Imperial College London
Department of Bioengineering

Principles of sensorimotor control and learning in complex motor tasks

Anastasia Sylaidi

Submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy in Bioengineering of Imperial College London
and the Diploma of Imperial College London

Declaration

I herewith certify that all material in this dissertation which is not my own work has been properly acknowledged.

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Anastasia Sylaidi

Abstract

The brain coordinates a continuous coupling between perception and action in the presence of uncertainty and incomplete knowledge about the world. This mapping is enabled by control policies and motor learning can be perceived as the update of such policies on the basis of improving performance given some task objectives. Despite substantial progress in computational sensorimotor control and empirical approaches to motor adaptation, to date it remains unclear how the brain learns motor control policies while updating its internal model of the world.

In light of this challenge, we propose here a computational framework, which employs error-based learning and exploits the brain's inherent link between forward models and feedback control to compute dynamically updated policies. The framework merges optimal feedback control (OFC) policy learning with a steady system identification of task dynamics so as to explain behavior in complex object manipulation tasks. Its formalization encompasses our empirical findings that action is learned and generalised both with regard to a body-based and an object-based frame of reference. Importantly, our approach predicts successfully how the brain makes continuous decisions for the generation of complex trajectories in an experimental paradigm of unfamiliar task conditions. A complementary method proposes an expansion of the motor learning perspective at the level of policy optimisation to the level of policy exploration. It employs computational analysis to reverse engineer and subsequently assess the control process in a whole body manipulation paradigm.

Another contribution of this thesis is to associate motor psychophysics and computational motor control to their underlying neural foundation; a link which calls for further advancement in motor

neuroscience and can inform our theoretical insight to sensorimotor processes in a context of physiological constraints. To this end, we design, build and test an fMRI-compatible haptic object manipulation system to relate closed-loop motor control studies to neurophysiology. The system is clinically adjusted and employed to host a naturalistic object manipulation paradigm on healthy human subjects and Friedreich's ataxia patients. We present methodology that elicits neuroimaging correlates of sensorimotor control and learning and extracts longitudinal neurobehavioral markers of disease progression (i.e. neurodegeneration).

Our findings enhance the understanding of sensorimotor control and learning mechanisms that underlie complex motor tasks. They furthermore provide a unified methodological platform to bridge the divide between behavior, computation and neural implementation with promising clinical and technological implications (e.g. diagnostics, robotics, BMI).

Acknowledgements

I would first like to thank my thesis advisor, Aldo Faisal, for giving me the chance to work on this project and encouraging my research. His support provided guidance but also left room for my own intuition and allowed me to grow as a scientist. I feel particularly privileged to have been part of the stimulating environment and vision within his lab, which have provided the ideal balance between pioneering thinking and anchoring to reality.

I would like to thank Reiko Tanaka for her invaluable mentoring, support and encouragement throughout the course of my PhD. I am also grateful to Richard Festenstein for his vital insights and our very useful interaction within the clinical context of my work. Within the same context I am very happy and thankful to have collaborated with Sathjii Nageshwaran, Albert Busza, Julie Fitzpatrick and Stavros Athanasopoulos who have always been greatly helpful. My special thanks go to Pedro Lourenco for his unique collaborative spirit, his professionalism, his kindness and his always valuable insights and help that made our interaction very fruitful.

I would like to thank the Bioengineering Department of Imperial College London for supporting financially the three first years of my PhD with the Colin Caro Fellowship. I am also deeply thankful to the Foundation for Education and European Culture for its crucial financial support during the fourth and last year of my PhD. Thanks to this support I was able to successfully complete my research project.

Very special thanks go to my colleagues (present and past) at the Brain and Behavior Lab: Ali Neishabouri, Alessandro Ticchi, Andreas Thomik, Ariadne Whitby, Chin-Hsuan Lin, Constantinos Gavriel, Ekaterina Abramova, Feryal Behbahani, Ira Ktena, Luke Dickens, Margarita Kotti, Marta Garnelo, Matthieu Komorowski, Raj Iyer, Romy Lorenz, Scott Taylor, William Abbott for our inspiring discussions, for their support and professional ethics which made this PhD a very rewarding experience.

I would additionally like to thank the following colleagues beyond my narrow academic affiliation

for their help, advice, insightful contributions and willingness to share knowledge: Alessandro Allievi, Arichi Tomoki, Spyros Masouros, Eleni Bazigou.

Furthermore, I would like to thank Irini Spyropoulou, Tina Grigorakou, Maria Arvanitaki, Zoi Arvanitaki, Yeshna Yildiz, Christina Nikita, Zafeirios Fountas, Zacharias Fotos, Nikitas Thomareis, Samantha Matsa, Chryssa Thoua, Greg Wayne, Tassos Stasinopoulos, Stelios Giamarelos, Christos Bergeles, George Fagogenis, Konstantinos Trantopoulos, Nikos Petropoulos who independently of geographical and professional borders have always been a source of care, enthusiasm and positive reinforcement. I am particularly thankful to Konstantinos Moutoussis for our mind shaping interaction, his invaluable advice and sincere, generous encouragement throughout this journey. I would also like to thank Mina Bouras, who has contributed substantial and dedicated support tightly linked to the very initiation and completion of this PhD, amongst other professional and personal endeavours.

My special gratitude goes to Aristides Baltas who has been a mentor and inspiration. Our insightful interactions and his belief in my intellectual instincts have been great gifts for my path at a critical stage of its shaping.

Last but foremost, I would like to deeply thank my family for its unparalleled contribution to all my quests, including the completion of this PhD. Their love and support, but also their values and personal paths are the greatest legacy and reference in my life. I would individually like to thank my parents, Vasilis and Katerina, my siblings, Alexandros and Chrisi, my grandmothers Tasia Sylaidi and Chrysoula Makri, and my uncle and aunt, Ioannis and Eleni Karatzas. Finally, I would like to thank my grandfather Georgios Makris, whose love and memory keep him as my everlasting companion and inspiration.

Geschrieben steht: »Im Anfang war das Wort!«
Hier stock ich schon! Wer hilft mir weiter fort?
Ich kann das Wort so hoch unmöglich schätzen,
Ich muß es anders übersetzen,
Wenn ich vom Geiste recht erleuchtet bin.
Geschrieben steht: Im Anfang war der Sinn.
Bedenke wohl die erste Zeile,
Daß deine Feder sich nicht übereile!
Ist es der Sinn, der alles wirkt und schafft?
Es sollte stehn: Im Anfang war die Kraft!
Doch, auch indem ich dieses niederschreibe,
Schon warnt mich was, daß ich dabei nicht bleibe.
Mir hilft der Geist! Auf einmal seh ich Rat
Und schreibe getrost: Im Anfang war die Tat!

J.W.GOETHE, *Faust*

Contents

List of Figures

Acronyms

BIC Bayesian Information Criterion

BOLD Blood Oxygenated Level Dependent

CNS Central Nervous System

EPE End Point Error

EV Explanatory Variable

FE Fixed-Effects

fMRI Functional Magnetic Resonance Imaging

FRDA Friedreich's Ataxia

GLM General Linear Model

HRF Hemodynamic Response Function

IAM Ideal Actor Model

IDOF Internal Degrees of Freedom

IFM Internal Forward Model

LDSI Linear Dynamical System with Inputs

LQR Linear Quadratic Regulator

MCMC Markov Chain Monte Carlo

ME Mixed-Effects

NM Naive Model

OFC Optimal Feedback Control

PLM Policy Learning Model

RE Random-Effects

RL Reinforcement Learning

RMSE Root-mean-squared-error

ROI Region of Interest

RT Reaction Time

SARA Scale for the Assessment and Rating of Ataxia

SE Standard Error

VR Virtual Reality

1 Introduction

At the 1976 Summer Olympics in Montreal, Nadia Comaneci was the first gymnast in Olympic history to be awarded the perfect score of 10 for her performance on the uneven bars. Her routine manifested what can be perceived as a perfectly balanced use of physical strength, flexibility, agility and motor coordination. Such level of motor skills involves hundreds of muscles and a well-orchestrated control of sequential acts, which can be affected even by the slightest variability of preceding movements. More importantly, performance in such contexts is the outcome of a remarkable learning process, which expands over long, systematic training and which progressively optimises and automatises motor execution.

Skills cultivated and demonstrated by accomplished athletes and other performers, such as professional dancers or musicians may represent the acme of learning (Globerson and Nelken 2013). However, even in the case of trivial daily activities, humans generally display a striking ability to learn a wide range of complex motor skills and generalise them to novel task conditions.

Broadly speaking, learning is the acquisition and integration of information, which is not contained in the prespecification of a system. Biological systems exhibit a trade-off between prespecified traits and learning (Roth 2007). For instance, living systems operate commonly based on hardwired low-level reflexes and high-level instincts, but the most intelligent and adaptive ones (e.g. mammal brains) develop many of their behavioural features through learning while interacting with their environment. This interaction stores experience about the world in neural representations, which based on widely accepted views (Wolpert et al. 2011; Krakauer and Mazzoni 2011) can subsequently be used to propel decisions upon actions and make predictions of their consequences.

This mechanism is consistent with a Bayesian brain hypothesis that formulates the interplay between perception and action as a constructive process driven by internal or generative models of the task and the world (Knill and Pouget 2004; Körding and Wolpert 2004; Pouget et al. 2013; Mamassian et al. 2002; Berkes et al. 2011). Historically, it can be traced back to the 19th century proto-Bayesian propositions of Helmholtz on active object recognition guided by previously formed visual representations (Helmholtz 1866). Conceptually, this notion bares an even older and tenuous association to Kant's views on perception as an understanding of the

external world founded both on experience and a priori concepts (Kant 1781).

Crucially, contemporary perspectives of perception and action in neuroscience, which posit the existence of internal models of the world, usually assume that the world encompasses the body (Wolpert et al. 1995; Kawato and Wolpert 1998; Vaziri et al. 2006). The empirical validation of this scenario (Vaziri et al. 2006; Berniker and Kording 2008) suggests that learning is influenced by aspects of an agent's body state, beyond the physiological constraints and prespecifications of the brain itself. This approach presents similarities to rising trends in artificial intelligence and specifically to the theory of embodied cognition that has deviated from previous highly symbolic, context-dependent and inefficient algorithms and resorted to neuro-inspired, generic computational systems with better sense organs, which can solve and learn from problems of perception and locomotion directly (Krichmar and Wagatsuma 2011).

1.1 An introduction to the sensorimotor system

Elucidating motor learning requires an understanding of the sensorimotor system. Sensorimotor functions are realized in a neuronal, kinematic and behavioural context. They reflect a complex interaction between the biomechanical properties of the body and distributed control by the circuits of the central nervous system (CNS). The CNS has a hierarchical organization, which spans over three levels: the cortex (highest), the brainstem and the spinal cord (lowest) (Scott 2004). The spinal cord is responsible for a basic stereotypical and automatised motor repertoire comprising mainly of multi-joint patterns. As the mid-level, the brainstem operates on the spinal repertoire and enhances it to better control posture. Finally, the cortex constitutes a network of multiple areas linked to sensory processing, motor planning and control. It thereby underlies the largest and most adaptive portion of movement patterns (Kandel et al. 2000). This thesis mainly focuses on examining mechanisms that underlie complex voluntary tasks and therefore constrains its interest to the processes supported by cortical regions.

Cortical function is generally viewed as a continuous coupling between perception and action in the presence of uncertainty and incomplete knowledge about the world (Faisal and Wolpert 2009; Faisal et al. 2008; Orbán and Wolpert 2011). This mapping is enabled by control policies and motor learning can be perceived as the update of such policies on the basis of improving performance given some task objectives (Izawa et al. 2008; Shadmehr and Krakauer 2008; Wolpert et al. 2011). The primary motor cortex (M1) is widely considered to play a key role in this theory by driving volitional motor control in goal-directed tasks (Scott 2004). Furthermore, the cerebellum has also been systematically linked to principles of motor learning (Houk et al. 1996; Albus 1971; Marr 1969; Mazzoni and Krakauer 2006; Taylor et al. 2010; Rabe et al. 2009; Galea et al. 2011), primarily as the locus of internal and inverse models of the motor apparatus

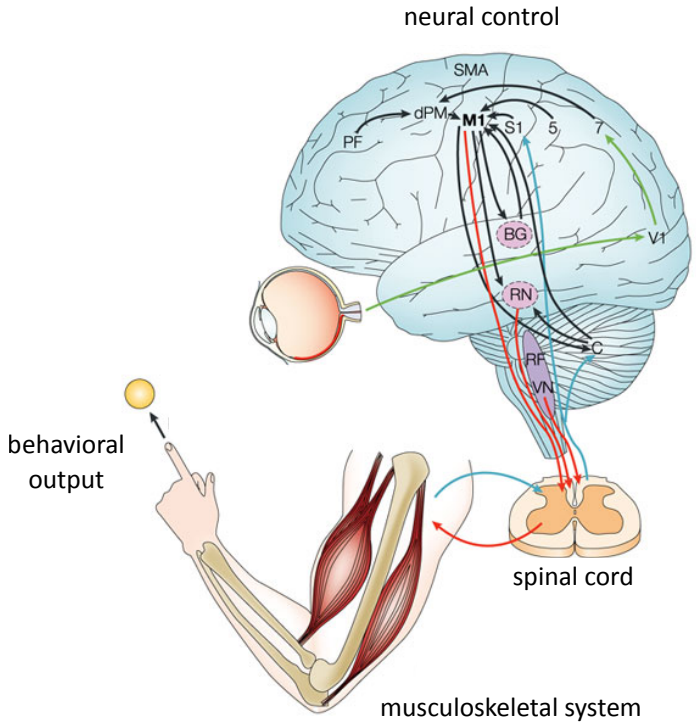

Figure 1.1: Overview of the sensorimotor system (adapted from Scott 2004): The highest level of motor control is supported by the cerebral cortex, usually based on the interrelated contribution of the primary motor cortex (M1), the primary somatosensory cortex (S1), parietal cortex area 5 (5), the basal ganglia (BG), the cerebellum (C), cerebellar pathways, the red nucleus (RN), the primary visual cortex (V1), region 7 of the posterior parieral cortex (7), the dorsal premotor cortex (dPM), supplementary motor areas (SMA) and the prefrontal cortex (PF). The middle level comprises mainly brainstem regions like the reticular formation (RF) and vestibular nuclei (VN) and the lowest level is the spinal cord which directly regulates the behavioural repertoire produced by the musculoskeletal system.

(Wolpert et al. 1998; Kawato et al. 2003; Albus 1971; Marr 1969; Kawato et al. 1987; Jordan and Rumelhart 1992). Both these areas and their relationship to other brain regions have been the focal point of systematic research; however there are many questions that remain unanswered. Does M1 control motion by coding low-level features of the musculoskeletal system or high-level features related to task goals? What is the purpose of control policies? Importantly, *what are the mechanisms that underlie the revision of control policies during learning*?

1.2 Principles of sensorimotor control and learning

In order to address these elusive topics in sensorimotor control and learning it is meaningful to structure their content into three basic divisions (Wolpert et al. 2011): (i) the neural representa-

tions which determine how a task is learned and generalised to new conditions (Krakauer et al. 2006; Ghahramani et al. 1996), (ii) the selection of motor strategies and (iii) the learning mechanism itself, which refers to the way in which the brain exploits its internal task representations and a 'learning signal' to update performance.

1.2.1 Action representations

The question of action representation translates to how the brain captures information about the body and environment within a given task context. This defines a mapping between sensory and motor variables. Former work in neuroscience has suggested that action can be planned in at least two possible reference frames (Shadmehr and Mussa-Ivaldi 1994; Diedrichsen 2007; Diedrichsen et al. 2010; Ingram et al. 2010): an intrinsic coordinate system centered on the body's actuators and sensors, and an extrinsic coordinate system related to task conditions, environmental settings and/or object properties. Although evidence has been found for the existence of both frames (Shadmehr and Mussa-Ivaldi 1994; Malfait et al. 2002, 2005; Shadmehr and Moussavi 2000; Burgess et al. 2007; Criscimagna-Hemminger et al. 2003; Krakauer et al. 2000; Ahmed et al. 2008a; Brayanov et al. 2012; Berniker et al. 2014), there is still no unified theory for internal representations. Moreover, the formation of motor memory remains elusive particularly within the context of complex naturalistic motor behaviour (e.g. manipulation of physical objects, Ingram and Wolpert 2011). This means that there is an enduring challenge to study this topic beyond un-natural lab settings and the simple hand reaching tasks commonly examined, so as to ensure that task conditions are interchangeable with respect to general insights. Here, we attempt to address this challenge with experimental paradigms that introduce life-like manipulations in the instructed task objectives.

1.2.2 Selection of motor strategies

The brain employs motor strategies to deal with the inferred structure (representation) of a task (Vaziri et al. 2006; Gold and Shadlen 2007). These strategies determine the form of the desired output. They also represent decisions which rely on the extraction of information from the environment, which is argued to also constitute an active sampling process of sensory input with the aim of minimising prediction errors (Friston et al. 2011).

