3 Model implementation

We applied the above inference and learning algorithms to a sequence of noiseless observations, $y_{1:l}$, assigned values of 0 (null-field trials) or $\pm$ 1 (force-field trials), corresponding to the actual perturbations delivered by the robot. We also examined adding zero-mean Gaussian noise (variance r) to the delivered sequence of perturbations to include observation noise. However, because the observation noise that participants perceived is unknown to us, it must be marginalized out (e.g., using Monte Carlo integration). The inclusion of observation noise also results in a stochastic objective function, which requires alternative parameter optimization techniques. For simplicity, we only report fits to noiseless observations as the addition of observation noise did not qualitatively change the model fit or the outcome of model comparison.

To update the estimate of the continuous state, we applied the standard Kalman filter recursive equations (Kalman, 1960). The prior estimate of the continuous state was initialized with mean, $\hat{\boldsymbol{x}}_{0|0}$, and covariance, $\boldsymbol{V}_{0|0}$, equal to zero and the mean steady-state covariance matrix across Kalman filters, respectively. The prediction error of each Kalman filter, $y_t - \hat{y}_t^{(j)}$, was set to zero on channel trials.

We set $\boldsymbol{h}^{(j)}$ and $\boldsymbol{\Pi}$ as described in the text and initialized $\hat{\boldsymbol{\Phi}}_0$ as a symmetric matrix:

$$\hat{\boldsymbol{\Phi}}_0 = \begin{pmatrix} \beta & 1-\beta \\ 1-\beta & \beta \end{pmatrix}, \tag{B.12}$$

where $\beta = 0.5 + \epsilon$ and $\epsilon = 10^{-6}$. Thus the learner starts naive with respect to how cues relate to contexts. $\hat{\boldsymbol{\Phi}}_t$ was updated recursively using the online expectation maximization algorithm. The small amount of jitter, ϵ, was added to allow symmetry breaking. We used $q_t = 1$ and $q_t = 2$ for the left and right control points, respectively.

Appendix C

The nonparametric switching state-space model

Here we develop a switching state-space model with an unknown, potentially infinite number of contexts and cues. Before we describe the nonparametric switching state-space model, we first introduce the underlying theory of Dirichlet and hierarchical Dirichlet processes (Teh et al., 2006).

C.1 The Dirichlet process

A Dirichlet process (DP) is a distribution over a countably infinite random probability measure on a parameter space Θ. The Dirichlet process is the infinite-dimensional generalisation of the Dirichlet distribution. Whereas samples from the Dirichlet distribution assign probability mass to each of the n items in $\{1, \ldots, n\}$, samples from a Dirichlet process assign probability mass to a countably infinite subset of a parameter space Θ. We write $G_0 \sim \mathrm{DP}(\gamma, H)$ if G_0 is a random distribution distributed as a Dirichlet process with concentration parameter γ and base distribution H. The base

distribution H is the mean of the Dirichlet process ($\mathbb{E}[G_0] = H$) and the concentration parameter γ can be understood as the inverse variance, which governs how much G_0 varies around H. A random distribution drawn from a DP can be expressed using a stick-breaking representation (Sethuraman, 1994):

$$G_0 = \sum_{k=1}^{\infty} \beta_k \delta_{\theta_k} \qquad \theta_k \mid H \sim H, \tag{C.1}$$

where δ_{θ_k} denotes a delta-Dirac mass located at θ_k and β_k is the weight assigned to that mass. The base distribution H furnishes a prior over the parameters θ_k. The weights are sampled via a stick-breaking construction:

$$\beta_k = \beta'_k \prod_{l=1}^{k-1} (1 - \beta'_l) \qquad \beta'_k \mid \gamma \sim \text{Beta}(1, \gamma). \tag{C.2}$$

We denote this distribution by $\boldsymbol{\beta} = (\beta_k)_{k=1}^{\infty} \sim \text{GEM}(\gamma)$. Note that $\sum_{k=1}^{\infty} \beta_k = 1$. This representation makes explicit that draws from a Dirichlet process are countably infinite and discrete (even if H is continuous).

A common application of the Dirichlet process is as a prior on the parameters and mixing coefficients of a mixture model with an unknown, potentially infinite number of components. This is known as the DPMM. In this setting, each θ_k represents the parameters of component k and $\boldsymbol{\beta}$ represents the vector of mixing coefficients. To generate datapoint i, we first draw parameter $\bar{\theta}_i \sim G_0$ and then draw observation $y_i \sim F(\bar{\theta}_i)$, where $F(\bar{\theta}_i)$ is the distribution of y_i given $\bar{\theta}_i$.