In order to optimise performance for a given task structure, the brain needs to select efficient motor control policies that are meant to satisfy the perceived task objectives in the face of redundancy. This means that the brain is faced with more degrees of freedom than constraints in solving a motor control problem, which allows different strategies to achieve completion of the same task. A possible path to derive a solution to such redundant control problems is provided

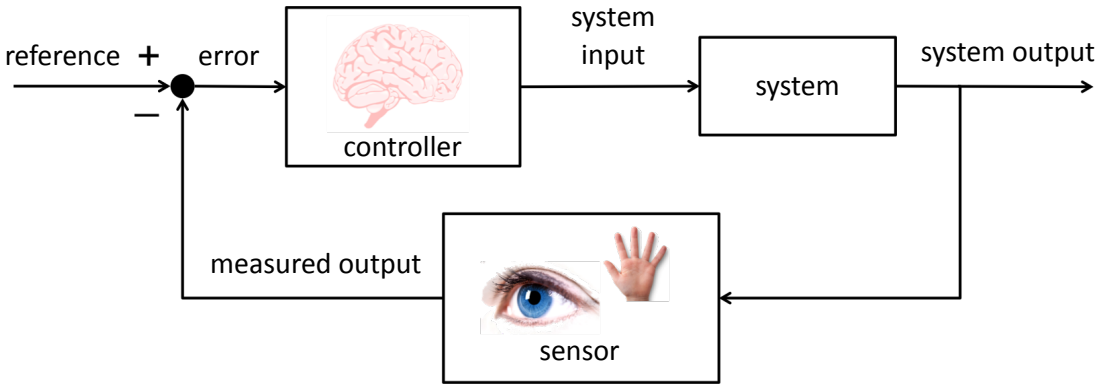

Figure 1.2: Capturing motor control in an OFC loop. Based on the representation of a task, the controller (brain) interacts with the affected system (e.g. body, object) via a control signal. As a consequence, the system's state changes and this update is measured by the brain's sensory system. The perceived new system state is compared against the brain's predictions on action consequences and the difference thereof (measured error) is fed into the controller to further drive the interaction with the environment in an optimal fashion.

by optimal control theory. In fact, the biological relevance of optimal feedback control (OFC) (Fig. 1.2) becomes particularly evident when treating the sensorimotor system as a product of evolution, development and adaptation, all of which are effectively optimisation processes striving to upgrade performance. Furthermore, policy selection has been examined by a growing body of empirical work, which showed that motor behaviour and particularly the formation of control signals is consistent with OFCT (Todorov and Jordan 2002; Todorov 2004; Scott 2004; Diedrichsen 2007). Therefore, in the present thesis we hypothesise that OFCT is a valid approach to explain the emergence of control signals within a motor learning framework.

1.2.3 The mechanism of motor learning

The description of the motor learning process depends on the type of signal that the brain uses to update behavioural outcomes. We review here two main categories of such signals that have been extensively examined in the past.

Error-based learning This process exploits the difference between the predicted consequences of action and the actual motor output. The estimation of this difference has been proposed as the driving force and a metric of adaptation in numerous studies (Pélisson et al. 2010; Shadmehr and Mussa-Ivaldi 1994; Flanagan and Wing 1997; Donchin et al. 2003; Diedrichsen et al. 2005). In the computational approach to motor learning here we assume that the brain exploits a

gradient of the error along which it performs updates on the internal action representations.

Reinforcement learning This constitutes typically a reward-based process (Sutton and Barto 1998) which in contrast to error-based learning is unsigned and hence non-suggestive of the direction of behavioural change. The mere goal of a reinforcement learning agent is to gather as much reward as possible based on the history of its actions and potentially also some stochastic action selection criteria. Reinforcement learning is thus better suited to describe paradigms with complex sequence of actions where the ultimate behavioural reward is distanced from the reward of the individual actions. In these cases sequential decisions can be easier made based on the history of experienced states and actions instead of an error describing the cumulative motor output. Given that the experimental paradigms in this thesis do not instruct such highly structured tasks, we do not employ RL for the formalisation of our motor learning hypothesis. However, we argue that a future generalisation of our approach can accommodate the fusion of our low-level error-based description of learning (for motor execution optimisation) and a high-level RL description of learning (responsible for the selection of control policy combinations which can subsequently be optimised at the low-level).

1.2.4 The challenge of expanding existing motor learning frameworks

Despite the substantial interest that motor learning has attracted within the past two decades, its formalisation and examination has still not been proposed in an inclusive manner which encompasses simultaneously the two major learning components presented above: a progressive update of internal action representations and a revision of control policies. The key point that remains unclear is *how the brain learns control policies while it upgrades its internal model of the world*.

In fact, most studies employing error-based accounts of motor performance updates confine the study of learning to the gradual adaptation of task kinematics and dynamics parameters without elucidating the parallel progression of motor strategies. In some cases this type of adaptation is driven by the assumption that the motor system's goal is to produce a specific motor output, e.g. straight, minimum jerk trajectories (Berniker and Kording 2008).

Other empirical or theoretical studies, that actually integrate the aspect of optimal control policies in the investigation of complex motor tasks, make use of models that assume full knowledge of task and world dynamics (i.e. fixed action representations). Such models essentially predict motor performance at the end-point of learning and thereby sidestep a full interpretation of motor learning (Nagengast et al. 2009; Mordatch et al. n.d.).

1.3 From computational models and motor psychophysics to the neural foundation of behaviour

Although in recent years there has been significant progress in the study of sensorimotor functions through various methodological approaches, less advancement has been achieved in associating motor psychophysics and computational motor control to their underlying neural foundation. This pending association appears as a necessary step towards tuning any theory or behaviour driven hypotheses on sensorimotor control and learning according to physiological constraints.

A growing number of studies have attempted to address this challenge with the use of neuroimaging, and fMRI in particular, which they employed during either very simple hand reaching movements or non-specific open-loop manipulations (e.g. finger tapping). The main reason for this restriction in paradigm design lies in the technical limitations embedded in fMRI technology, which is often incompatible to advanced motion tracking systems (mainly due to electromagnetic interference and safety issues), that could monitor more complex motor behaviour.

In light of this background, the vision of bridging the divide across behaviour and neurophysiology features four main prerequisites which we seek to address here: (i) The development of alternative advanced motion tracking systems which are compatible to fMRI use. (ii) The development of experimental platforms that can accommodate closed-loop motor control studies on complex, goal-directed and feedback-driven tasks. (iii) The extraction of high-resolution behavioural metrics that capture the transient dynamics of motor learning and the development of efficient techniques to regress these metrics against their concurrent neural activity. (iv) The design and introduction of new experimental paradigms, which expand previous lines of research in motor neuroscience to more naturalistic task contexts (Ingram and Wolpert 2011).

1.3.1 Neurobehavioural correlates of sensorimotor processes: a clinical and technological challenge?

Apart from its intellectual significance in understanding the underpinnings of complex motor behaviour, eliciting neurobehavioural correlates of sensorimotor control and learning also presents significant clinical relevance. In fact, the examination of many neurodegenerative diseases and neurological disorders necessitates a grasp of both the sensorimotor system in baseline healthy cases as well as the nature and effects of its perturbation in pathological cases. Specifically in the context of neurodegeneration, conventional diagnostic tools often rely on subjective evaluations of clinical staff, which may perpetuate an inherent bias. This reflects a clear need for the development of multimodal biomarkers (e.g. neurobehavioural metrics of complex motor behaviour), that will be capable of objectively and efficiently diagnosing, monitoring and/or even predicting

disease progression (Skovronsky et al. 2006; Gotovac et al. 2014).

Questions accompanying this motivation are: How can we use empirical and computational approaches to drive the design of markers that can collaboratively capture the subtle and noisy behavioural and neurophysiological changes manifested in the course of neurodegeneration? How can we associate these markers to clinically established diagnostic tools? Finally, how can these markers be used to support diagnosis and monitoring on an individualised patient basis? Here, we examine these points in a clinical paradigm on Friedreich's ataxia (FRDA), an autosomal recessive inherited disease, which causes progressive damage to the nervous system (Dürr et al. 1996; Gibilisco and Vogel 2013).

On another front, the extraction of neurobehavioural correlates of complex sensorimotor processes and their integration into models poses a technological value. To date, artificial systems (e.g. robots) have been successful in carrying out distinct activities within well-defined and known task and environmental contexts. However, this ability is compromised significantly under unfamiliar task conditions, in which case robotic devices lag behind human performance. Empirically valid models of motor control and learning thus offer a promising solution towards upgrading AI behaviour in a wide array of complex and potentially newly encountered tasks.

Similarly, we argue that progress in contemporary brain machine interfaces (BMI, Scott 2006; Hatsopoulos and Donoghue 2009) can benefit from the successful implementation of neurobehaviourally inspired motor control and learning models. Conventional BMIs operate based on arbitrary mappings between cortical activation and motor control signals (Shih et al. 2012). While these decoding principles may be sufficient to support specific activities (e.g. target-driven cursor or robotic arm steering), it becomes very challenging to generalise them to the completion of more sophisticated naturalistic tasks (e.g. grasp and complex manipulation of physical objects). In an attempt to evade this limitation, new generations of BMIs could incorporate -like the brain- memory based prediction algorithms; that is internal models of the task and world which capture systemically the relations between actions and their results (Vaadia and Birbaumer 2009).

1.4 Organisation of thesis

The goal of this thesis is to develop computational and empirical approaches that can reflect and trace principles of sensorimotor control and learning in complex motor tasks. In chapters 2-3 we attempt to build a framework of motor learning, whose assumptions and predictions are reinforced by experimental evidence.

In particular, Chapter 2 investigates candidate mechanisms of action representation during life-

like manipulations of physical objects with and without internal degrees of freedom. It provides evidence that motor learning and generalisation of task dynamics occur in a way compatible to the existence of composite motor memories; defined both in a body-based and object-based reference frame.

Chapter 3 integrates these findings to formulate a computational model of action representation and motor learning that fuses control policy learning with a system identification of task related parameters. This framework poses a new approach to exploiting the brain's inherent link between forward models and feedback control so as to estimate near optimal (due to incomplete task knowledge), dynamically updated policies. Its predictions are tested against the transient dynamics of motor learning in an experimental paradigm, which exposes human subjects to unfamiliar task conditions.

Chapter 4 presents a methodological platform to track and examine exploratory behaviour in control policy selection during the course of motor learning. Our approach involves reverse engineering the control process in a whole-body manipulation paradigm and techniques for assessing the stability of learning within and across trials.

Chapters 5-7 we employed parts of our empirical approaches and computational insights from the previous sections to investigate the neural foundation of sensorimotor control and learning. To this end, we first built f2MOVE; an fMRI-compatible haptic object manipulation system for closed-loop motor control studies. Chapter 5 presents the development of the system and its adjustment in a clinical setting (Clinical Imaging Facility, Imperial College London, Hammersmith Hospital, London, UK).

Chapter 6 transfers a life-like object manipulation paradigm on healthy human subjects into neuroimaging sessions and presents an approach to elicit neurobehavioural metrics of motor learning. behavioural measures are namely employed as regressors against fMRI signal so as to extract neural correlates of learning and support a model of the region-specific characterisation of early-fast and late-slow performance updates.

Chapter 7 employs the same clinical paradigm on Friedreich's ataxia patients (FRDA) to gain insight into the impact of neurodegeneration on sensorimotor functions. In fact, it examines this impact on a longitudinal basis (through multiple visits) to describe disease progression on a behavioural and implementation level. The behavioural marker toolbox designed for the baseline study on healthy subjects, is expanded here to encompass a combination of behavioural metrics, that can track on an individual patient basis subtle and noisy performance changes, which are consistent to clinically diagnosed trends. Furthermore, the chapter presents a method to elicit the behaviour-specific evolution of neural activity during the course of the disease. Finally, it outlines an approach to advancing the characterisation of the obtained neurobehavioural markers

via network-level analysis.

Chapter 8 summarises and discusses thesis conclusions and proposes future research directions.

Relevant publications

Parts of Chapter 2 have been presented at the Conference for Neural Control of Movement (NCM, 2012) and published at Frontiers in Computational Neuroscience (Proceedings of Bernstein Conference for Computational Neuroscience, 2012) (Sylaidi and Faisal 2012).

Parts of Chapter 6 have been presented and published at the conference for Computational and Systems Neuroscience (Cosyne, 2015), at Areadne (Conference for Encoding and Decoding of Neural Ensembles, 2014), at the Decision Making Meeting (Bristol, 2014), at the Conference for Neural Control of Movement (NCM, 2014) and the Annual Meeting of the Society for Neuroscience (SFN, 2013).

Within the context of our clinical work in Chapter 5, Chapter 6 and Chapter 7 two full length publications have been accepted by the 7th International IEEE/EMBS Conference for Neural Engineering (2015) and another one by the 4th International Congress on Neurotechnology, Electronics and Informatics (Neurotechnix, 2014) (Sylaidi et al. 2015; Gavriel et al. 2015; Rodriguez et al. 2014).

2 Naturalistic object manipulation is learned based on composite action representations

2.1 Introduction

Sensorimotor integration can be understood as a course of transformations between sensory inputs and motor commands. The study of motor control attempts to shed light on these transformations by unraveling how they are acquired and represented in the brain. This translates to a question of action representation in the brain, which remains a central challenge in motor neuroscience. How does the brain represent information about the body and environment within a given task context? More importantly, how are these representations realized for the purposes of skilled naturalistic behaviour, beyond the wide-spread laboratorary experiments which predominantly focus on reaching movements during a limited number of perturbations?

Elucidating the mechanism of action representation reveals in turn important insight into how the brain generalizes knowledge to novel task conditions (Krakauer et al. 2006; Ghahramani et al. 1996). It has been shown, for instance, that when acquiring a new motor skill, existing action representations may be exploited and adapted and new ones may be formed (Haruno et al. 2001; Miall 2002; Wolpert and Flanagan 2001; Wolpert and Kawato 1998). This generalization process appears to be flexible and dependent on task context (Krakauer et al. 2006). Krakauer et al. 2006 showed that humans transfer knowledge based on their history of prior action, namely the statistics of how they have previously used the various parts of their body and the resemblance that these statistics carry to the new motor task.

2.1.1 Categorically distinct or mixed and adaptive action representations?

Former work in neuroscience has suggested that action can be planned in at least two possible coordinate systems (Shadmehr and Mussa-Ivaldi 1994; Diedrichsen 2007; Diedrichsen et al. 2010; Ingram et al. 2010). One of these constitutes an intrinsic reference frame centered on the body's actuators and sensors, while the other one is an extrinsic frame of reference related to task conditions, environmental settings and object properties. Although evidence has been found for the existence of both frames as well as for their combination, there is still no unified theory for

internal representations.

In a study that tested categorically distinct hypotheses of extrinsic and intrinsic coordinate frames, Shadmehr and Mussa-Ivaldi 1994 investigated how the central nervous system learns to control movements in different dynamical conditions and how its motor policies are represented. Their experiments manifested that human subjects model imposed forces using a combination of computational elements whose output can be broadly tuned inside but also outside a training region in the motor state space. In specific, they showed that subjects are able to generalize any learned dynamics of a motor task to a novel workspace (rotated arm position) provided that the imposed force-field is rotated with the arm (unchanged force orientation with regard to hand). These results led the authors to conclude that the mapping between motion and forces is based on a body-centred (intrinsic) coordinate system of afferents and actuators.

This conclusion has been both verified and contradicted by a number of subsequent studies, depending on the type of perturbations subjects adapt to and the type of task context to which they have to transfer the initially acquired knowledge. An overview of these studies obtains evidence for both intrinsic (Malfait et al. 2002, 2005; Shadmehr and Moussavi 2000) and extrinsic (Burgess et al. 2007; Criscimagna-Hemminger et al. 2003; Krakauer et al. 2000) action representations. Another study showed that subjects select either intrinsic or extrinsic action representation reference frames depending on their familiarity with the task dynamics (e.g. object dynamics) (Ahmed et al. 2008a). Furthermore, there is a body of work that advocates the narrative of local adaptation, according to which the brain does not generalise efficiently to very unfamiliar task conditions (Burgess et al. 2007; Donchin et al. 2003; Gandolfo et al. 1996; Krakauer et al. 2000; Lackner and Dizio 1994; Mattar and Ostry 2007, 2010). In agreement with this narrative, other studies on movement error generalisaton argue that the brain employs representations, which encode a very narrow mapping associated with specific task contexts (Ingram et al. 2010; Kadiallah et al. 2012; Thoroughman and Taylor 2005).

En masse, the results from a large amount of studies suggest that motor learning may not be simply explained based on categorically distinct representations associated to extrinsic and intrinsic reference frames. Instead, there appears to be a distinct flexibility in the selection of representation bases and a sensitivity of generalisation tuned to the encountered task circumstances. Indeed, there is work suggesting that actions are planned based on a combination of extrinsic and intrinsic variables (Brayanov et al. 2012; Berniker et al. 2014; Diedrichsen 2007). Diedrichsen 2007 proposed that motor coordination should generalize to a representation level defined by low level state estimates of effectors (body, hands) and to one defined by high level state estimates of the object being manipulated during a task. This suggests that internal models rely on a hierarchical organization of action representations. However, even these approaches to mixed representation schemes have not solidified a generic formalization of internal represen-

tations. Moreover, they have not managed to address the question of action representation in naturalistic motor behaviour, as opposed to the un-natural task context usually introduced in experiments by the use of a robotic manipulandum.

2.1.2 Naturalistic approaches to behaviour: biological and artificial systems

Recent studies in human sensorimotor control have started following a more naturalistic approach to designing experimental conditions. One aspect of this trend is expressed in the introduction of real-world objects in the examined task settings, since object manipulation emerges as a dominant component of naturalistic behaviour in daily life (Ingram and Wolpert 2011; Brown 1986; Parker and Gibson 1977; Piaget 1954). Arguably, object manipulation poses a challenge for sensorimotor control since grasping an object can notably modify the body dynamics (Atkeson and Hollerbach 1985; Bock 1990; Lacquaniti et al. 1982). Therefore, an increasing number of studies instruct an interaction with physical objects with familiar dynamics (Flanagan and Beltzner 2000; Gordon et al. 1993; Johansson and Westling 1988; Nowak et al. 2007) or simulated objects with familiar (Witney and Wolpert 2003; Witney et al. 2000) or unfamiliar dynamics (Caithness et al. 2004; Gandolfo et al. 1996; Howard et al. 2008, 2010; Malfait et al. 2002; Shadmehr and Brashers-Krug 1997; Shadmehr and Mussa-Ivaldi 1994; Tcheang et al. 2007). The studies employing physical objects are nevertheless commonly limited to rigid body physics and do not investigate adaptation to objects with internal degrees of freedom (IDOF). Overall, none of the two types of experimental approaches to object manipulation (physical and simulated objects) have managed to unravel the mechanisms of action representation in naturalistic task contexts.

Sensorimotor mechanisms that underlie complex object manipulation tasks have also been a focus of engineering approaches. Work in robotics has provided various motor control formalisations, including adaptive critic neural network-based object grasping controllers (Jagannathan and Galan 2004) and hierarchical optimal control models, which achieve complex object manipulations (Simpkins and Todorov 2011). These studies often set interesting assumptions on body and object representation and draw on intelligent computational properties found in biological systems. However, they are usually not regressed against human behaviour and therefore do not offer a computational interpretation of the actual brain processes.

2.1.3 Core contributions

In the present study, we investigate three candidate mechanisms of action representation, each of which poses a different hypothesis on how intrinsic and extrinsic variables are combined to support a complex motor task. Two of the mechanisms propose a hierarchical relationship between

body-based and object-based representations of the task, whereas the third one captures task dynamics through a composition of independent extrinsic and intrinsic action representations. We test the validity of each mechanism against human behaviour in two object manipulation paradigms. The novelty of our approach lies primarily in the formalisation of our action representation scenaria which reflect the composite contribution of intrinsic (body-based) and extrinsic (object-based) variables. It further lies in our experimental design which attempts to mimic naturalistic task conditions by instructing the interaction with real-world objects with complex internal dynamics or no internal degrees of freedom. Crucially, our work presents a new inclusive investigation of how previous, categorically distinct theories of action representation can be used as parts of a unifying framework to support motor planning for the purposes of skilled naturalistic behaviour.

2.2 Aims and Methods

The primary aim of this section is to investigate how sensorimotor transformations are represented by the brain in a naturalistic object manipulation paradigm. This translates to a question of how the brain represents the body and object for the purposes of skilled motor behaviour. We tested three candidate mechanisms of action representation, which pose hierarchical and composite assumptions on the organization of intrinsic and extrinsic variables of task dynamics. We examined which of these competing scenaria best predicts the generalisation of learned motor skills (training session) to novel task conditions (testing session). In particular, we probed how human subjects transfer acquired knowledge from a single training context to two different testing contexts of either varying body configuration or varying object configuration. Our experimental paradigm instructs the interaction with real-world objects in an attempt to escape conventional laboratorary settings, which predominantly focus on simple hand reaching movements and commonly impose un-natural task contexts with the use of robotic manipulanda (howbeit, see crefnaturalistic for recent advancements in robotic manipulandum technology that simulates physical object dynamics).

2.2.1 Formalising action representation scenaria

Our approach to describing scenaria of action representation employs reference frames previously used in literature, primarily as categorically distinct hypotheses (see Section 2.1.1). One of these reference frames constitutes an intrinsic coordinate system centered on the body's actuators and sensors, while the other one is an extrinsic coordinate system related to task conditions, environmental settings and object properties. For the purposes of the present study, we name the former a joint-based reference frame (J) and the latter an object-based reference frame (O) (Fig. 2.1). A task learned within the joint-based frame can be generalised across multiple joint

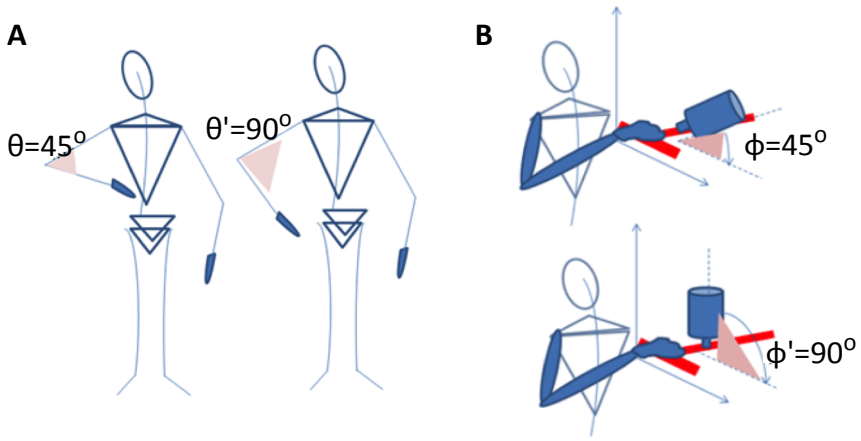

Figure 2.1: (**A**) A joint-based reference frame. If task dynamics are represented in this coordinate system, then the brain can generalise acquired knowledge across multiple similar joint-configurations. (**B**) An object-based reference frame. If task dynamics are represented in this coordinate system, then the brain can generalise acquired knowledge across multiple similar object-configurations.

configurations, whereas a task learned within the object-based frame can be generalised across multiple object configurations. Based on previous evidence on local learning (Burgess et al. 2007; Donchin et al. 2003; Gandolfo et al. 1996; Krakauer et al. 2000; Lackner and Dizio 1994; Mattar and Ostry 2007, 2010), we can argue that this notion is valid for joint and object configurations in a relative proximity to the training task conditions.

We assume that while the body interacts with an object, the brain learns the dynamics of the task in three possible ways which can also be perceived as three distinct hierarchical Bayesian models (Gelman et al. 2014), using variables T denoting the complete task dynamics, J the joint based reference frame and O the object based reference frame:

1. The dynamics of the body are learned in dependence of the dynamics of the object (serial architecture, Fig. 2.2 (i)). In this hypothesis, the brain learns to generalise across multiple joint configurations, but with regard to one distinct object configuration. This relationship can be formalised probabilistically as the belief over the complete task dynamics $P(T|JO)$ that can be composed by the conditionally dependent factors $P(T|J)$, $P(J|O)$ and a prior belief over the object based reference frame $P(O)$:

$$P(T|JO) \sim P(T|J) \cdot P(J|O) \cdot P(O) \tag{2.1}$$

2. The dynamics of the object are learned in dependence of the dynamics of the body (serial

architecture, Fig. 2.2 (ii)). In this hypothesis, the brain learns to generalise across multiple object configurations, but with regard to one distinct body configuration. Similarly to the previous case, this mechanism of action representation can be captured probabilistically based on a composition of the conditionally dependent factors $P(T|O)$, $P(O|J)$ and a prior belief over the joint based reference frame $P(J)$.

$$P(T|JO) \sim P(T|O) \cdot P(O|J) \cdot P(J) \tag{2.2}$$

3. The dynamics of the object are learned independently from the dynamics of the body and combined ad hoc so as to capture the dynamics of the task in a composite fashion (parallel architecture, Fig. 2.2 (iii)). The Bayesian architecture that reflects this mechanism of action representation therefore depends on a composition of the independent factors $P(T|J)$, $P(T|O)$ and the prior beliefs over both the joint based reference frame $P(J)$ and the object based reference frame $P(O)$:

$$P(T|JO) \sim P(T|J) \cdot P(T|O) \cdot P(J) \cdot P(O) \tag{2.3}$$

2.2.2 Empirical validation logic

In order to understand which of these three architectures constitutes the underlying mechanism of learning in object manipulation tasks we designed an experimental approach based on the logic of Figure 2.2. The first part of the design (Experiment A, see Fig. 2.2) includes an object manipulation experiment with one training and multiple testing sessions, throughout which joint (body) coordinates remain constant. That is, the workspace of the subjects is stable in all task contexts imposed on the subjects. On the other hand, the testing sessions impose varying object configurations; that is the external forces imposed on the subjects' end-effector (e.g. hand) through the manipulated object vary in orientation (due to changing object positioning). Based on our introductory definitions of action representation scenaria (see Section 2.2.1), subjects are expected to be able to transfer succesfully any learned task dynamics (training session) to novel object configurations (testing session), if their motor system generalises in an object-based reference frame (consistent to Fig. 2.2 (ii)). If subjects do not transfer learned dynamics successfully to novel object positioning conditions, it can be inferred that they are potentially capable of generalising in a joint-based coordinate frame (Fig. 2.2 (i)). This result is also consistent with our composite action representation scenario in which the brain generalises both in a joint-based and an object-based frame of reference (Fig. 2.2 (iii)).

To distinguish between cases (i) and (iii), an inverse version of our initial experimental design

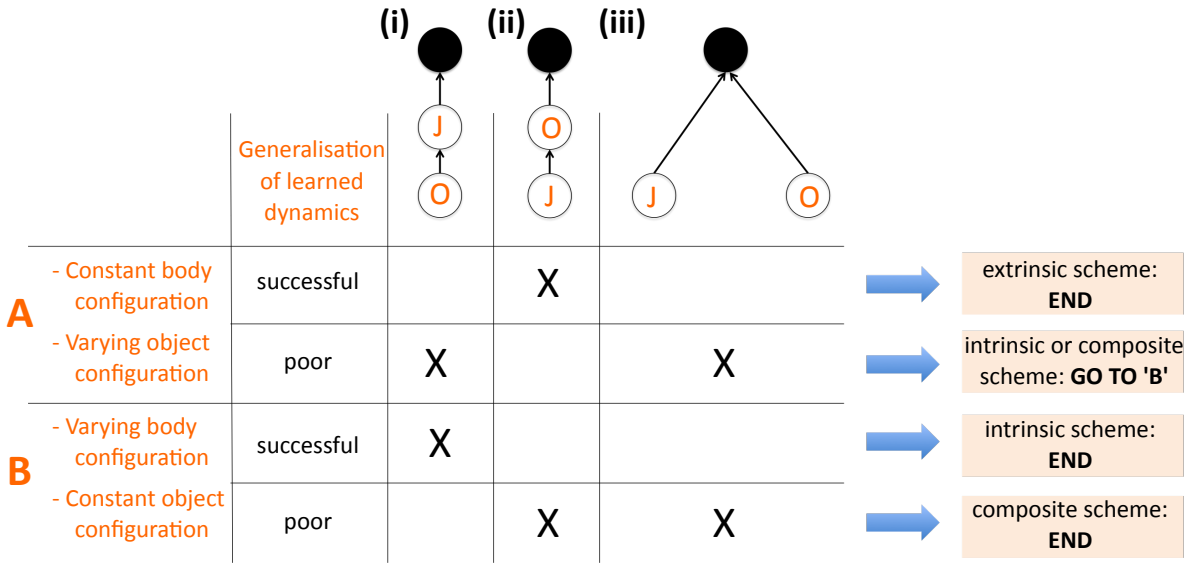

Figure 2.2: This schematic illustration reflects the logic of our experimental approach. Our design employs an object manipulation paradigm which includes one training and multiple testing sessions. Human subjects learn an object manipulation task under distinct conditions of joint and object configuration. Then they are asked to perform multiple testing sessions in which task conditions are perturbed. Their ability to generalise learned dynamics to novel task contexts is evaluated and used as evidence for or against the existence of our three candidate action representation theories ((**i**)-(**iii**), also see Section 2.2.1). Experiment A includes testing sessions which preserve constant joint coordinates and impose varying object configurations. If its results exclude scenario (ii), Experiment B is employed. The latter includes testing sessions which preserve constant object coordinates and impose varying joint configurations. Its results can distinguish between the empirical validity of the two remaining scenaria.

can be employed (Experiment B, see Fig. 2.2). The latter needs to examine subjects' ability to transfer learned dynamics when the external forces applied on their body through the manipulated object remain constant in orientation. At the same time, their workspace and consequently their joint configuration varies across different testing sessions. If subjects are able to transfer learned dynamics successfully to these novel task conditions, we can assume that they generalise in a joint-based reference frame. Otherwise, a composite action representation scheme remains as the prevailing interpretation of behaviour. In summary, Experiment A is used to distinguish the empirical validity of scenario (ii) from the remaining two scenaria. If the obtained results provide evidence of poor generalisation, then scenario (ii) is excluded and Experiment B can be employed to distinguish between the empirical validity of scenaria (i) and (iii).

2.2.3 Experimental setup and data acquisition

Subjects were asked to sit at a 3D virtual reality rig, which projected the experimental task via a 120 Hz, 1280x720 native resolution monitor and a mirror system (Fig. 2.3). They were instructed to hold an object in their right hand and were able to receive visual feedback of the object motion on a screen positioned at the level of their eyes. The screen displayed a simulated view of the object inside the workspace. Two sensors were placed at the two object extremes and allowed tracking of its location based on an electromagnetic motion tracking device (LIBERTY, Polhemus), operating in real time at a 240 Hz update rate.

Two objects were designed and used in the present experimental study. They both resembled plastic bottles of 1 l capacity, which were mounted on a plastic T-shaped handle. Their mounting allowed a rotation of the object at variable angles about the handle, which in turn enabled changes of the object orientation with regard to subjects' hand. One object type was half filled with water and sealed, which assigned internal degrees of freedom to its dynamics. The second object type matched the first one in shape, volume and mass, but possessed no internal degrees of freedom (Fig. 2.3). Instead of liquid content, a 1 kg wax layer was uniformly distributed on its internal surface. Throughout all experimental sessions certain body properties were taken into account (arm joint distances, relative positioning of right shoulder with respect to eyes) to callibrate and normalise the instructed 3D workspace (home and target locations). Furthermore, arm positioning inside the rig was stabilized at the same pre-selected angles on the saggital, coronal and transverese plane for all subjects. This positioning was used as the reference body configuration in the training session and the testing session of Experiment A (Fig. 2.2). It was partially perturbed in the testing session of Experiment B (Fig. 2.2) since the latter imposed task conditions of varying joint configuration.

The design of the experiment, including any necessary visual feedback and data acquisition, was programmed using a custom made motorlib C code library and opengl. All acquired motion data was analysed using MATLAB (The Mathworks Inc., Natick, Massachusetts, USA).

2.2.4 Experimental protocol

Thirty right-handed subjects (21 men; 9 women; mean age, 25.27 years) participated in this study, which was subdivided in two paradigms, each of which instructed interaction with a different object type (either an object type I, of complex internal dynamics or an object type II, with no internal degrees of freedom). Fifteen subjects participated in the first paradigm and fifteen in the second paradigm. Each subject attended only one of the two paradigms. Each participant attended and completed both experiments included in each paradigm (Experiment A and B, see Fig. 2.2). Subjects were naive to the purpose of the experiments and all provided

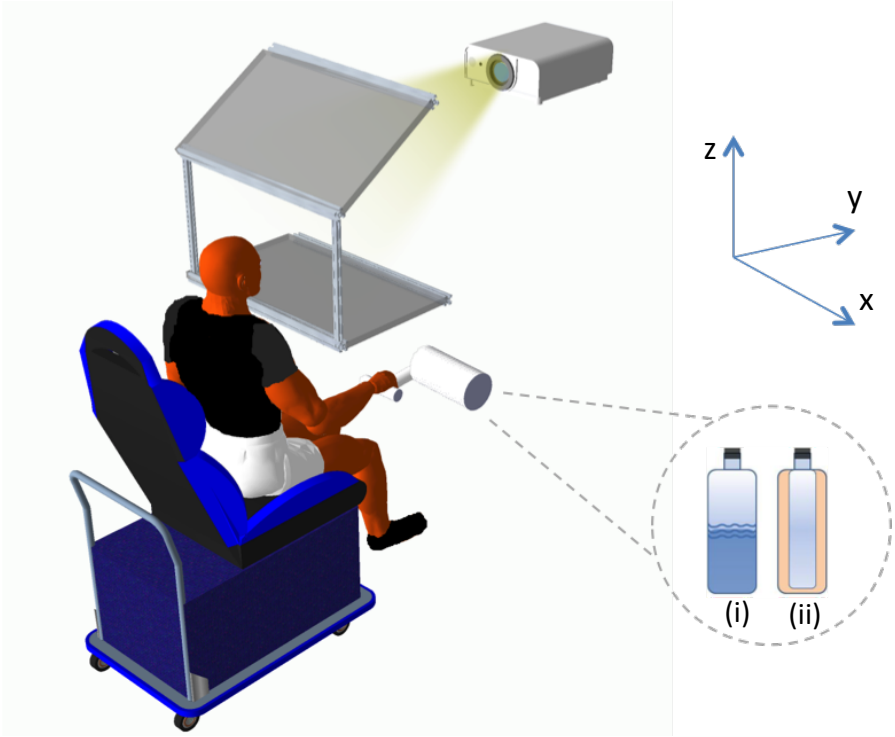

Figure 2.3: Both experimental paradigms used in this study were hosted by a 3D virtual reality rig. The rig projected task conditions via a 120 Hz, 1280x720 native resolution monitor and a mirror system. The first paradigm instructed the manipulation of (**i**) a naturalistic object that resembled the dynamics of a half-filled 1 l bottle of water (internal degrees of freedom), whereas the second paradigm instructed the manipulation of (**ii**) an object with no internal degrees of freedom which matched the first one in shape, volume and mass. Both object types were stabilized on a T-shaped handle via a customised mounting system that allowed their rotation about the handle axis and thereby their orientation change with regard to the subjects' right hand. Subjects received feedback of the object positioning and their overall performance only from a virtual reality screen positioned at the level of their eyes.

informed consent consistent with the policies of the Imperial College London Ethics Committee.

Both experimental paradigms employed the same design, differing only in the object types used for manipulation. This design divided each paradigm into two experiments (Experiment A and B, see Fig. 2.2), each of which was further subdivided into five phases. In each of them participants had to rotate the provided object between a home and a target orientation that were visualised on the virtual reality screen. Rotation was instructed on the xy plane (Fig. 2.4 A). At the beginning of each trial the object was stabilized at distinct orientations defined on a plane vertically configured with respect to the hand (0°, 45°, 90°, 135°, 180° in phases (i), (ii), (iii), (iv) and (v) in Figure 2.4 A respectively).

Phase (i) was used as a training session whereas the other four phases were all used as testing sessions with variable task conditions. The training session's task context was retained the same for Experiments A and B. In particular, phase (i) (Fig. 2.4 A) contained 40 trials and was intended to familiarise subjects with the task of rotating the object between 0° and 315° (in a clockwise and anti clockwise fashion). A larger number was avoided to prevent significant fatigue symptoms which were observed in preliminary testing of our paradigm's completion feasibility. At the beginning of each trial subjects were presented with a visual feedback of the instructed home orientation of the object (in this case a cylinder of (0,0,0) origin, 0.5 cm radius and 15 cm length along the 0° axis). The object's visualisation followed the basic shape and dimension features of the physical object used in the paradigm (bottle-like form of 4.5 cm radius and 15 cm length). As soon as subjects stably aligned the object with the home orientation, a target orientation appeared on the screen (in this case defined along the 315° axis), which subjects had to reach, so as to subsequently return to the home orientation and terminate the trial. The required rotation stability was defined within a range of 0.5 cm accuracy around the instructed rotation centre (a red sphere of 1 cm radius at the origin of the home and target orientations). Subjects had to hold the proximal end of the object within this distance from the rotation centre and for movements which did not satisfy this stability criterium the trial was terminated and reported as failed.