The hierarchical Dirichlet process

In the hierarchical Dirichlet process, a second Dirichlet process is defined with base distribution G_0 and concentration parameter α. Hence, in the hierarchical Dirichlet process, the base distribution is itself distributed as a Dirichlet process. A draw from

this Dirichlet process $G_j \sim \mathrm{DP}(\alpha, G_0)$ can again be expressed using a stick-breaking representation:

$$\begin{aligned} G_j &= \sum_{t=1}^{\infty} \tilde{\pi}_{jt} \delta_{\tilde{\theta}_{jt}} \qquad \tilde{\theta}_{jt} \mid G_0 \sim G_0 \\ \tilde{\boldsymbol{\pi}}_j \mid \alpha &\sim \mathrm{GEM}(\alpha), \end{aligned} \tag{C.3}$$

where $\tilde{\boldsymbol{\pi}}_j = (\tilde{\pi}_{jt})_{t=1}^{\infty}$. An alternative representation of G_j is

$$G_j = \sum_{k=1}^{\infty} \pi_{jk} \delta_{\theta_k} \qquad \boldsymbol{\pi}_j \mid \alpha, \boldsymbol{\beta} \sim \mathrm{DP}(\alpha, \boldsymbol{\beta}), \tag{C.4}$$

where $\boldsymbol{\beta}$ and $\boldsymbol{\pi}_j = (\pi_{jk})_{k=1}^{\infty}$ are interpreted as distributions on the positive integers. This representation is made possible as G_0 and G_j have support at the same points: $(\theta_k)_{k=1}^{\infty}$. The hyperparameters of the hierarchical Dirichlet process are H, γ and α.

One important application of the hierarchical Dirichlet process is as a prior on the parameters and state transition matrix of a hidden Markov model (HMM) with an unknown, potentially infinite number of states. This is known as the HDP-HMM and forms the basis of the NSSSM. In the HDP-HMM, each θ_k represents the emission parameters of state k, and each $\boldsymbol{\pi}_j$ represents a state-specific transition distribution that defines the probability of transitioning from state $z_{t-1} = j$ to state z_t. The global transition distribution $\boldsymbol{\beta}$ is the mean of each state-specific transition distribution ($\mathbb{E}[\boldsymbol{\pi}_j] = \boldsymbol{\beta}$), and the concentration parameter α governs how much each $\boldsymbol{\pi}_j$ varies around $\boldsymbol{\beta}$. To generate datapoint t, we first draw state $z_t \sim \boldsymbol{\pi}_{z_{t-1}}$ and transition from state z_{t-1} to state z_t and then draw observation $y_t \sim F(\theta_{z_t})$ conditioned on parameter θ_{z_t}.

We now describe an alternative representation of the hierarchical Dirichlet process known as the Chinese restaurant franchise (CRF), which will come in use when performing inference in the hierarchical Dirichlet process.

The Chinese restaurant franchise

Consider a Chinese restaurant franchise with J restaurants. Each restaurant has an infinite number of tables, and each table serves one dish from a global menu. Upon entering restaurant j, customer i sits at an occupied table with probability proportional to the number of people already sitting at that table or sits at a new table with probability proportional to α:

$$p(t_{ji}|t_{j1},\ldots,t_{ji-1},\alpha) \propto \sum_{t=1}^{m_{j\cdot}} \tilde{n}_{jt\cdot}\delta(t_{ji},t) + \alpha\delta(t_{ji},m_{j\cdot}+1), \tag{C.5}$$

where $\tilde{n}_{jtk}$ denotes the number of customers in restaurant j sitting at table t eating dish k, and m_{jk} denotes the number of occupied tables in restaurant j serving dish k. Marginal counts are represented with dots. For example, $\tilde{n}_{jt\cdot} = \sum_k \tilde{n}_{jtk}$ denotes the number of customers in restaurant j sitting at table t, and $m_{j\cdot} = \sum_k m_{jk}$ denotes the number of occupied tables in restaurant j.