Feedback on the subjects' performance was illustrated on the screen at the end of each trial. It was provided in the form of error measurements proportional to the cumulative displacement of the object's proximal end from the instructed rotation centre across all trial recordings. Subjects were asked to minimize this cumulative error throughout the experiment. The instructed time limit for trial completion was 2.4 s and participants received a visual message for success or failure if they completed the rotations before or after this threshold respectively.

Phases (ii)-(v) (Fig. 2.4 A) were intended to examine the ability of subjects to generalize any learned task dynamics (phase (i)) to novel task conditions. The task contexts introduced in

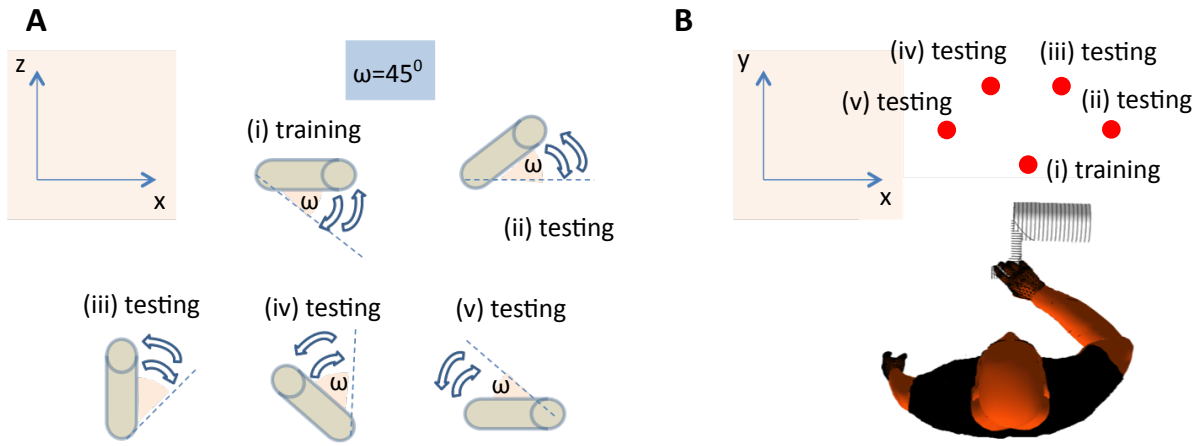

Figure 2.4: Experimental design (**A**) Experiment A is subdivided in one training session (i) and four testing sessions (ii)-(v). Subjects learn how to perform a rotational movement about a given rotation axis within a given accuracy while stably holding an object. In the testing sessions they are examined at how well they transfer the learned task dynamics to novel conditions, under which the workspace is retained fixed while the object positioning with regard to the hand varies. (**B**) Experiment B similarly employs a training session on object manipulation (same conditions to Experiment A). Subsequently it presents subjects with four testing sessions, during which the object positioning with regard to the hand remains stable, while the workspace and thereby the joint configuration required for task completion vary. In both experiments the testing phases are presented to subjects in a pseudorandom order to counterbalance the effects of task sequence on human performance.

these testing sessions, each of which contained eight trials, preserved a fixed workspace (object proximal end origin at (0,0,0) and varied the object orientation with regard to the hand which in turn varied the subjects' rotation range within the given workspace. The new instructed rotation ranges were [0°-0°], [45°-0°], [90°,45°], [135°,90°], [180°,135°] for phases (ii), (iii), (iv) and (v) respectively (Fig. 2.4 A). Between consecutive testing phases subjects also completed shorter versions of the training phase (20 trials) so as to re-adapt to the initially learned task dynamics.

In Experiment B a similar design was used to examine how subjects transfer acquired knowledge to novel task contexts. Participants were in fact presented with the same traning phase used in Experiment A. However, subsequently they were instructed to complete a new group of testing phases for a fixed object positioning with regard to the hand but a varying workspace. The new object proximal end origins were namely displayed at a 15 cm distance from (0,0,0) of the training session and at a 36°, 72°, 108°, 144° counter-clockwise orientation for the new testing phases (ii), (iii), (iv), (v) respectively (Fig. 2.4 B). This task arrangement was aimed at encouraging the recruitment of new joint configurations.

2.2.5 Performance measures and generalisation criteria

We selected a measure of evaluating task performance based on translational displacement of the object handle (in specific, of the object proximal end) relative to the instructed rotation centre. We namely estimate the cumulative displacement in each trial, T, in the form of a root mean square error:

$$T = RMSE(P_{obj}, P_{rc}) = \sqrt{\frac{\sum_{i=1}^{n} (P_{obj,i} - P_{rc})^2}{n}} \tag{2.4}$$

where n refers to the maximum number of recordings during a trial, P_{obj} to the three dimensional position of the object's proximal end and P_{rc} to the fixed three dimensional position of the instructed rotation centre. This cummulative estimation of translational error constitutes a stability measure during the rotational movement and is provided as a quantitative feedback to subjects on their VR screen after the completion of each trial. Subjects are aware that their primary task goal is to minimise this error, which is inherently linked to an additional speed requirement, since for longer trial duration the total translational displacement rises by accumulation.

Based on Equation (2.4) we can evaluate the course of learning of subjects during training. We can namely estimate the cummullative displacement for each trial throughout the whole training period and average the derived error progressions across subjects.

We also use Equation (2.4) to determine naive performance measures for both the training and testing sessions. The motivation is to quantitatively capture subjects' first response to the newly encountered experimental conditions before they started adapting to them. This is determined as the mean cummulative translational displacement across the first two trials of each experimental phase (*naive testing error*). As a control, we compare this naive performance estimate to the mean error estimate of the following six trials in each session in order to exclude a chance (involuntary) performance scenario. The latter would be indicated if the mean error of the latter six trials is significantly higher than the mean error of the former two; an observation that would also be inconsistent to the expected learning trend. In this outlier-case, we substitute the outlying performance with the first following non-outying performance (amongst the six subsequent trials).

Lastly, Equation (2.4) is employed to calculate the mean error estimate at the end of the training phase (*end performance training error*, across the 15 last trials). In order to evaluate generalisation of learned dynamics from training to testing context, we compare the end performance estimates for the training phase to the naive performance estimates of the testing phases. Poor

generalisation would be implied by a statistically significant increase of error from the former to the latter. Furthermore, the scenario of partial generalisation can be investigated by a comparison between the naive performance estimates of the training session (*naive training error*) and the whole of the testing session (all four phases). If subjects generalise partially from training to testing conditions, their initial performance at each testing phase is expected to be lower than the naive performance of the training phase.

2.2.6 Statistical analysis

To perform the statistical comparisons described above between training and testing performance (Section 2.2.5) we apply non-parametric statistical analysis since we cannot assume our data to be normally distributed. We use Wilcoxon rank sum test to examine pairwise the differences between the end performance of the training phase and the naive performance of each of the testing phases. We introduce post-hoc analysis correction of the significance level, because our approach includes multiple pairwise comparisons which might individually show a false significant difference. This problem arises from the fact that an increase in the number of hypotheses in a test, also increases the probability of witnessing a rare event, and thus the probability of rejecting the null hypothesis when it's true. We deal with this risk using Bonferroni corrections to re-adjust the significance level of our statistical comparison.

The same statistical technique is employed to perform comparisons between the naive performance in the training phase and the naive performance in each of the testing phases.

2.3 Results

We formalised three candinate mechanisms of action representation that reflect different hypotheses on how intrinsic and extrinsic variables of task dynamics can be organised to support motor planning and execution. Our three approaches to action representation draw on previous empirical evidence on the existence of body-based and object-based coordinate systems for the definition of task dynamics (Shadmehr and Mussa-Ivaldi 1994; Diedrichsen 2007; Diedrichsen et al. 2010; Ingram et al. 2010). Their generic aim is to pose plausible scenaria on how the brain implements sensorimotor transformations during complex motor tasks. We particularly designed these three mechanisms to encompass and expand the question of action representation to the context of skilled naturalistic behaviour, which has mainly recently attracted attention and therefore calls for targeted empirical studies towards a better understanding of its underlying sensorimotor processes.

Our present study employed two paradigms to investigate how motor planning is represented during the manipulation of an object with complex internal dynamics and an object with no in-

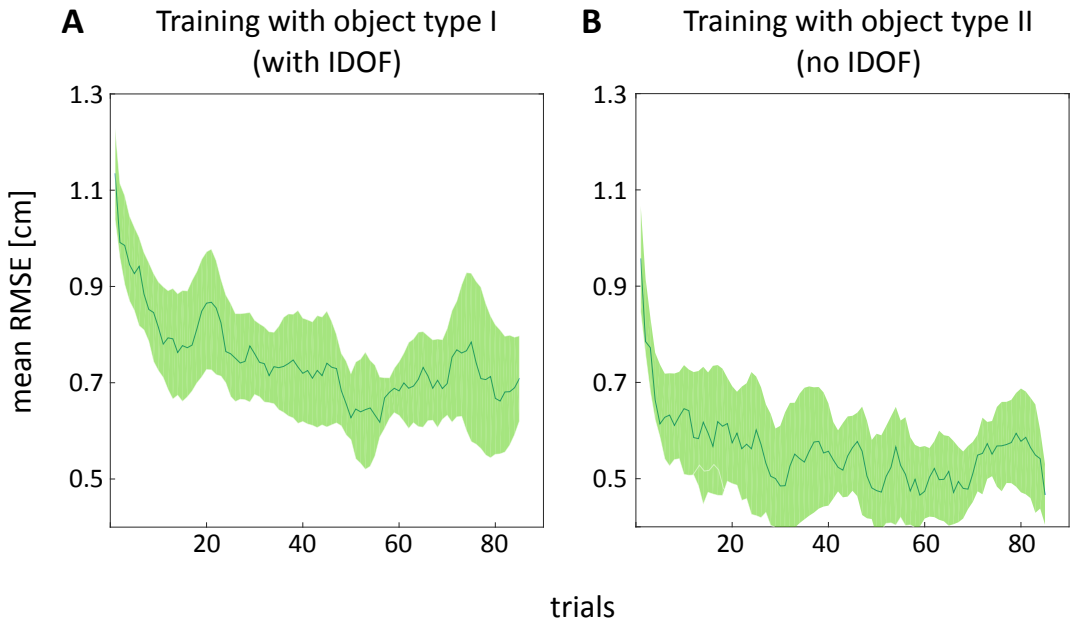

Figure 2.5: Learning object manipulation dynamics (**A**) The learning curve for the training session ($t_{1/2}$=6.5 trials) that instructs interaction with a real-world object with internal degrees of freedom. Each plot estimate illustrates the mean cummulative translational displacement across all subjects. It represents the error between the object's proximal end and the instructed rotation centre and is displayed to subjects after the completion of each trial as quantitative performance feedback on the VR screen, which needs to be minimized. (**B**) The equivalent learning curve for the training session ($t_{1/2}$=3.2 trials) that instructs interaction with an object of no internal degrees of freedom. Performance improvement appears to occur faster in this paradigm as compared to (A). Both training sessions (A and B) encounter a small performance deteriotation towards the final trials, which can be presumably interpreted as an effect of fatigue.

ternal degrees of freedom. Each paradigm was further subdivided into two experimental sessions (Experiment A and B, see Figure 2.2), each of which was intended to test the empirical validity of our proposed action representation mechanisms. In Experiment A we instructed subjects to complete a training session during which they familiarised themselves with the task context. The latter involved matching a rotational task to a given accuracy while holding an object in the right hand.

2.3.1 Learning naturalistic object manipulation

We measured human performance in terms of pivot point displacement; that is the translational displacement of the object's proximal end with regard to the instructed rotation centre. This behavioural measure was acquired as a cummulative estimate throughout all recordings of a

trial (based on Equation (2.4)). It was displayed quantitatively through visual feedback to the participants to reinforce the verbal instructions they received before the initiation of experiment to rotate the object as quickly as possible while holding a stable rotation centre.

We first tested whether subjects' training period features a gradual improvement of performance given the provided performance measures. Indeed, subjects converge to progressively smaller cummulative translational errors (Fig. 2.5). This evolution constitutes evidence of learning the dynamics of the task to which subjects are exposed. Notably, learning the dynamics of the object with no internal degrees of freedom (Fig. 2.5 B) occurs at a faster rate ($t_{1/2}$=3.2 trials) than learning the dynamics of the object with internal degrees of freedom (Fig. 2.5 A) which also appears to converge to a higher error plateau ($t_{1/2}$=6.5 trials). These observations could be justified if subjects possess richer prior experience in manipulating objects with no IDOF (similar to object type II,Fig. 2.5 B), than in manipulating objects with IDOF (similar to object type I,Fig. 2.5 A).

2.3.2 Generalising knowledge to novel object or joint configurations

After training subjects at specific joint and object configuration (Section 2.2.4) we exposed them to test conditions of the varying object positioning with regard to the hand (Experiment A). We thereby attempted to test whether subjects transfer successfully learned dynamics to novel task contexts, following the logic of Figure 2.2 and Figure 2.6 A. In this experiment successful generalisation of already acquired knowledge to the testing conditions is only consistent with action representation scenario (ii) (Fig. 2.6 A). The latter suggests that learning the task occurs within an object-based reference frame which in turn is defined with regard to a distinct joint configuration.

In order to evaluate human behaviour at the testing phases relatively to the training phase we measured the naive performance of the testing phases (*naive testing error*) and compared it against the end performance of the training phase (*end performance training error*, equation (2.4)). This is illustrated in Figure 2.6 B,C as the mean translational displacement between the object's proximal end and the instructed rotation center for the final 15 trials of the training phase (initial object orientation at 0°) and for the first 2 trials of each of the testing phases (initial object orientation at 45°, 90°, 135°, 180° for phases (ii)-(v) in Section 2.2.4 A respectively).

We examined the pairwise differences between the end performance estimates of the training phase and the naive performance estimates of each testing phase using Wilcoxon rank sum tests. Thus, each comparison included 15 mean displacement estimates for each phase, that corresponded to 15 subjects who participated in Experiments A and B of the paradigm based on object type I. We applied Bonferroni corrections to our analysis which readjusted the statistical

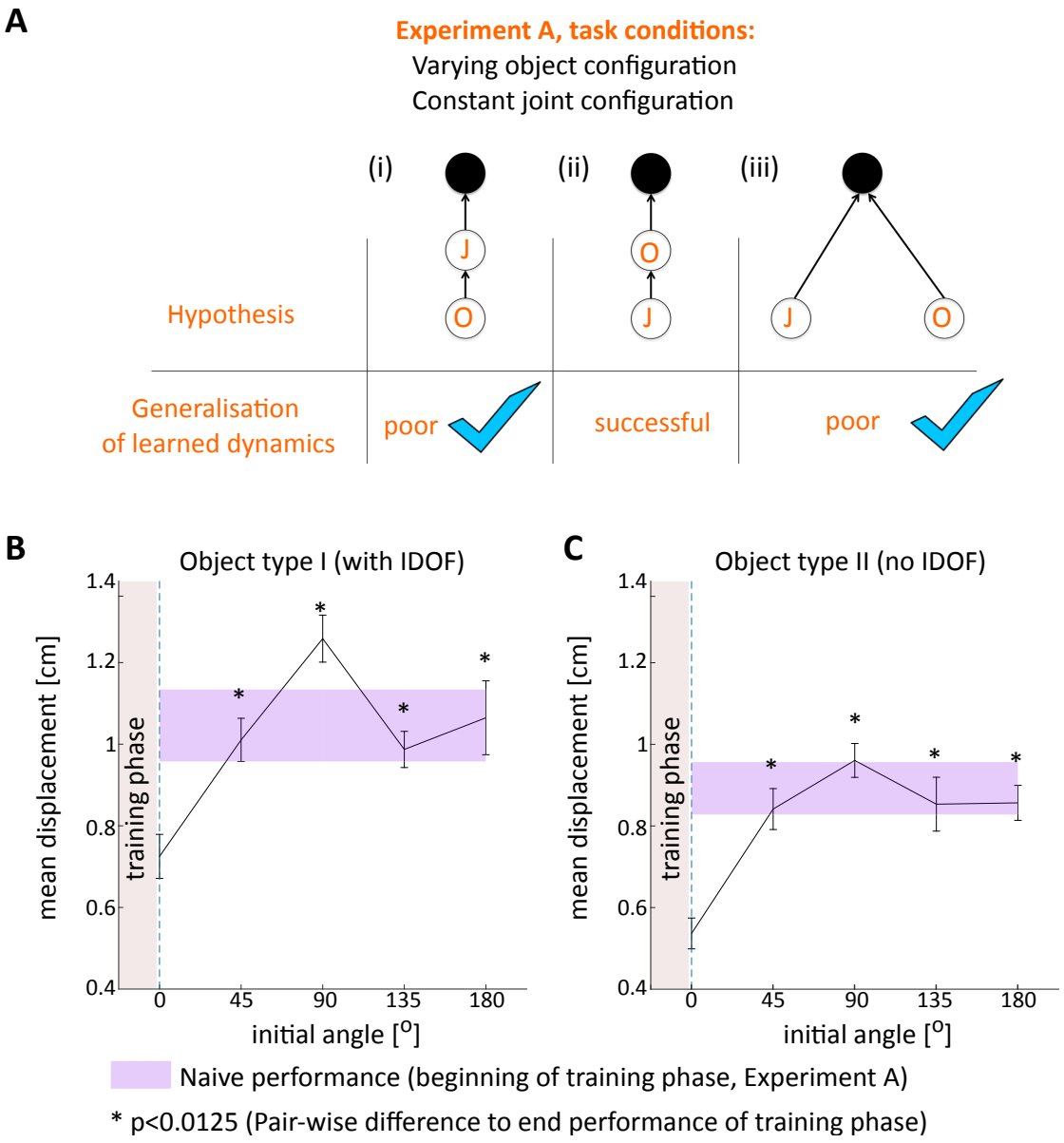

Figure 2.6: (**A**) Experiment A tests how subjects transfer any learned dynamics from training phase to testing sessions which introduce varying object configurations. The only proposed scenario which predicts successful knowledge transfer is (ii) (object-based action representation). (**B**) Findings for the first paradigm, which instructs manipulation of object type I (with IDOF), show that mean translational displacement estimates at the beginning of the testing phases (naive testing errors) are significantly higher than mean translational displacement estimates at the final part of the training phase (p<0.0125, Wilcoxon rank sum tests with Bonferroni corrections). This result suggests poor generalisation to novel task contexts and disvalidates action representation mechanism (ii). At the same time, further non-parametric analysis (see Section 2.3) shows that naive testing errors are not significantly lower than the mean translational displacement estimates at the beginning of the training phase (naive training error). This suggests that there is no clear evidence for partial generalisation, which would arguably bear some consistency with the composite action representation mechanism (iii). (**C**) Similar findings are obtained for the paradigm which instructs manipulation of object type II (no IDOF). Here, there is an evidently smaller variation of naive testing error levels.

significance level of the comparison to 0.0125 from 0.05. The results of the analysis revealed significant decrease in performance levels from training to testing phases, with $p<0.0125$ for all pairwise comparisons. This suggests poor generalisation of learned dynamics from training to testing conditions and excluded action representation scenario (ii) from our investigation (Fig. 2.6 A,B).