The first customer to sit at a table chooses a dish for that table with probability proportional to the number of other tables in the franchise serving that dish or chooses a new dish with probability proportional to γ:

$$p(k_{jt}|\underline{k}_1,\ldots,\underline{k}_j,\gamma) \propto \sum_{k=1}^{K} m_{\cdot k}\delta(k_{jt},k) + \gamma\delta(k_{jt},K+1), \tag{C.6}$$

where $\underline{k}_j = \{k_{j1},\ldots,k_{jm_{j\cdot}}\}$ and K is the total number of unique dishes being served at occupied tables in the franchise.

In the HDP-HMM, each restaurant corresponds to a state $z_t = j$, and each dish corresponds to a parameter θ_k. The dish served at table t in restaurant j is distributed as $k_{jt} \mid \boldsymbol{\beta} \sim \boldsymbol{\beta}$, where $\boldsymbol{\beta}$ now represents the overall popularity or ratings of the dishes. Multiple tables in multiple restaurants can serve the same dish.

State persistence

The discrete state of the world often persists for many time steps before switching. This suggests that self-transitions are common than transitions between states. One shortcoming of the HDP-HMM is that it cannot differentiate self transitions from transitions between states, as the expected transition probabilities are independent of the previous state ($\mathbb{E}[\pi_{jk}] = \beta_k$). To address this issue, Fox et al. (2011b) developed the *sticky HDP-HMM*. In this generalisation of the HDP-HMM, the Dirichlet process prior over the state-specific transition distribution (equation C.4) becomes

$$\boldsymbol{\beta} \mid \gamma \sim \text{GEM}(\gamma) \qquad \boldsymbol{\pi}_j \mid \alpha, \kappa, \boldsymbol{\beta} \sim \text{DP}\left(\alpha + \kappa, \frac{\alpha\boldsymbol{\beta} + \kappa\delta_j}{\alpha + \kappa}\right). \tag{C.7}$$

Here a self-transition bias κ has been added to the jth component of $\alpha\boldsymbol{\beta}$. This increases the prior probability of self-transitions $\mathbb{E}[\pi_{jj}]$ by an amount proportional to κ. By setting $\kappa = 0$, the original HDP-HMM is recovered.

The sticky HDP-HMM can be understood by extending the metaphor of the Chinese restaurant franchise to include loyal customers. Each restaurant has a specialty dish whose index is the same as that of the restaurant (e.g., dish j is the speciality dish of restaurant j). This dish is available in all restaurants, but it is more popular in the dish's namesake restaurant. As we shall see, this leads to family loyalty to a restaurant. We refer to z_{t-1}, z_t and z_{t+1} as the grandparent, parent and child, respectively. The parent enters restaurant j, determined by the grandparent $z_{t-1} = j$, and sits at table $t_{ji} \sim \tilde{\boldsymbol{\pi}}_j$ that serves dish $k_{jt_{ji}}$. The dish the parent eats determines the restaurant that the child eats at, that is $k_{jt_{ji}} = z_t$. The increased popularity of the specialty dish means that children are more likely to eat at the same restaurant as their parent. Hence, multiple generations often eat at the same restaurant, demonstrating family loyalty

To simplify inference in the Chinese restaurant franchise with loyal customers, a distinction is made between *considered* dishes and *served* dishes. The first customer

to sit at a table chooses a dish for the table without acknowledging the increased popularity of the specialty dish of the restaurant. However, with some probability, this considered dish is overridden (perhaps by a waiter's suggestion) and the specialty dish is served instead. This process is described as follows:

$$\begin{aligned} \bar{k}_{jt} \mid \boldsymbol{\beta} &\sim \boldsymbol{\beta} \\ w_{jt} \mid \alpha, \kappa &\sim \text{Ber}\left(\frac{\kappa}{\alpha+\kappa}\right) \\ k_{jt} &= \begin{cases} \bar{k}_{jt} & \text{if } w_{jt} = 0 \\ j & \text{if } w_{jt} = 1, \end{cases} \end{aligned} \tag{C.8}$$

where $\bar{k}_{jt}$ is the considered dish, w_{jt} is an override variable, Ber is the Bernoulli distribution and k_{jt} is the served dish. When $\kappa = 0$, the considered dish is always served as $w_{jt} = 0$.

The increased popularity of the specialty dish leads to the following modified dish ratings:

$$k_{jt} \mid \alpha, \kappa, \boldsymbol{\beta} \sim \frac{\alpha\boldsymbol{\beta} + \kappa\delta_j}{\alpha+\kappa}. \tag{C.9}$$

C.2 The nonparametric switching state-space model

C.2.1 The generative model

We use the HDP-HMM as a prior on the context transition and cue emission matrices of a switching state-space model with an unknown, potentially infinite number of contexts and cues. For the context transition matrix, we use the sticky HDP-HMM to incorporate a self-transition bias. For the cue emission matrix, we use the standard HDP-HMM, as there is no notion of a self transition. We describe the model in its

most general form, in which all parameters of the linear dynamical systems are random variables to be inferred.

In the nonparametric switching state-space model, the environment is in one context c_t at any time t. The context evolves as a first-order hidden Markov process:

$$c_t \mid c_{t-1}, (\boldsymbol{\pi}_i)_{i=1}^{\infty} \sim \ \boldsymbol{\pi}_{c_{t-1}}, \tag{C.10}$$

where each context-specific transition distribution $\boldsymbol{\pi}_i$ is distributed as

$$\boldsymbol{\beta} \mid \gamma \sim \mathrm{GEM}(\gamma) \qquad \boldsymbol{\pi}_j \mid \alpha, \kappa, \boldsymbol{\beta} \ \sim \mathrm{DP}\left(\alpha + \kappa, \frac{\alpha\boldsymbol{\beta} + \kappa\delta_j}{\alpha + \kappa}\right). \tag{C.11}$$

The context can emit an observable cue q_t (a discrete symbol) according to emission probabilities given by

$$q_t \mid c_t, (\boldsymbol{\phi}_i)_{i=1}^{\infty} \sim \ \boldsymbol{\phi}_{c_t}, \tag{C.12}$$

where each context-specific emission distribution $\boldsymbol{\phi}_i$ is distributed as

$$\boldsymbol{\beta}' \mid \gamma' \sim \mathrm{GEM}(\gamma') \qquad \boldsymbol{\phi}_i \mid \alpha', \boldsymbol{\beta}' \sim \mathrm{DP}(\alpha', \boldsymbol{\beta}'). \tag{C.13}$$

Each context is associated with a latent perturbation that evolves independently of all other perturbations as a linear dynamical system:

$$x_t^j = a_j x_{t-1}^j + d_j + w_t^j \quad w_t^j \sim \mathcal{N}(0, s_j), \tag{C.14}$$

where x_t^j is the perturbation, a_j is the state transition coefficient, d_j is the state drift and s_j is the variance of the process noise, all associated with context j.

At each time step, the perturbation associated with the current context emits an observation (the state feedback):

$$y_t = x_t^{c_t} + v_t \quad v_t \sim \mathcal{N}(0, r_{c_t}), \tag{C.15}$$

where r_{c_t} is the variance of the measurement noise associated with the current context.

The parameters of the state transition function (equation C.14) are sampled from a normal-inverse-gamma distribution:

$$\boldsymbol{f}_j, s_j \mid \boldsymbol{\mu}_0, \boldsymbol{\Lambda}_0, \nu_0, \eta_0 \sim \mathcal{NIG}(\boldsymbol{\mu}_0, \boldsymbol{\Lambda}_0, \nu_0, \eta_0), \tag{C.16}$$

where $\boldsymbol{f}_j = [a_j \; d_j]^\intercal$, $\boldsymbol{\mu}_0 = [\mu_\mathrm{a} \; \mu_\mathrm{d}]^\intercal$, $\boldsymbol{\Lambda}_0$ is a 2×2 diagonal precision matrix, and ν_0 and η_0 are shape and scale parameters that can be interpreted as having arisen from $2\nu_0$ effective prior observations having variance η_0/ν_0. The parameter of the observation function (equation C.15) is sampled from a separate inverse-gamma distribution:

$$r_j \mid \epsilon_0, \tau_0 \sim \mathcal{IG}(\epsilon_0, \tau_0). \tag{C.17}$$

We examined two versions of the NSSSM that made different assumptions about the distribution $\mathcal{N}(\mu_0, \sigma_0^2)$ of a perturbation the first time it is encountered. In one version (employed in the spontaneous recovery simulations in Chapters 4 and 5), we assumed that the perturbation is distributed according to the initial distribution $\mathcal{N}(0, 0.5)$. In another version (employed in the context estimation simulation in Chapter 6), we assumed that the perturbation is distributed according to its stationary distribution $\mathcal{N}(d_j/(1-a_j), s_j/(1-a_j^2))$.