We subsequently probed the possibility of partial generalisation, which is arguably consistent with our composite action representation mechanism (iii). In the case of partial generalisation, we might still be able to observe a significant difference between naive performance of the test phases and end performance of the training phase. On the other hand, this scenario would also reveal a significant decrease from naive performance at the training phase to naive performance at the testing phases (see Equation (2.4)). That is, we would expect subjects' initial mean translational errors at the training phases to be significantly lower than the initial mean translational error of the training phase (even if they are still larger than the end performance error of the training phase). Three out of four pairwise comparisons based on Wilcoxon rank sum tests revealed no significant difference in naive performance between training and testing phases ($p>0.0125$ for test object positioning at 45°, 135°, 180°, after post hoc analysis corrections). The fourth comparison (for test object positioning at 90°) showed that the naive error in the testing phase was significantly higher than the naive error in the training phase. This can be justified by the different distribution of liquid for different object configurations, which introduces a different range of forces on the hand each time. Both types of results (Fig. 2.6 B) suggest that for the given paradigm there is no clear evidence for partial generalisation.

We performed the same analysis steps on the data acquired from Experiment A in the second paradigm, which instructed manipulation of an object with no IDOF (Fig. 2.6 C). Similarly to the first paradigm, our findings indicate a significant difference between naive performance of testing sessions and end performance of training session (significant decrease of testing performance, $p<0.0125$ for all four Wilcoxon rank sum tests, after Bonferroni corrections). Furthermore, naive error comparisons between training and testing phases showed no significant difference, which again does not provide evidence for partial generalisation.

Notably, naive error levels at the testing phases vary differently in the two paradigms (Fig. 2.6 B and C), with significant differences noticed primarily amongst the testing phases that use object type I (with IDOF). This is reasonable, considering the greater complexity of this object's internal dynamics, that introduce substantial force perturbations on the hand. Hence, if subjects do not recruit an object-based mapping between experienced forces and perceived states during their training period, they are not able to generalise successfully any learned task dynamics to testing conditions. In summary, for both paradigms Experiment A reveals poor generalisation, which disproofs action representation mechanism (ii) (Fig. 2.6 A) and leads us to a further

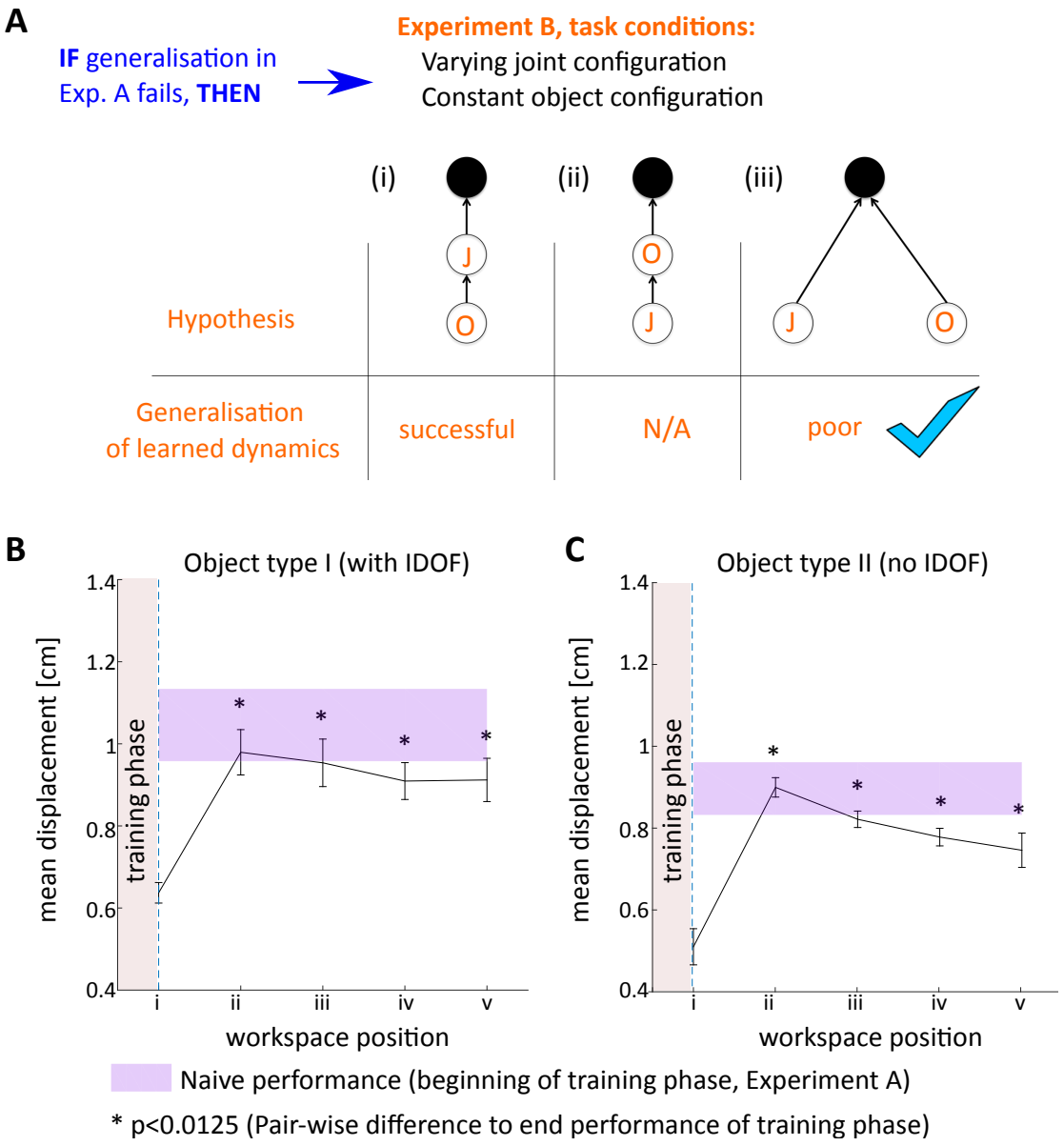

Figure 2.7: (**A**) Experiment B tests how subjects generalise any learned dynamics from training to testing sessions which introduce varying joint configurations (workspace origins (ii)-(v), see Section 2.2.4). From the two remaining action representation mechanisms, (i) is the one that predicts successful generalisation under these testing conditions. (**B**) Analysis reveals a significant increase from end performance training error to naive testing errors for the paradigm which instructs manipulation of a liquid-filled object (p<0.0125, Wilcoxon rank sum tests with Bonferroni corrections). (**C**) Similar findings are obtained for the paradigm which instructs manipulation of a solid object. In summary, both paradigms reveal poor generalisation of learned dynamics to novel joint configurations, thereby disvalidating action representation mechanism (i) and supporting the existence of a composite action representation mechanism (iii). It can be argued that this finding is in agreement with the evidence for partial generalisation, which we obtain for some testing conditions in this experiment. In fact, a number of naive testing errors in the second paradigm (object type II) appear to be significantly lower than the naive training error (p<0.0125 for workspace origins (iv) and (v), Wilcoxon rank sum test with Bonferroni corrections)

investigation of the two remaining mechanisms based on Experiment B.

In Experiment B subjects repeat the same training session that was used in Experiment A, to refamiliarise themselves with its task conditions. They are subsequently asked to perform a series of four testing phases which -inversely to the previous design- vary workspace positioning (and thereby joint positioning), while preserving object positioning fixed with regard to the hand (Section 2.2.4). A successful generalisation of learned task dynamics in this case can be only predicted by our first action representation scenario (i), which proposes that sensorimotor transformations are realised with regard to a joint-based reference frame.

In order to examine whether subjects transfer acquired knowledge succesfully to the new task conditions (workspace origins (ii)-(v), see Section 2.2.4 and Fig. 2.7 B,C), we employ again a comparison between the end performance error of the training phase and the naive errors of the testing phases. Statistical pairwise comparisons reveal that the difference between the former and the latter is significant (significant decrease of testing performance, $p<0.0125$ for all four comparisons, Wilcoxon rank sum test with Bonferroni corrections) for both experimental paradigms (Fig. 2.7 B,C), suggesting that subjects generalise poorly any learned dynamics to the testing sessions.

In an attempt to examine partial generalisation, we perform further pairwise comparisons between the naive testing errors and the naive training error. The comparisons this time lead to mixed results. That is, several naive testing performances appear to bare no significant difference to the naive training performnace ($p>0.0125$ for phases (ii), (iii), (iv), (v) of the paradigm on object type I and for phases (ii), (iii) of the paradigm on object type II, Wilcoxon rank sum test with Bonferroni corrections). On the other hand, a number of naive testing errors appear to be significantly lower than the naive training error ($p<0.0125$ for phases (iv), (v) of the paradigm on object type II, Wilcoxon rank sum test with Bonferroni corrections). This provides evidence for partial generalisation in some cases (primarily for some manipulations of the object type II).

2.3.3 Composite action representation: an adaptive, weighted mechanism?

In summary, Experiments A and B for both paradigms (liquid-filled and solid object manipulation) reveal poor generalisation of learned task dynamics to novel task conditions of varying object or joint configurations. This observation provides evidence against the existence of our first two action representation proposals (see (i), (ii) in Figure 2.2) and supports our composite action representation scenario (iii). The latter posits that task representations are independently learned within object-centered and joint-centered coordinate frames and ad-hoc combined during motor actions.

The design of the present study further allows us to obtain overall performance comparisons

between all used experimental instances. Such comparisons are performed based on each experiment's *generalisation index*. The latter is determined as the difference between the end performance training error and the mean naive testing error averaged across all testing sessions. Based on our definition of cummulative translational displacement for each trial, T, (see Equation (2.4)), the *generalisation index* for each subject can be approximated as:

$$g = \sqrt{\frac{\sum_{s=S-E+1}^{S} \left(T_{end,1,s} - \frac{\sum_{k=2}^{K} \bar{T}_{naive,k}}{K-2+1} \right)^2}{E}} \tag{2.5}$$

with

$$\bar{T}_{naive} = \frac{\sum_{l=1}^{L} T_{naive,l}}{L} \tag{2.6}$$

where S denotes the maximum number of trials in the training phase; therefore index s runs through the last fifteen (E=15) trials of the training phase. $T_{end,1}$ denotes based on Equation (2.4) the translational error for each of these trials and captures a sample end performance of the training phase. k is the index of experimental phases (K=5); with 1 corresponding to the training phase (subscript of $T_{end,1}$) and 2-5 to the testing phases (subscripts of $\bar{T}_{naive,k}$). $\bar{T}_{naive,k}$ represents the mean translational error across the first two trials (L=2) of the testing phases and thus captures the naive performance of each testing phase (with l standing for the testing phase trial index). In order to estimate performance shifts (*p.c.*) due to task context changes we normalise each generalisation index g to its corresponding mean end performance training error and estimate its percent change between different paradigms (see Figure 2.8, where a decrease of g is illustrated as the equivalent increase of performance):

$$p.c. = \left[\frac{\hat{g}_{p2} - \hat{g}_{p1}}{\hat{g}_{p1}} \right] \cdot 100\% \tag{2.7}$$

where $\hat{g}_{p2}$ and $\hat{g}_{p1}$ denote the normalised generalisation index of the second and first task context in the comparison respectively.

In Figure 2.8 A we notice that when transitioning from the liquid-filled object (type I) paradigm to the solid object (type II) paradigm there is an increase in generalisation performance which is predominant for Experiment A. This means that for the solid object paradigm subjects tend to generalise better; that is they generalise less unsuccesfully than when interacting with a liquid-

filled object. This is arguably consistent with the hypothesis that an object-based reference frame plays an enhanced role during interaction with a solid object. This is reasonable, considering that for familiar objects (e.g. solid objects with no IDOF like the one used in our paradigm) our brain probably uses more economical representations, which capture object use regardless of small shifts in object configuration with regard to the body.

When transitioning from Experiment A to Experiment B, we also observe an increase in generalisation performance, which is more significant for the liquid-filled object paradigm (Fig. 2.8 B). This may be suggestive of the enhanced role of a joint-based reference frame in action representation, when the task instructs manipulation of an object with IDOF (liquid-filled in our case). Again, this view can be interpreted as the brain's potential difficulty in representing action economically in object-centered coordinates, when the task introduces notably additional complexity for different object configurations.

Admittedly, the task condition changes introduced by the testing phases of the two experiments (A and B) in the present study are not callibrated. That is, despite their feasibility controls, it cannot be argued that the selected object configuration variations bring about similar performance shifts as the selected joint configuration variations. Therefore, our interpretation of

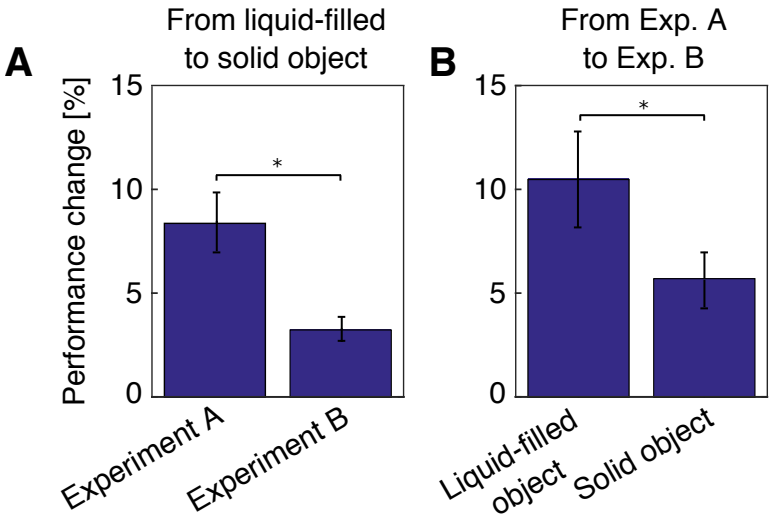

Figure 2.8: Performance changes between paradigms (**A**) From the liquid-filled object paradigm to the solid object paradigm there is an increase of generalisation performance, primarily evident in Experiment A. (**B**) From Experiment A to Experiment B there is an increase of generalisation performance, primarily evident in liquid-filled object paradigm. The statistical significance of across-paradigm differences is estimated based on Wilcoxon rank sum tests with $p<0.05$.

a weighted action representation mechanism should be merely treated as one possible justification of the observed performance changes and can drive future empirical approaches, that can validate and enrich it.

2.4 Discussion

The findings of the present study support the existence of a motor learning mechanism, in which naturalistic object manipulation tasks are represented simultaneously both with regard to a body-based and to an object-based reference frame. They also imply that generalisation of learning to novel body and object configurations as well as the weighting of each reference frame's contribution to action representation depend on the complexity of the object's internal dynamics.

Our conclusions are based on the investigation of three candidate mechanisms of action representation, each of which poses an assumption on how the brain employs extrinsic (object-specific) and intrinsic (body-specific) variables during naturalistic object manipulation tasks. This translates to a question of how the brain represents body and object dynamics to support skilled motor behaviour. Our approach drew on previous work, which suggested the existence of body-based or object-based coordinate systems for the definition of task dynamics (Shadmehr and Mussa-Ivaldi 1994; Diedrichsen 2007; Diedrichsen et al. 2010; Ingram et al. 2010). We expanded previous categorically distinct hypotheses of such reference frames to formalise three mechanisms, which posed either a hierarchical or a parallel synthesis of different types of task variables. Our design exposed human subjects to training and testing conditions of task dynamics and examined how well they generalise knowledge from the former to the latter.

One of the main motivations of the present work lies in the study of naturalistic motor behaviour. Despite a rising interest in naturalistic approaches to designing experimental conditions (Ingram and Wolpert 2011), the unravelling of action representation mechanisms in life-like motor tasks remains an open challenge in neuroscience. Our study, hence, employed physical objects in real-world tasks; namely liquid-filled and solid objects in rotational movements resembling daily activities, e.g. pouring water from a half-filled bottle. Our findings verify our hypothesis that such tasks are learned based on a collaborative contribution of body-based and object-based representations. They arguably also suggest an enhanced contribution of body-based representations for manipulating objects with internal degrees of freedom (IDOF) and object-based representations for manipulating objects with no IDOF (see Section 2.3.3). However, due to lack of calibration between our different testing paradigms, our hypothesis on weighting can be treated as one possible explanation of the observed performance changes and thereby encourage further investigations and empirical validations.

Nevertheless, the narrative of weighted representations lies in agreement with previous work, which showed that object dynamics, can be flexibly represented in different coordinate frames by the brain (Ahmed et al. 2008a). Specifically, according to Ahmed et al. 2008a familiar dynamics are represented in object-centered coordinates and unfamiliar dynamics in arm-centered coordinates. Considering that Ahmed et al. 2008a simulated object dynamics and placed focus specifically on the representation of object dynamics (as opposed to whole task dynamics in our approach), we can conceptually view our composite action representation mechanism as a generalisation of this study's conclusions.