C.2.2 Inference

The goal of the learner is to use state feedback and contextual cues to infer the current context, the number of contexts, the context transition probabilities, the cue emission probabilities, the perturbations and the parameters of the state transition and observation functions associated with the perturbations. To perform inference, we use a Monte Carlo simulation strategy based on a resample-sampling framework known as particle learning (Carvalho et al., 2010). Central to particle learning is the

essential state vector:

$$Z_t = (\{c_{1:t}, C_t, Q_t, \{n_{ij,t}\}, \{n'_{jk,t}\}, \boldsymbol{\beta}_t, \boldsymbol{\beta}'_t, s^x_t, s^\theta_t, \theta\}^{(i)})_{i=1}^N,$$

which contains the sampled context sequence $c_{1:t}$, the number of instantiated contexts C_t, the number of observed cues Q_t, the number of transitions $\{n_{ij,t}\}$ from each context i to each context j, the number of emissions $\{n'_{jk,t}\}$ of each cue k in each context j, the global transition distribution $\boldsymbol{\beta}_t$, the global emission distribution $\boldsymbol{\beta}'_t$, the sufficient statistics (mean and variance) for the perturbations s^x_t, the sufficient statistics (equation C.36) for the parameters of the perturbations s^θ_t and the parameters of the perturbations θ.

The algorithm is initialised by setting $C_0 = 0, Q_0 = 0, \boldsymbol{\beta}_0 = 1, \boldsymbol{\beta}'_0 = 1$ for all particles.

Resample particles

For clarity, we suppress the indices of individual particle throughout most of the text. At time t, resample particles with weights proportional to the predictive distribution:

$$\begin{aligned} w_t &\propto p(y_t, q_t | c_{1:t-1}, q_{1:t-1}, s^x_{t-1}, \theta, \boldsymbol{\beta}'_{t-1}, \alpha', \boldsymbol{\beta}_{t-1}, \alpha, \kappa) \\ &= \sum_{c_t=1}^{C_{t-1}+1} p(y_t, q_t, c_t | c_{1:t-1}, q_{1:t-1}, s^x_{t-1}, \theta, \boldsymbol{\beta}'_{t-1}, \alpha', \boldsymbol{\beta}_{t-1}, \alpha, \kappa) \\ &= \sum_{c_t=1}^{C_{t-1}+1} p(y_t | c_t, s^x_{t-1}, \theta) p(q_t | c_{1:t}, q_{1:t-1}, \boldsymbol{\beta}'_{t-1}, \alpha') p(c_t | c_{1:t-1}, \boldsymbol{\beta}_{t-1}, \alpha, \kappa). \end{aligned} \tag{C.18}$$

The context transition term of the predictive distribution is obtained by integrating out the transition distribution $\boldsymbol{\pi}_{c_{t-1}}$ and can be written as

$$p(c_t | c_{1:t-1}, \boldsymbol{\beta}_{t-1}, \alpha, \kappa) = \frac{\alpha\beta_{c_t,t-1} + \kappa\delta(c_{t-1}, c_t) + n_{c_{t-1}c_t,t-1}}{\alpha + \kappa + n_{c_{t-1},t-1}}, \tag{C.19}$$

where $n_{c_{t-1}c_t,t-1}$ denotes the number of transitions from context c_{t-1} to context c_t up to time $t-1$. Marginal counts are represented with dots. For example, $n_{c_{t-1}\cdot,t-1} = \sum_{c_t} n_{c_{t-1}c_t,t-1}$ is the number of transitions out of context c_{t-1} up to time $t-1$.

The cue emission term is obtained by integrating out the emission distribution $\boldsymbol{\phi}_{c_t}$ and has a similar formulation to the context transition term but with no self-transition bias:

$$p(q_t|c_{1:t}, q_{1:t-1}, \boldsymbol{\beta}'_{t-1}, \alpha') = \frac{\alpha'\beta'_{q_t,t-1} + n'_{c_t q_t,t-1}}{\alpha' + n'_{c_t\cdot,t-1}}, \tag{C.20}$$

where $n'_{c_t q_t,t-1}$ denotes the number of emissions of cue q_t in context c_t up to time $t-1$, and $n'_{c_t\cdot,t-1} = \sum_{q_t} n'_{c_t q_t,t-1}$ is the number of cues emitted in context c_t up to time $t-1$.

The state feedback term depends on the estimate of the perturbation and is given by

$$p(y_t|c_t, s^x_{t-1}, \theta) = \mathcal{N}(y_t|\hat{y}^{c_t}_t, p^{c_t}_t), \tag{C.21}$$

where $\hat{y}^{c_t}_t$ and $p^{c_t}_t$ are the mean and variance of the predicted observation for the perturbation associated with context c_t (see Algorithm 1). For a new context (i.e., $c_t = C_{t-1}+1$), $\hat{y}^{c_t}_t = \mu_0$ and $p^{c_t}_t = \sigma^2_0 + r_{c_t}$, where r_{c_t} is sampled from the prior (equation C.17).

Propagate the context

To propagate the context, sample $c_t \in \{1, \ldots, C_{t-1}+1\}$ from the posterior:

$$p(c_t|y_t, q_t, c_{1:t-1}, q_{1:t-1}, s^x_{t-1}, \theta, \boldsymbol{\beta}'_{t-1}, \alpha', \boldsymbol{\beta}_{t-1}, \alpha, \kappa). \tag{C.22}$$

Update the number of instantiated contexts and observed cues:

$$C_t = \begin{cases} C_{t-1} + 1 & \text{if } c_t = C_{t-1} + 1 \\ C_{t-1} & \text{otherwise} \end{cases} \qquad Q_t = \begin{cases} Q_{t-1} + 1 & \text{if } q_t = Q_{t-1} + 1 \\ Q_{t-1} & \text{otherwise.} \end{cases} \tag{C.23}$$

Update the context transition and cue emission counts:

$$n_{ij,t} = \begin{cases} n_{ij,t-1} + 1 & \text{if } c_{t-1} = i, c_t = j \\ n_{ij,t-1} & \text{otherwise} \end{cases} \qquad n'_{jk,t} = \begin{cases} n'_{jk,t-1} + 1 & \text{if } c_t = j, q_t = k \\ n'_{jk,t-1} & \text{otherwise.} \end{cases} \tag{C.24}$$

If c_t is a new context, sample $b \sim \text{Beta}(1, \gamma)$ and update the global transition distribution:

$$\beta_{j,t} = \begin{cases} \beta_{j,t-1} & \text{if } j < C_t \\ b\beta_{j,t-1} & \text{if } j = C_t \\ (1-b)\beta_{j,t-1} & \text{if } j = C_t + 1. \end{cases} \tag{C.25}$$

Otherwise, set $\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1}$.

If q_t is a new cue, sample $b' \sim \text{Beta}(1, \gamma')$ and update the global emission distribution:

$$\beta'_{k,t} = \begin{cases} \beta'_{k,t-1} & \text{if } k < Q_t \\ b'\beta'_{k,t-1} & \text{if } k = Q_t \\ (1-b')\beta'_{k,t-1} & \text{if } k = Q_t + 1. \end{cases} \tag{C.26}$$

Otherwise, set $\boldsymbol{\beta}'_t = \boldsymbol{\beta}'_{t-1}$.

Propagate the sufficient statistics for the perturbation

Update the sufficient statistics for the perturbation (mean and variance) using the Kalman filter equations (Algorithm 1).

Algorithm 1 The recursive update equations $\mathcal{K}(\cdot)$ for the sufficient statistics for the continuous state

Time update
if $j \leq C_{t-1}$ **then**
$\hat{x}^j_{t|t-1} = a_j \hat{x}^j_{t-1|t-1} + d_j$
$v^j_{t|t-1} = a_j v^j_{t-1|t-1} a_j + q_j$
else if $j = C_{t-1} + 1$ **then**
$\hat{x}^j_{t|t-1} = \mu_0$
$v^j_{t|t-1} = \sigma_0^2$
end if
$\hat{y}^j_t = \hat{x}^j_{t|t-1}$
$p^j_t = v^j_{t|t-1} + r_j$
Measurement update
if $c_t = j$ **then**
$k^j_t = v^j_{t|t-1} / p^j_t$
$\hat{x}^j_{t|t} = \hat{x}^j_{t|t-1} + k_t(y_t - \hat{y}^j_t)$
$v^j_{t|t} = (1 - k^j_t) v^j_{t|t-1}$
else
$\hat{x}^j_{t|t} = \hat{x}^j_{t|t-1}$
$v^j_{t|t} = v^j_{t|t-1}$
end if