The findings of our study further suggest that the brain encounters difficulty in generalising to very different circumstances than the ones already encountered through extended exposure. Indeed, subjects transfer learning to novel configurations of an object with complex internal dynamics less successfully than to novel configurations of an object with no IDOF (Section 2.3.2). This is consistent with behavioural evidence in several studies that advocate the narrative of local learning, which posits that the brain uses narrow representations of tasks, defined locally around experienced sensorimotor mappings (Berniker et al. 2014; Ingram et al. 2010; Burgess et al. 2007; Donchin et al. 2003; Gandolfo et al. 1996; Krakauer et al. 2000; Lackner and Dizio 1994; Mattar and Ostry 2007, 2010). Notably, the aspect of weighting in a composite action representation scenario offers flexibility, which might counterbalance the consequences of local learning on generalisation performance.

Additionally, the existence of weighted contributions would pose interesting links to the problem of credit assignment; that is how the brain interprets and responds to errors. Indeed, credit assignment constitutes a key area of open research in motor neuroscience. It translates to a challenge of decision making on how the brain should assign the credit of an error across body dynamics, object dynamics or even across past actions (Wolpert and Landy 2012; Berniker and Kording 2008; White and Diedrichsen 2010; Wei and Körding 2009; Dam et al. 2013; Verduzco-Flores and O'Reilly 2015). An interesting future direction of work would thus be to investigate how credit assignment can explain and validate the weighting of extrinsic and intrinsic variables in our composite action representation mechanism. Such an approach could expand our formalisation by bridging the level of sensorimotor mapping and learning to a context of decision making. Moreover, this expansion can be enriched by introducing a probabilistic perspective in the interpretation of error sources. Particularly, a Bayesian framework (Wolpert and Landy 2012; Körding and Wolpert 2004) would capture how the allocation of credit (and in turn the weighting of different action representations) is dependent not only on sources that are most consistent with the error, but also on previous experience.

As a composite action representation scenario, our proposed mechanism is in agreement with previous findings, which suggested that the brain relies on mixtures of extrinsic and intrinsic

variables (Brayanov et al. 2012; Berniker et al. 2014; Diedrichsen 2007) to support skilled motor tasks. Our work attempts an inclusive investigation of this scenario and expands its validity to the context of naturalistic object manipulation. As such, it constitutes a further contribution to formalising a unifying framework to capture the intricate and still -to a great extent- obscure workings of motor planning and learning.

3 A model-based learning rule that predicts control policy updates in complex motor tasks

3.1 Introduction

The brain coordinates a continuous coupling between perception and action in the presence of uncertainty and incomplete knowledge about the world (Faisal and Wolpert 2009; Faisal et al. 2008; Orbán and Wolpert 2011). This mapping is enabled by control policies and motor learning can be perceived as the update of such policies on the basis of improving performance given some task objectives (Izawa et al. 2008; Shadmehr and Krakauer 2008; Wolpert et al. 2011). This process links photoreceptor signals to high-level goals, to real-time control inputs applied for muscle activation and task execution. However, the mechanisms that underlie the revision of control policies during learning remain elusive, particularly in complex motor tasks, e.g. tasks that involve the manipulation of an object of unknown dynamics.

The open question of policy learning addresses the understanding of motor behaviour as the latter is illustrated in human daily activities, in which seemingly simple motor tasks often appear to employ complex adaptation processes. The progression of such processes is particularly evident in sports history, in which motor strategies are sometimes revised over long periods of time and performance is improved through exploration-exploitation trade-offs. An exemplar case is demonstrated by the history of high-jump, that evolved over many decades. From the early 'straight-on' and 'scissors' techniques in the 19^{th} century to the 'Fosbury-flop' golden standard at the end of the 1960s, policy learning and performance optimization occurred over multiple generations and thousands of athletes' careers. So, how does the brain learn control policies while it seeks to adapt to a task?

3.1.1 Learning: from cognitive decisions to motor execution

The study of learning in the context of high-level action selection and decision making has revealed several mechanisms by which symbolic control policies are learned during the repetitive

execution of a task. Decision theory and reinforcement learning have namely provided formal frameworks that allow us to capture behaviour in cognitive tasks (Dayan and Daw 2008; Bülthoff 1996a; McNamara and Houston 1980; Glimcher 2004; O. 1985; Bertsekas 2001; Bellman 1957; Bertsekas and Tsitsiklis 1995; Berry and Fristedt 1985; Gittins et al. 2011; Green and Swets 1966; Kording 2007; Mangel and Clark 1988; Montague et al. 1996; Puterman 2014; Sutton and Barto 1998; Lai 1998; Bülthoff 1996b) through policy upgrade based on reward-prediction errors. Notably, both model-based and model-free versions of such methods of making choices have been found to be consistent with behaviour and therefore constitute vital tools for describing and predicting psychological data and their neural underpinnings (Gold and Shadlen 2007; Niv et al. 2006). Nevertheless, the question remains how the brain bridges cognitive decision making to low-level execution commands. In other words, how can the brain simultaneously learn which abstract actions to select ("turn bicycle to the left") and which control policies will coordinate our muscles accordingly ("flexion of left and extension of right muscles")?

Motor neuroscience has been to date primarily focussed on studying and describing motor learning in terms of the low-level continuous control of actuators (e.g. forces, torques, movement displacements) (Berniker and Kording 2008; Chen-Harris et al. 2008; Yang et al. 2011; Tee et al. 2010). Furthermore, work in motor adaption has quantitatively explained how the sensorimotor system reacts to unexpected perturbations in the experienced task dynamics (Braun et al. 2009a). Such adaptation has been either interpreted as an error-based improvement in performance in response to altered conditions (Krakauer and Mazzoni 2011; Shadmehr et al. 2010), or as a mechanism that attributes a form of implicit reward to the incremental reduction in sensory prediction errors, which in turn leads to an effective reinforcement process (Izawa and Shadmehr 2011). Yet, such mechanisms of updating task kinematics and dynamics (Shadmehr and Mussa-Ivaldi 1994) within or across trials have not clearly explained the simultaneous generation and progression of control policies.

The question of motor learning has also been addressed by a number of studies which investigate how a previously encountered task can be (re-)learned at a much faster rate, a phenomenon referred to as savings (Kitago et al. 2013; Shadmehr and Moussavi 2000; Malfait et al. 2002, 2005). It is speculated that this adaptation employs recalling and updating already learned control policies. In fact, it has been shown that the brain can extract the common underlying structure of variable tasks, which enables savings and structure-specific facilitation of the already learned motor control process in future trials of similar tasks (Braun et al. 2010, 2009b; Acuña and Schrater 2010). Such a process relies on the exploration of parameters during learning along lower-dimensional task-relevant parameter subspaces. It thus appears to be more suggestive of a mechanism of model-based learning, which enables the brain to retain a few internal models across multiple different tasks, the differences amongst which can be captured by abstract meta-

parameters (Cothros et al. 2006; Braun et al. 2010). Still, studies supporting these frameworks have not projected the question of learning from the context of adaptation to the level of policy selection.

Policy selection has been in fact examined by a growing body of empirical work, which showed that motor behaviour and particularly the formation of control signals is consistent with Optimal Feedback Control theory (OFC) (Todorov and Jordan 2002; Todorov 2004; Scott 2004; Diedrichsen 2007). OFC approaches have been successful in predicting control strategies and motor performance at the end-point of learning (Nagengast et al. 2009), thereby bypassing the aspect of motor learning while the brain operates in a context of unfamiliar task dynamics.

3.1.2 Expanding motor learning to control policy learning

Here, we present a computational framework, which expands the formalization of motor learning from the level of task parameter optimization to control policy update. We propose that in unfamiliar tasks the brain makes continuous decisions for the generation of complex trajectories by learning Optimal Feedback Control (OFC) based on the identification of unknown intrinsic and extrinsic task parameters (Sylaidi and Faisal 2012; Brayanov et al. 2012; Ahmed et al. 2008b; Berniker et al. 2014). We use a model-based approach that captures the way in which an actuator (hand) learns how to control an abstract tool (object) of unknown dynamics. Crucially, while learning an optimal controller directly is a sophisticated non-linear dynamic programming problem, our approach postulates that the brain only needs to learn the unknown task parameters, in a system identification process that in turn determines the control policies. It thus advocates the existence of a simple policy learning mechanism on which the brain relies to learn the near optimal control of complex motor tasks.

3.1.3 Core contributions

The novelty of our theory is that it predicts on a trial-by-trial basis the temporal dynamics of human behaviour measured during a novel experimental paradigm on unintuitive object manipulation. It thereby stands in contrast to previous adaptation models that assumed full knowledge of world dynamics and only managed to predict the end-point of learning. Furthermore, while a significant body of work in motor adaptation relies on the assumption of fixed control policies that pursue a distinct target trajectory (e.g. minimum jerk), our framework employs policies that aim at the fulfilment of a broader number of task objectives, presumably similar to the goals set by the CNS during a motor task.

In the work that follows we employ the term motor learning to refer to the mechanism used by the brain to tune world representations and control policies. Notably, the concept of motor learning

expands to the way the brain selects the very structure of world representations and control policies. The two cases of learning could arguably be perceived as a low level of motor execution (action/strategy optimisation) and a high level of cognitive planning (action/strategy selection) respectively. Eventhough our framework explores the former, we believe that a unification of both mechanisms could successfully capture a fuller spectrum of learning and should be the focus of future studies.

3.2 Aims and Methods

The objective of this section is to shed light on the mechanisms that underlie the selection and update of motor control policies in the brain during motor learning, which to date remain unclear, particularly during the execution of complex manipulation tasks. To this end we designed an empirical study, which investigated human subjects' behaviour during an unintuitive object manipulation task. We described this familiarization with a task of unknown dynamics with an OFC framework that explains motor learning in complex motor tasks as a simultaneous update of task representations and action policies. Our model of task dynamics was driven by our findings in Chapter 2 on composite action representations. Our policy learning model (PLM) was regressed against experimental data and its predictions of transient learning dynamics were compared to those of a naïve model (NM) and of an ideal actor model (IAM).

3.2.1 Experimental setup and data acquisition

The experimental setup used for the present study is illustrated in Figure 3.1 A. Subjects were asked to sit in a chair facing a 3D virtual reality rig, which projected the experimental task via a 120 Hz, 1280x720 native resolution monitor and a mirror system. Participants received any visual feedback on a glass screen positioned at the level of their eyes. Underneath the screen, subjects placed and moved their arms on a second glass layer determining the borders of the experimental workspace. In fact, the examined tasks confined subjects' hand movements on this 2D glass plane. Any hand movements inside the setup were captured by electromagnetic motion tracking technology (LIBERTY, Polhemus), operating in real time at a 240 Hz update rate. In particular, during the experiment subjects held one electromagnetic motion tracking sensor in their right hand, thereby allowing an approximation of their right arm joint locations, since their right shoulder joints were held stable at a predefined position close to the chair. The design of the experiment, including any necessary visual feedback and data acquisition, was programmed using a custom made motorlib C code library and opengl. All acquired motion data was analysed using MATLAB (The Mathworks Inc., Natick, Massachusetts, USA).

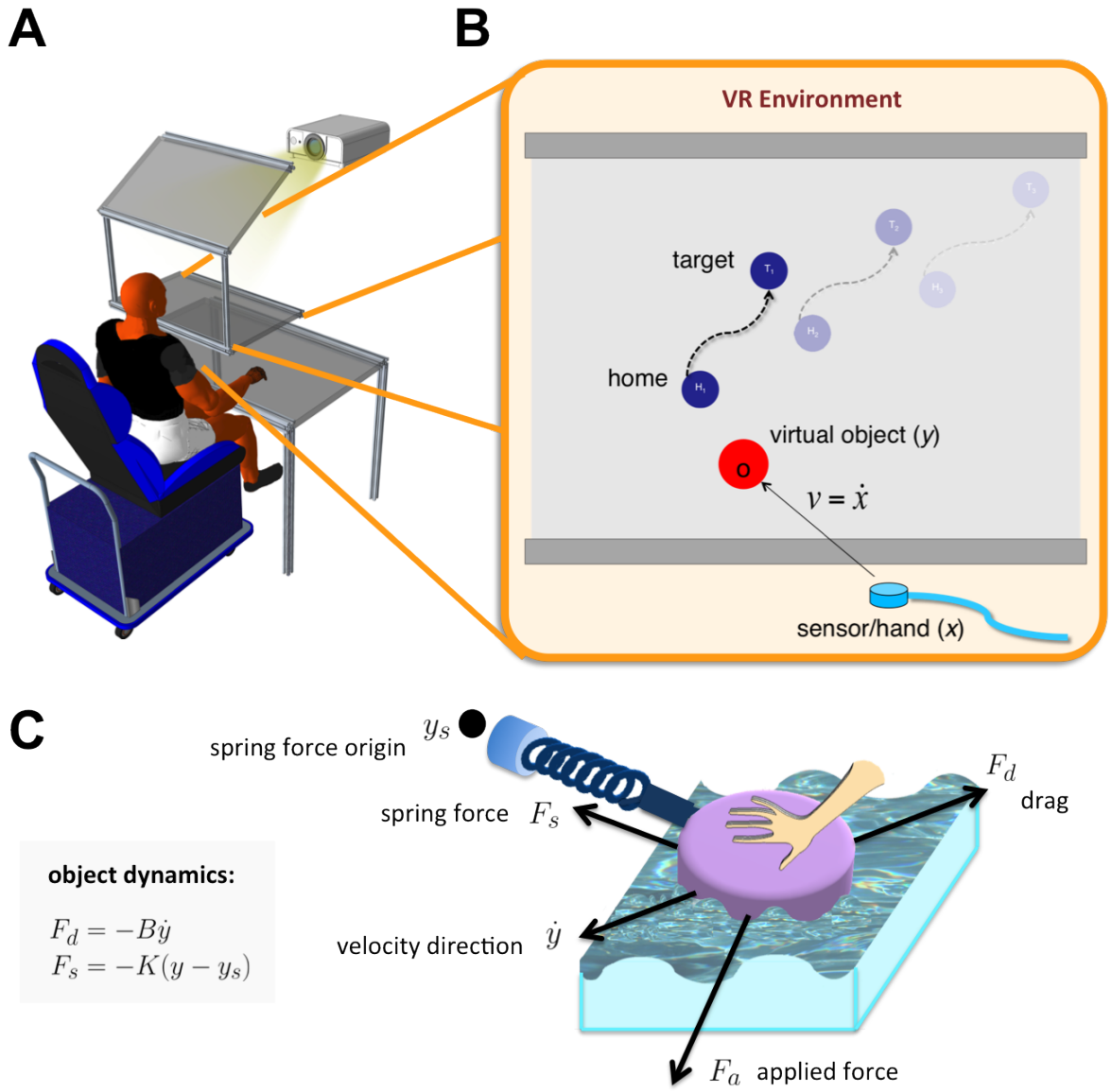

Figure 3.1: (**A**) Human subjects sit inside a virtual reality setup and are instructed to move a manipulated cursor between home and target positions. The movements are captured and displayed on a virtual reality screen with the help of an electromagnetic sensor that subjects hold in their right hand. (**B**) Three sets of target hitting sets are examined. In the first experimental session the manipulated cursor represents the subjects' right hand whereas in the second experiment it represents a virtual object. Thus the first experimental session constitutes a simple hand reaching paradigm, whereas the second one involves object manipulation. (**C**) The object manipulation session is of unintuitive nature, which allows us to study learning from a naïve starting point. Hand velocity maps to a control force on the object state. The object dynamics reflect the impact of a spring force F_s, a damping force F_d which can be used as an approximation to the friction caused by drag, and also depend on the force applied by each subject's right hand F_a, which is proportional to the hand's velocity, as imposed by $v = \dot{x}$. The latter proportionality resembles the forces applied on an object moving in a liquid environment (when Reynold's number is low $Re \ll 1$).

3.2.2 Experimental protocol

Fifteen right-handed subjects (8 women; 7 men, mean age, 26.9 years) participated in the main part of this study, which consists of two experimental sessions, one of which examines hand reaching movements and the other one object manipulation. Additionally, ten right-handed subjects (5 men; 5 women; mean age, 24.5 years) participated in a modified version of the main study, which was used as a control experiment to test subjects' behaviour during long training sessions.

All participants of both the main study and the control study were instructed to complete both examined experimental sessions (hand reaching paradigm and object manipulation paradigm). The subjects were naive to the purpose of the experiments and all provided informed consent consistent with the policies of the Imperial College London Ethics Committee.

Subjects completed two subsequent experimental sessions. In the first one they were required to perform hand reaching movements between starting and target positions. In the beginning of each trial their hand had to be aligned in the starting position for 0.5 s with 0.5 cm accuracy. As soon as this criterion was met, the instructed target was displayed to the participants. Successful hits were considered the ones occurring within 1 s with 1cm accuracy and allowing stabilization of the hand on the target for 0.3 s. Slow attempts were disregarded as unsuccessful. Visual feedback of success or failure was provided accordingly in the form of a short written message. Three categories of trials were chosen by estimating the starting position for each subject's arm features based on a 2-joint-arm configuration of either $\vartheta_{1H} = 70°$, $\vartheta_{2H} = 100°$ or $\vartheta_{1H} = 60°$, $\vartheta_{2H} = 100°$ or $\vartheta_{1H} = 50°$, $\vartheta_{2H} = 100°$ respectively. Similarly, the target position was estimated based on a 2-joint-arm configuration of either $\vartheta_{1T} = 60°$, $\vartheta_{2T} = 90°$ or $\vartheta_{1T} = 50°$, $\vartheta_{2T} = 90°$ or $\vartheta_{1T} = 40°$, $\vartheta_{2T} = 90°$ respectively (Fig. 3.1 B). The subjects were tested in such hand reaching movements for 120 trials (40 for each of the three trial types, all of which were displayed in a random order). We used this experimental session to acquire behavioural data that would allow us to estimate the body dynamics of participants on an individual basis. We namely aimed at achieving this for a task context with which human subjects are familiar from previous experience in daily activities. Consequently, such estimations are linked to the brain's existing representations of body dynamics and can be subsequently used as prior knowledge in the simulations of our motor learning framework.

In the second experimental session subjects were instructed to complete a series of object manipulation tasks. They moved inside the workspace used for the initial hand reaching experiment and were presented with the same series of home and target positions. Furthermore, subjects were required to complete the same stabilization phase so as to align their hand in the starting position before the initiation of each trial. As soon as stabilization was achieved, subjects were

introduced to new task dynamics in which the manipulated cursor that denoted the hand until that point, would start representing a virtual object. The object state was controlled by each subject's hand velocity and experienced the effect of predetermined spring and friction forces. Subjects had to move the object to the target as quickly as possible with 1 cm accuracy and stabilize it there for 0.3 s in order to complete a successful attempt. At the end of each trial visual feedback of success of failure (trial duration more than 10 s) was provided, as well as a quantitative measure of performance, estimated by a pre-defined cost function. As in the first experiment, the second session was completed after 120 counterbalanced successful trials (40 for each of the three trial types). A pilot experiment showed that task completion for the three selected target hitting sets was feasible with at least 25% success for the first 30 attempts. For other examined target hitting sets positioned at different angles and more distant areas of the workspace, this feasibility criterion was not met and pilot-subjects reported difficulty in employing the right hand configurations for task completion.

A modified version of the study described above was repeated on a second group of participants. In this version, all conditions of both experimental sessions were preserved, apart from the number of trials in the object manipulation paradigm. In particular, the number of trials for each of the three trial types rose from 40 to 120 and the object manipulation session was thereby completed after 360 counterbalanced successful trials. This long-training paradigm was employed to investigate subjects' performance and learning evolution for an extended period of time. It was also used to test the hypothesis that with greater experience subjects' behaviour converges closer to the predictions of an ideal actor model, which is based on full knowledge of task dynamics.

3.2.3 Modelling task dynamics

The examined system in the object manipulation paradigm is modeled as two coupled linear state space models that reflect hand and object dynamics respectively.

$$\dot{x} = A \cdot x + B \cdot u \tag{3.1}$$

$$\dot{y} = C \cdot y + D \cdot v \tag{3.2}$$

A, C constitute the state matrices and B, D the input matrices. The hand and object states, x and y respectively, represent position and velocity in Cartesian space. Their update depends on control signals u and v. We initially model the human upper limb as a two-link, two-degree-

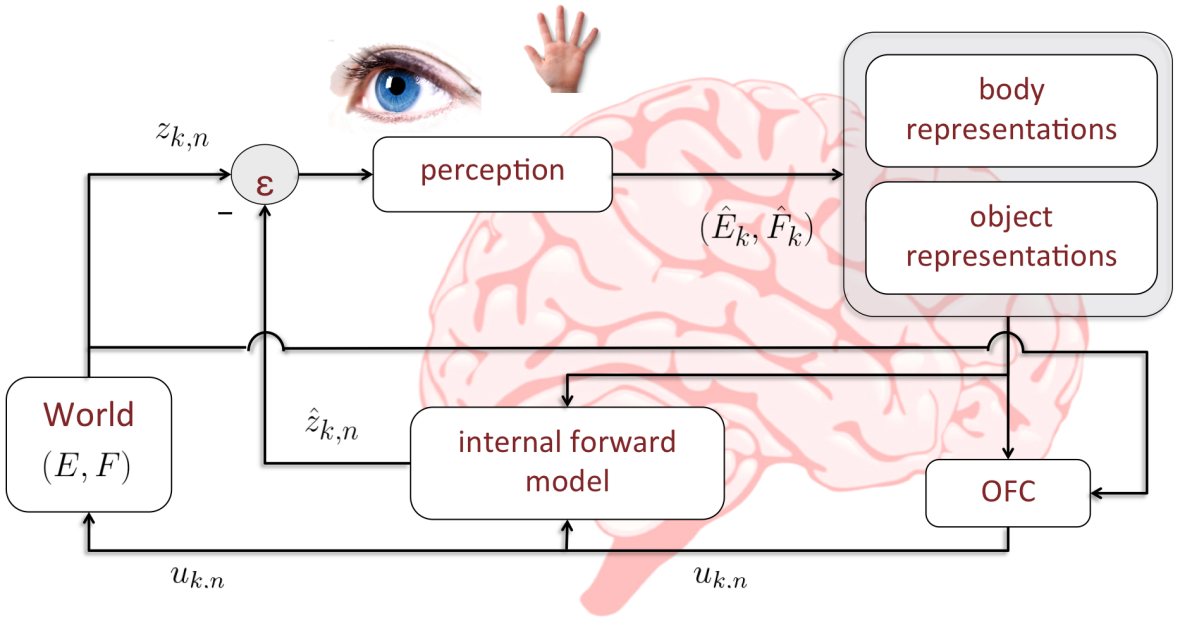

Figure 3.2: The representation of task dynamics refers to assumed body and object dynamics $(\hat{E}_k, \hat{F}_k)$ as well as the assumed task objective captured by a cost function that the experimental design imposes on the participants (equation (3.11)). The assumed task dynamics determine the form of the control signal $u_{k,n}$ for each trial k which is updated at each timepoint n, based on an OFC mechanism that optimizes the performance using a linear quadratic regulator (LQR). In turn, the produced control signal gives rise through an internal forward model (IFM) a prediction about the future system state $\hat{z}_{k,n}$. It also produces a motor command that leads to an interaction with the real world dynamics. The latter includes both body and object dynamics (E,F). The accumulated difference between the predicted $\hat{z}_{k,n}$ and produced system state $z_{k,n}$ throughout a whole trial forms an observed error ε, which can guide through sensory mapping an update of the assumed world dynamics.

of-freedom system and map it from the joint angle space to the end-effector Cartesian space. Based on this forward kinematics approach the hand state is estimated as

$$x_1 = l_1 \cdot cos(\vartheta_1) + l_2 \cdot cos(\vartheta_1 + \vartheta_2) \tag{3.3}$$

$$x_2 = l_1 \cdot sin(\vartheta_1) + l_2 \cdot sin(\vartheta_1 + \vartheta_2) \tag{3.4}$$

where ϑ_1,ϑ_2 and l_1,l_2 are the joint angles and lengths respectively. The latter are measured for each subject before the initiation of the experimental sessions. We are able to describe our task context with a linear dynamical system since movement during trial completion is performed in limited space range. For more expanded movement distances, our task context would be more accurately captured by a non-linear dynamical system. However, even in this case the

problem of learning the non-linear task dynamics could be methodologically broken down into a combination of smaller, locally linear system identification problems (Abramova et al. 2011), such as the one we formalise and use in the present study.

The object dynamics reflect the impact of a spring force and friction caused by drag (Fig. 3.1 C). They also depend on the force applied by each subject's right hand, which is proportional to the hand's velocity. The parametric definition of these dynamics is

$$C = w_C \cdot \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{L_{fric}}{M} & 0 & -\frac{L_{spr}}{M} & 0 \\ 0 & 0 & 0 & -\frac{L_{fric}}{M} & 0 & -\frac{L_{spr}}{M} \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{3.5}$$

$$D = w_D \cdot \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.6}$$

where $\frac{L_{fric}}{M} = 2$, $\frac{L_{spr}}{M} = 1$ the friction and spring coefficients normalized by the mass M and $w_C = 10$, $w_D = 40$ are state and input matrix weights. The latter were selected after preliminary testing with a broader range of values, some of which obstructed the feasibility of the task. To avoid oscillatory activity, object dynamics need to support a critically damped system. Namely the damping ratio ζ, which returns the system to equilibrium, requires

$$\zeta = \frac{L_{fric}}{2 \cdot \sqrt{M L_{spr}}} = 1 \tag{3.7}$$

The values for the normalized friction and spring coefficients are thus determined such that they respect condition (3.7). The coupling rule that enables a composite representation of hand and object dynamics is $v = \dot{x}$. Hand velocity thereby maps to a control force on the object state. Both this condition as well as the nature of object dynamics render an unintuitive task (3.1), which allows us to study learning from a naïve starting point.

Based on the coupling rule between body and object dynamics, (3.1) and (3.2) form a unified state-space body-object representation

$$\dot{z} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} A & 0 \\ DA & C \end{bmatrix} \cdot z + \begin{bmatrix} B \\ DB \end{bmatrix} \cdot u$$
$$= E \cdot z + F \cdot u \tag{3.8}$$

where z the 10-dimensional state vector that consists of the body and object dynamics information:

$$x = [x_1; x_2; \dot{x}_1; \dot{x}_2] \tag{3.9}$$

$$y = [y_1; y_2; \dot{y}_1; \dot{y}_2; {y_1}'; {y_2}'] \tag{3.10}$$

with $[{y_1}'; {y_2}'] = [y_1 - y_{1o}; y_2 - y_{2o}]$ denoting the offset of the object position with regard to the spring origin $(y_{1o}; y_{2o})$ and $u = [u_1; u_2]$ the 2-dimensional control vector. Body and object dynamics are represented in the brain as illustrated by Figure 3.2. These representations constitute approximations of the real world dynamics, whose update is driven by sensory prediction errors produced throughout a trial-by-trial modification of a motor-to-sensory mapping (forward model) (Mazzoni and Krakauer 2006).

3.2.4 Simulating optimal control and learning

In the present study we assume that motor outputs come about based on the brain's representations of world dynamics (body and object in our case). In particular, this process is realised via the internal feedback loop in Figure 3.2, in which the brain employs efference copies of continuous motor commands to predict the sensory state of the hand and object. These predictions are in turn compared against the consequences of acting inside the actual world dynamics. The mechanism is described using optimal feedback control (OFC) theory. In order to formulate an optimal feedback controller we define an objective function that captures the instructions of the examined task as quadratic costs:

$$J = 0.5 \cdot z^T(t^*) + \int_{t_0}^{t^*} (z - z^*)^T \cdot Q \cdot (z - z^*) dt + \int_{t_0}^{t^*} (u - u^*)^T \cdot R \cdot (u - u^*) dt \tag{3.11}$$

where t^*, z^*, u^* are the time point, system state and control signal at the target location

respectively. The first cost term reflects the need to stabilize the object at the target position, the second one penalizes large end point errors and the third one specifies a motor cost, where

$$Q = w_Q \cdot I^{10x10} \cdot [1;1;1;1;1;1;1;1;0;0] \cdot dt \tag{3.12}$$

$$R = w_R \cdot I^{2x2} \cdot dt \tag{3.13}$$

$$N = w_N \cdot Q/dt \tag{3.14}$$

with $w_Q = 800$, $w_R = 800$, $w_N = 100$ are the cost weights and dt=0.008 s is the discrete time step for which the state update is simulated. For the purposes of this study Q, R, N are kept fixed. The estimates of the objective function are optimized based on the effect of a linear quadratic regulator (LQR) (Kwakernaak Huibert/Sivan 1972). Here, due to the time constrains of the examined tasks, we implement a finite-horizon-discrete-time optimal control approach, in which the feedback gain is estimated iteratively backwards in time:

$$u_{n+1,k} = -K_{n,k}(\hat{E}_k, \hat{F}_k, Q, R, N) \cdot \ z_{n,k} \tag{3.15}$$

The feedback gain K and the motor command u are updated for each time step n within each trial k and throughout all trials of the experiment. It depends on the imposed objective function parameters Q, R, N and on the assumed system dynamics parameters $\hat{E}_k, \hat{F}_k$ which differ from the actual system dynamics E, F in (3.8) and are thus revised after every trial. The feedback controller (LQR) used to provide a solution to the set of linear differential equations that describe our system, is given by

$$K_n = (R + F^T \cdot P_n \cdot F)^{-1} \cdot F^T \cdot P_n \cdot E \tag{3.16}$$

and P_k is estimated backwards in discrete time by the dynamic Ricatti equation (Kwakernaak Huibert/Sivan 1972)

$$P_{n-1} = Q + E^T \cdot (P_n - P_n \cdot F \cdot (R + F^T \cdot P_n \cdot F)^{-1} F^T \cdot P_n) \cdot E \tag{3.17}$$

for terminal condition $P_{n^*} = Q$.

The selected motor command leads to an interaction with the world (real body and object dynamics (E,F)) and an update of the system state z. The error that steers the revision of the

brain's world representations $(\hat{E}_k, \hat{F}_k)$, records the observed deviation between this produced system state z the predicted system state $\hat{z}$ throughout each trial:

$$\varepsilon_k = \sum_{n=0}^{n^*} (\hat{z}_n(\hat{E}_k, \hat{F}_k) - z_n(E_k, F_k))^2 \tag{3.18}$$

where n^* denotes the time step at the target location. Prediction and action are determined based on equation (3.8) and equation (3.15) in the discrete time domain as:

$$\hat{z}_n = \hat{E}_k \cdot z_{n-1} - \hat{F}_k \cdot K_{n-1,k}(\hat{E}_k, \hat{F}_k, Q, R, N) \cdot z_{n-1} \tag{3.19}$$

$$z_n = E \cdot z_{n-1} - F \cdot K_{n-1,k}(\hat{E}_k, \hat{F}_k, Q, R, N) \cdot z_{n-1} \tag{3.20}$$

The error ε_k is used in a gradient descent approach to motor learning that determines the trial-by-trial estimation of the system dynamics parameters $\hat{E}_k$ and $\hat{F}_k$ as

$$\hat{E}_{k+1} = \hat{E}_k + n_1 \cdot \frac{\partial \varepsilon_k}{\partial \hat{E}_k} \tag{3.21}$$

$$\hat{F}_{k+1} = \hat{F}_k + n_2 \cdot \frac{\partial \varepsilon_k}{\partial \hat{F}_k} \tag{3.22}$$

where n_1, n_2 are the learning rates for the state and input matrix respectively, which are estimated throughout trials so as to best predict the gradually progressing human learning. The gradient descent implementation was selected as a biologically plausible foundation of learning. It in fact presents links to recurrent neural network approaches to adaptation that operate on the basis of the Hebbian rule, which in turn reflects the gradient of an error function.

3.2.5 Analysis

For the purpose of comparing model predictions and experimental data, the evolution of human and simulated learning was estimated (Fig. 3.6). Human learning was captured as the average of the root mean squared errors between each experimental object trajectory and the experimental object trajectories of the ten last trials. The latter were used to capture experimental end performance given that they constituted the final stage of the produced cost plateaus (Fig. 3.4). For the estimation of each root mean squared error, the updated object trajectory was resampled so as to consist of as many discrete points as the compared end performance trajectory. Similarly, simulated learning was described as the average of the root mean squared errors between each

simulated object trajectory and the experimental object trajectories of the ten last trials. As in the human data analysis, in every comparison the modeled object trajectory was resampled to match in points the end performance trajectory. Finally, the ideal observer performance was illustrated as the average of the root mean squared errors between the single model prediction and the last ten experimental object trajectories.

The difference between end performance prediction errors of the ideal observer model and the learning model was evaluated based on Friedman's test for statistical significance. Furthermore, R^2 estimations were used for each subject to infer the correlation between the simulated and experimentally acquired learning curves (Fig. 3.6).

3.2.6 Regression

The learning model was initialized with hand dynamics parameters A,B that were estimated for each subject based on a parameter estimation method for linear dynamical systems with inputs (LDSI) (Ghahramani and Hinton 1996). This method allowed us to develop a regression algorithm of our linear dynamical model against the experimental hand reaching data. The latter included (i) hand position and velocity, which determined the state vector and (ii) hand acceleration, which in proportionality to the emerging torque, determined the control vector. The initial assumptions on object dynamics parameters C_{init}, D_{init} were estimated per participant and target hitting set as the values that allowed the best fit between the learning model and the mean experimental object trajectory across the first 2 trials (using R^2 and $RMSE$ metrics). The learning rates n_1, n_2 were updated for each trial so as to enable a best fit of the simulated object trajectory to the corresponding experimental trajectory. Their update was performed using a non-linear least-square optimization method, available in MATLAB's optimization toolbox. The number of time steps in each simulation was determined in accordance to the selected dt value and the duration of each experimental trial.

3.2.7 Model comparison

In order to test the performance of our motor learning framework in predicting the transient dynamics of human motor learning, we compared its predictions to two common alternative frameworks, which also predict human behaviour during complex motor tasks. In particular, we compared the predictions of our policy learning model (PLM) to the predictions of an ideal actor model (IAM) which assumes complete knowledge of task dynamics (Nagengast et al. 2009). In this case, the brain produces the following control signal based on OFC:

$$u_{n+1,k} = -K_{n,k}(E, F, Q, R, N) \cdot z_{n,k} \tag{3.23}$$

where the assumed dynamics E, F represent the actual body and object dynamics. The IAM's prediction is thus a single estimation for the human performance at the end of a training session that has presumably allowed convergence to full learning of the task. We use this single end-performance prediction to further estimate the differences between IAM's predicted object trajectory and the experimentally measured object trajectories for each of the last 10 trials of each subject's training session. The average of these differences represents IAM's mean prediction error. The latter is compared for our PLM's equivalent mean prediction error, derived from the difference between PLM's simulated object trajectories and the corresponding experimental object trajectories for the last 10 trials of the training session.

We furthermore compared PLM to a model-free framework of learning (naïve model, NM), whose learning progression can be formalized by the movement error between consecutive trials. This progression is compared for the first ten trials of each subject's training session (primary learning period) against the evolution of PLM's prediction errors for the same trials. Since NM's estimates represent motor output differences between consecutive trials, we assign them to half-trial timepoints and subsequently we interpolate them to determine the estimates that correspond to the exact trial timepoints for which PLM's prediction errors are determined.

The rate of PLM's evolution is assumed to both capture convergence to the experimental data and the model's inherent learning mechanism. Consequently, if the learning mechanism posed by PLM describes the brain's learning principles successfully and captures human behaviour accurately, we expect the evolution of its prediction errors to converge faster to a learning plateau than NM's progression.

3.3 Results

3.3.1 Formalising the representation of a complex motor task

We describe the linear dynamics of a body and object system as two coupled state space models representing body and object dynamics respectively, according to $\dot{x} = A \cdot x + B \cdot u$ and $\dot{y} = C \cdot y + D \cdot v$, where A, C are the state dynamics matrices and B, D the input control matrices which govern the dynamics of the body x and the govern of the object y, respectively. Movements of the hand (body) act on the object, such that hand velocity translates into forces on the object ($v = \dot{x}$), equivalent to swiping along an object with slippery surface. Subjects have to control the object motion so that it reaches a defined target position. Subjects are rewarded at the end of each trial based on the object stopping at the target within in the allotted time and with minimal force inputs applied (end point error and control forces determine quadratic cost terms).