Sample the global emission distribution

The posterior distribution of the global emission distribution is

$$(\beta'_1, \ldots, \beta'_{Q_t}, \beta'_{\tilde{q}}) \sim \mathrm{Dir}(m'_{\cdot 1}, \ldots, m'_{\cdot Q_t}, \gamma'), \tag{C.27}$$

where $\beta'_{\tilde{q}} = \sum_{k=Q_t+1}^{\infty} \beta'_k$. To sample $\boldsymbol{\beta}'$, first simulate a Chinese restaurant process to sample each m'_{jk}, which represents the number of tables in restaurant j serving dish k. For each $j \in \{1, \ldots, C_t\}$ and $k \in \{1, \ldots, Q_t\}$, set $m'_{jk} = 0$ and $n' = 0$. Then, for $i = 1, \ldots, n'_{jk}$ (i.e., for each customer in restaurant j eating dish k), sample

$$x \sim \mathrm{Ber}\left(\frac{\alpha' \beta'_k}{n + \alpha' \beta'_k}\right), \tag{C.28}$$

increment n', and if $x = 1$ increment m'_{jk}.

Conditioned on the number of tables serving dish k, $m'_{\cdot k}$, sample the global emission distribution

$$(\beta'_1, \ldots, \beta'_{Q_t}, \beta'_{\tilde{q}}) \sim \mathrm{Dir}(m'_{\cdot 1}, \ldots, m'_{\cdot Q_t}, \gamma'), \tag{C.29}$$

Sample the global transition distribution

The posterior distribution of the global transition distribution is

$$(\beta_1, \ldots, \beta_{C_t}, \beta_{\tilde{c}}) \sim \mathrm{Dir}(\bar{m}_{\cdot 1}, \ldots, \bar{m}_{\cdot C_t}, \gamma), \tag{C.30}$$

where $\beta_{\tilde{c}} = \sum_{j=C_t+1}^{\infty} \beta_j$. To sample $\boldsymbol{\beta}$, first simulate a Chinese restaurant process to sample m_{ij}, which represents the number of tables in restaurant i *serving* dish j. For each $i, j \in \{1, \ldots, C_t\}^2$, set $m_{ij} = 0$ and $n = 0$. Then, for $i = 1, \ldots, n_{ij}$ (i.e., for each customer in restaurant i eating dish j), sample

$$x \sim \mathrm{Ber}\left(\frac{\alpha\beta_j + \kappa\delta(i,j)}{n + \alpha\beta_j + \kappa\delta(i,j)}\right), \tag{C.31}$$

increment n, and if $x = 1$ increment m_{ij}.

To obtain $\bar{m}_{ij}$, which represents the number of tables in restaurant i *considering* dish j, subtract the number of override variables in restaurant i from m_{ii}. For each $i \in \{1, \ldots, C_t\}$, sample the number of override variables in restaurant i

$$w_{i\cdot} \sim \mathrm{Binomial}\left(m_{ii}, \frac{\rho}{\rho + \beta_i(1-\rho)}\right), \tag{C.32}$$

where $\rho = \kappa/(\alpha + \kappa)$, and set the number of tables in restaurant i considering dish j to

$$\bar{m}_{ij} = \begin{cases} m_{ii} - w_{i\cdot} & \text{if } i = j \\ m_{ij} & \text{otherwise.} \end{cases} \tag{C.33}$$

Conditioned on the number of tables considering dish j, $\bar{m}_{\cdot j}$, sample the global transition distribution

$$(\beta_1, \ldots, \beta_{C_t}, \beta_{\tilde{c}}) \sim \mathrm{Dir}(\bar{m}_{\cdot 1}, \ldots, \bar{m}_{\cdot C_t}, \gamma), \tag{C.34}$$

Propagate the sufficient statistics for the parameters

To propagate the sufficient statistics for the parameters, for each $j \in \{1, \ldots, C_t\}$, first sample x_{t-1}^j and x_t^j from

$$p(x_{t-1}^j, x_t^j | c_t, s_{t-1}^x, \theta, y_t) = p(x_t^j | x_{t-1}^j, c_t, \theta, y_t) p(x_{t-1}^j | c_t, s_{t-1}^x, \theta, y_t). \tag{C.35}$$

This can be achieved by sampling x_{t-1}^j from $p(x_{t-1}^j | c_t, s_{t-1}^x, \theta, y_t)$ and then sampling x_t^j from $p(x_t^j | x_{t-1}^j, c_t, \theta, y_t)$. The sampled values of x_{t-1}^j, x_t^j and c_t are used with the observation y_t to update the sufficient statistics for the parameters:

$$\begin{aligned}
s_{j,1,t} &= s_{j,1,t-1} + x_t^j \bar{\boldsymbol{x}}_{t-1}^j \\
s_{j,2,t} &= s_{j,2,t-1} + \bar{\boldsymbol{x}}_{t-1}^j \bar{\boldsymbol{x}}_{t-1}^{j\intercal} \\
s_{j,3,t} &= s_{j,3,t-1} + 1 \\
s_{j,4,t} &= s_{j,4,t-1} + x_t^{j2} \\
s_{j,5,t} &= s_{j,5,t-1} + \mathbb{1}(c_t = j) \\
s_{j,6,t} &= s_{j,6,t-1} + \mathbb{1}(c_t = j)(y_t - x_t^j)^2,
\end{aligned} \tag{C.36}$$

where $\bar{\boldsymbol{x}}_{t-1}^j = [x_{t-1}^j \; 1]^\intercal$ is an augmented state vector and $\mathbb{1}(\cdot)$ is the indicator function.

Sample the parameters

Conditioned on the updated sufficient statistics for the parameters, the posterior distributions at the end of trial t are

$$\boldsymbol{f}_j, s_j \mid \boldsymbol{\mu}_{j,t}, \boldsymbol{\Lambda}_{j,t}, \nu_{j,t}, \eta_{j,t} \sim \mathcal{NIG}(\boldsymbol{\mu}_{j,t}, \boldsymbol{\Lambda}_{j,t}, \nu_{j,t}, \eta_{j,t}) \qquad r \mid \epsilon_{j,t}, \tau_{j,t} \sim \mathcal{IG}(\epsilon_{j,t}, \tau_{j,t}), \tag{C.37}$$

where

$$\begin{aligned} \boldsymbol{\mu}_{j,t} &= \boldsymbol{\Lambda}_{j,t}^{-1}(\boldsymbol{\Lambda}_0\boldsymbol{\mu}_0 + s_{j,1,t}) & \epsilon_{j,t} &= \epsilon_0 + \frac{s_{j,5,t}}{2} \\ \boldsymbol{\Lambda}_{j,t} &= (\boldsymbol{\Lambda}_0 + s_{j,2,t}) & \tau_{j,t} &= \tau_0 + \frac{s_{j,6,t}}{2}. \\ \nu_{j,t} &= \nu_0 + \frac{s_{j,3,t}}{2} \\ \eta_{j,t} &= \eta_0 + \frac{1}{2}[\boldsymbol{\mu}_0^\intercal\boldsymbol{\Lambda}_0\boldsymbol{\mu}_0 + s_{j,4,t} - \boldsymbol{\mu}_{j,t}^\intercal\boldsymbol{\Lambda}_{j,t}\boldsymbol{\mu}_{j,t}] \end{aligned} \tag{C.38}$$

At the end of each trial, a new set of parameters is sampled from the updated posteriors.

The motor output

The motor output is modeled as the mean of the predictive distribution, which is calculated by marginalising out the context and summing over particles:

$$u_t = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C_{t-1}+1} \tilde{w}_t^{j(i)}\hat{y}_t^{j(i)}. \tag{C.39}$$

For particle i, the contribution of context j to the motor output is given by the prior probability of context j:

$$\tilde{w}_t^j = \frac{p(q_t|c_t = j, c_{1:t-1}, q_{1:t-1}, \boldsymbol{\beta}'_{t-1}, \alpha')p(c_t = j|c_{1:t-1}, \boldsymbol{\beta}_{t-1}, \alpha, \kappa)}{\sum_{c_t=j'}^{C_{t-1}+1} p(q_t|c_t = j', c_{1:t-1}, q_{1:t-1}, \boldsymbol{\beta}'_{t-1}, \alpha')p(j' = 1|c_{1:t-1}, \boldsymbol{\beta}_{t-1}, \alpha, \kappa)}. \tag{C.40}$$

On channel trials, we assume that participants observe a fictitious observation $y_t = u_t$.

Model implementation

For all simulations, we set γ and γ' to 0.1. We used $N = 10^3$ particles.

Although the number of contexts that the model can instantiate is unbounded, in practice, we limit the number to ten by truncating the stick-breaking process (equations C.11 and C.25).