3.3.2 Experimental evidence of learning

Human subjects (N=15) were instructed to complete two subsequent sessions in the main experimental study. The first one aimed at providing information about subjects' prior knowledge on body dynamics. The second one tested subjects' performance in an object manipulation task that utilized the examined prior knowledge on body dynamics. Both of these sessions were repeated in a revised version of the main study on a different group of subjects (N=10) so as to test the generalisation of the main study's conclusions to longer training contexts.

In the first experimental paradigm we instructed the main study's participants (N=15) to perform hand reaching movements between starting and target positions that were displayed on a virtual reality screen at the level of their eyes. Subjects held an electromagnetic sensor, which traced their hand position and visualized it as a circular cursor. Three different sets of hand reaching movements were examined in a random order (Fig. 3.1 B). On the whole, subjects completed the experiment after 120 succesful trials (40 for each of the three different sets). A succesful trial required hitting the target quickly and stabilizing the cursor on the target for 0.3 sec. Visual feedback of success or failure was provided to the subjects after the completion of each trial.

In the second experiment, we instructed the same group of subjects (N = 15) to manipulate an object of unintuitive dynamics, which was simulated inside the virtual reality setup (Fig. 3.1 A). Subjects were asked to move the object within a given time-window between the start and target locations used in the first experiment (Fig. 3.1 B). The manipulated object was displayed as a circular cursor. The object's state was controlled by the participants' hand, to which they had no visual access. To control for order effects the experimental trials were counterbalanced in that the three target hitting sets were presented to the subjects in a random order. The use of different home positions for each of these different trial types, instead of a center-out motion task, was deliberate, in that it allowed us to investigate object manipulation mechanisms captured by a model of linear system dynamics (see Section 3.2). In contrast to that, the use of center-out object manipulation movements was, under the given experimental conditions, more prone to introducing body manipulation that would be better captured by a non-linear dynamical system. The number of tested trial types was confined to three after preliminary experimental evidence that target hitting sets positioned at greater distances to the body, challenged the feasibility of the task. The reason for this was that in such cases subjects had to perform object manipulation movements, which either required inconvenient body configurations or were exceeding the provided workspace.

At the end of each trial, subjects were presented with a visual feedback on their success (in case of fast completion) or failure (in case of slow completion or no hit at all). Although slow

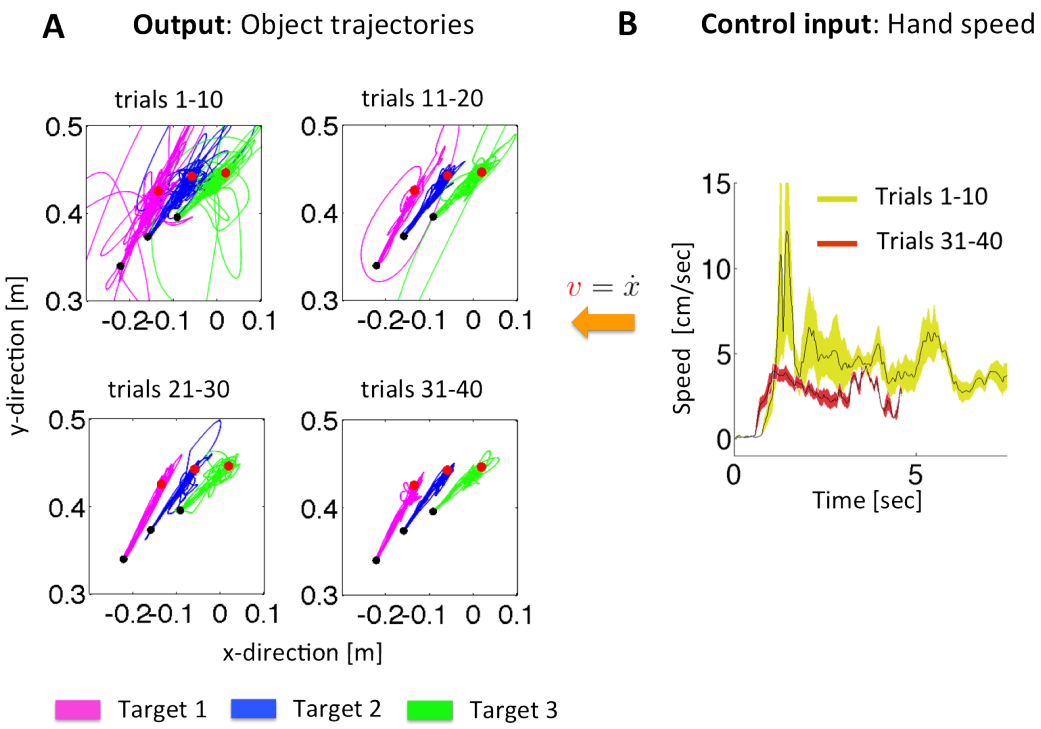

Figure 3.3: (**A**) Object trajectories throughout experiment (for one representative subject): The 2D object positioning is displayed for four consecutive phases of the experiment for three target hitting sessions (black dot: home, red dot: target). Subjects initiated the experiment with scattered movement patterns and eventually learned close to linear object trajectories. (**B**) Mean hand speed profiles estimated for the first (yellow) and last (red) ten trials 10 trials of the experiment (for each target hitting set). Subjects apparently converge to a bimodal hand speed profile that maps into control forces on the object state. The bimodal distribution of this speed profile illustrates the employment of a non-trivial motor control policy for the optimization of task completion.

attempts were disregarded as unsuccessful, subjects were allowed to move for a maximum of 7 sec within each trial, which provided them a longer time interval to familiarize themselves with the unknown object dynamics during the progression of a single target hitting attempt. Participants were also presented with a quantitative measure of their performance estimated by a predetermined cost function. The cost function formalized the instructions that subjects received before the initiation of the experiment. It thereby represented the objective to execute each trial as quickly and accurately as possible while avoiding immoderate energy consumption. Excessive movements were namely constrained with an effort penalty term increasing quadratically with the amplitude of the control signal. Additionally, successful hits required a stabilization of the object on the target for 0.3 sec. The investigated second experimental paradigm was completed after the successful execution of 120 trials (40 for each of the three different target hitting tests).

The unintuitive nature of the object manipulation task was primarily introduced through the

unknown object dynamics, which reflected the impact of friction and spring forces exercised on the system's state during hand movement (see Section 3.2). The origin of the applied spring force was a fixed point for each target hitting set, which did not coincide with the home or target location of the virtual object. The complexity of the examined tasks was received with increased motor variability during the initial period of the experiment.

This motor variability is evident in the first ten successful object manipulation trials for all three targets (Fig. 3.3 A,B), but also in the unsuccessful trials during the initial period of the experiment. It may reflect action exploration and thus facilitate motor learning as is suggested by recent studies (Wu et al. 2014) that pronounced a correlative link between motor variability and learning abilities. The dominance of exploratory behaviour in the beginning of the experiment is demonstrated in the fact the 96.67% of all failed trials in the study occurred within the period, which also contained the first ten successful trials (with 74.33% of the failed attempts even preceding the very first successful trial). The observed variability became more constrained while the subjects switched from unsuccessful to systematically more consecutive successful trials. During this process object trajectories became confined in a narrower motion range and utilized a smaller number of corrective movements to enable target approach. This transition appears to be in line with the selection of a structure of action that allows successful trial completion and can subsequently be optimized parametrically through an ongoing learning process. This is why, in the present study we interpret and model the portion of successful trials as behavioural data that unveils learning at the level of parametric optimization. We thereby assume a structure of action representation that has already been selected during the main body of unsuccessful attempts in the first phase of the experiment.

Figure 3.3 B suggests that throughout succesful trials within the object manipulation paradigm, subjects presumably tune their selected motor control strategy and converge to a specific hand speed pattern, which in turn is mapped into a control force on the object state. This end-performance speed pattern displays bimodality that suggests non-trivial motor control, in contrast to the bell-shaped distributions of velocity and speed profiles in simple hand reaching movements.

We further analysed behavioural evidence of motor learning in the gradual progression of the tracked subject performance. This was measured based on the values produced by the experimentally imposed quadratic cost function (Fig. 3.4). In all target hitting sets performance improved by reaching the fastest learning rate during the initial ten trials and afterwards converging in a slower tempo to a plateau that was maintained until the end of the experiment. While approaching this plateau (between trials 31-40) subjects converged from the initially spread-out movement to close to linear object trajectories observed in Fig. 3.3 A. Furthermore, the final learning outcomes commonly included loop end-movements right before goal achieve-

ment. Presumably, these loops allowed subjects to perform minor corrections on the linear object trajectory so as to refine its alignment with the target direction. Apart from higher learning rates, the first ten trials also revealed higher inter-subject variability. This variability decreased significantly throughout the experiment (also observed in Figure 3.4), a reason for which can be that subjects gradually converge to stereotypical behaviour.

3.3.3 A policy learning model, which familiarizes the brain with unknown world dynamics on a trial-by-trial basis

Based on assumptions about the body, the object and task dynamics, we posit that the brain exploits a control process regulated by a feedback gain that acts on the system's state to optimize a quadratic cost function. A forward model estimates the future sensory states using an efferent copy of the motor commands. An optimal controller produces the motor command that allows the body to perform the task. Experiencing the real system dynamics allows a comparison between predicted and actually produced action. Differences are mapped as errors through the sensory system and steer an update of the assumed system dynamics in a closed action-perception

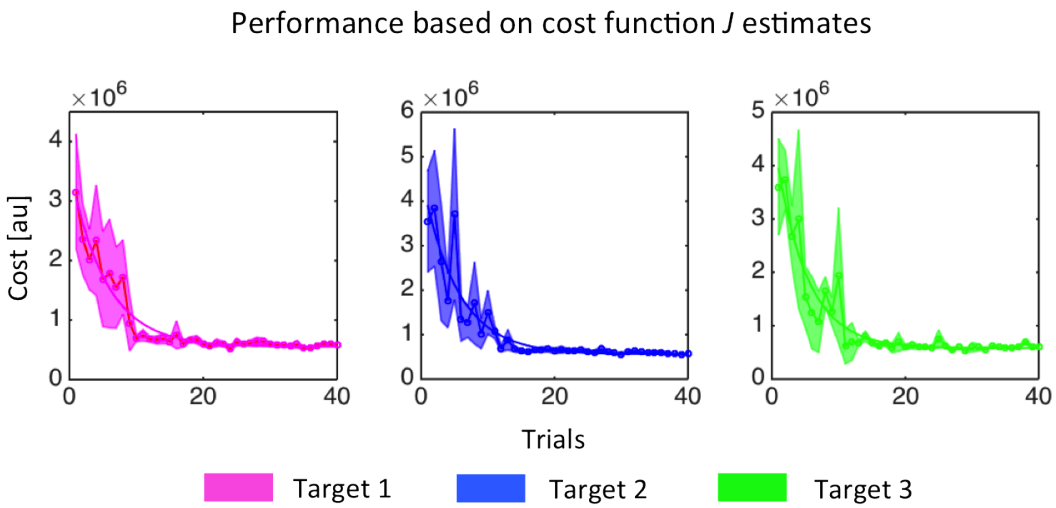

Figure 3.4: Performance evolution: Subjects performance is quantified with the help of a cost function (see equation (3.11)) imposed on the subjects during the experiment through visual feedback at the end of each trial. For all target hitting sets the costs converge to a plateau approximately after the first ten trials. Essentially subjects optimize the estimate of J that is presented to them as cost that needs to be minimised after the completion of each trial. They thereby optimize their performance in the instructions provided to them before the initiation of the experiment. These instructions include the minimisation of the distance to the presented targets as well as an end-point and control penatly, all of which are formalized in the cost function used for performance feedback within the actual experiment. Here, we observe that the evolution of this performance in cost estimates reveals learning which is consistent across subjects.

loop. Our approach thus describes motor learning as error-driven gradient descent steps in the space of unknown action dynamics parameters. Essentially this mechanism is repeated trial after trial and promotes the progressive adaptation to unknown system dynamics, which in turn updates the feedback gain parametrically and thereby supports learning at the control level.

Based on the observed progression of human motor performance for the whole of the experiment, we chose to approach learning throughout successful trials as a mechanism of parametric optimization. Our optimization applied to the assumed hand and object dynamics (see Section 3.2) as well as to the optimal feedback control policy which guided the system's predicted and produced motor outputs. We essentially formulated a locally linear system identification process that updated both system dynamics and control strategies. The update of the brain's assumptions on body and world dynamics was realized as a gradient descent process on state and input matrices. This process was determined by the root mean squared error between the predicted and produced object trajectories after they were rescaled to the same number of steps. In order to simulate motor learning we assumed that at the beginning of the experiment subjects possessed some prior knowledge of their body dynamics. This knowledge is presumably acquired through life-long experience on hand reaching movements. In order to quantify this prior knowledge, we estimated the parameters of each subject's body dynamics (A, B) for hand reaching movements in each of the three examined target sets from the behavioural data acquired in the first experiment (Fig. 3.5 A). In particular, we estimated individualized body parameters (Fig. 3.5 B) with a regression method (see Section 3.2.6) that searches for the best fit of the related state space model to the some provided states and control inputs (Ghahramani and Hinton 1996). The provided states are in this case the Cartesian hand position and velocity, while the control is determined by the torque emerging from hand acceleration.

Apart from the prior knowledge on body dynamics, we determined some initial guess on object dynamics for each participant, by estimating the parametric combination of the state matrix C and the input matrix D, which allowed a best fit of our motor learning model to human performance at the onset of the successful trials. We thus assumed that by the time the brain starts driving successful trial completion in the context of the instructed complex motor task, some partial prior knowledge of the body and world dynamics has already been established. This prior knowledge is estimated per individual and guides the simulation of each subject's motor behaviour throughout the experiment. In the trial-by-trial progression of the simulated object trajectories we distinguished features equivalent to the pattern of learning observed in the experimental data. In fact, given the individualized parametric adjustments of the motor learning model, these features were initially investigated in the predicted behaviour of each subject. In all cases the produced object trajectories evolved from an evidently unsystematic, overshooting motion pattern at the beginning of successful trials to an eventually close to linear

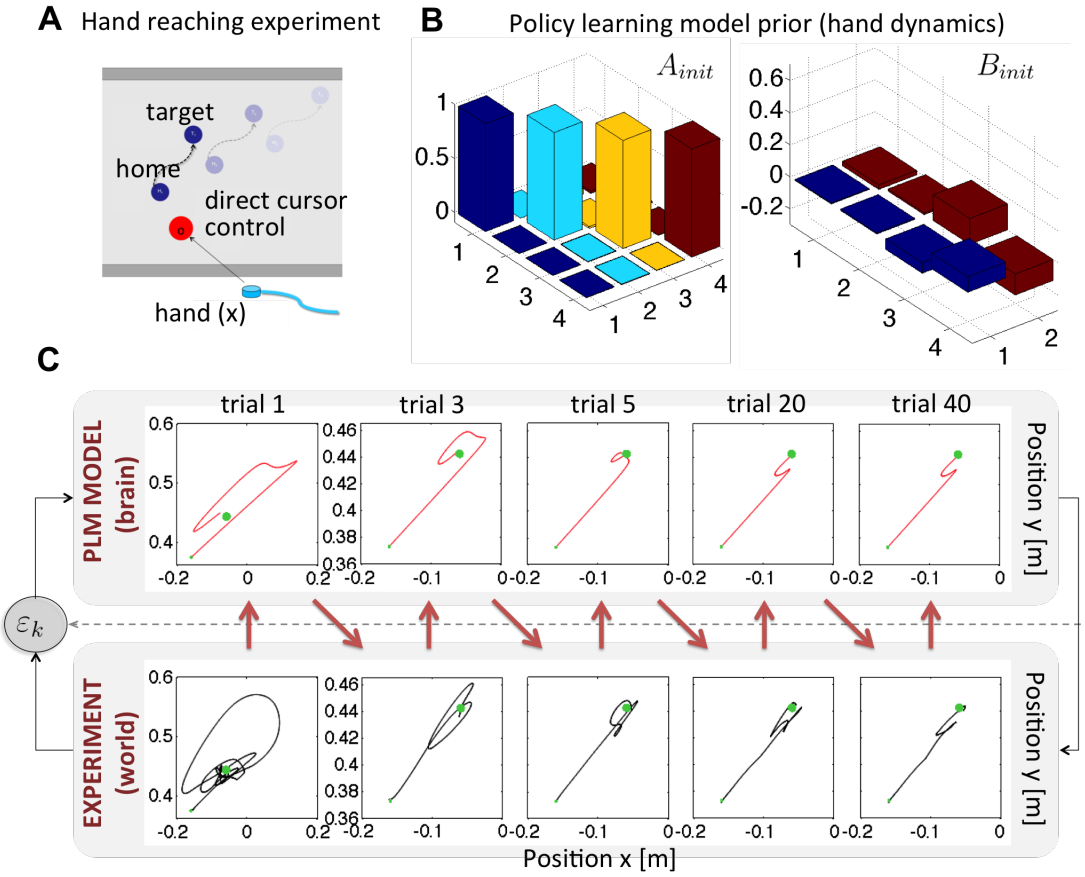

Figure 3.5: (**A**) A preliminary experimental session on simple hand reaching movements allows the acquisition of behavioural data for the same workspace and set of targets as the ones used in the subsequent object manipulation paradigm. (**B**) This behavioural data is used for the estimation of hand dynamics (state matrix A and input matrix B), which is assumed as pre-acquired knowledge of our motor learning model. The figure illustrates the estimation for one representative subject and target set. (**C**) Predicted object trajectories for the same representative case: Our policy learning model (PLM) predicts object movement that follows the human learning pattern (from spread out trajectories to close to linear object movements). Similarly to humans' end performance, the model completes each trajectory with a loop, which presumably allows the agent to decelerate before hitting the target. The estimated prior knowledge of A and B initializes the simulated behaviour at the beginning of the experiment. Each simulated trial is based on assumptions about the unknown world dynamics parameters (E,F) that are determined by an error between the previously predicted action and the actually produced object trajectory.

target hitting path. In fact, the object trajectory to which the motor learning model converges constitutes a straight route with an S-shaped ending right before the target that is consistent with the loop end-movement observed in the human trajectories (Fig. 3.5 C). Notably, the orientation of the loop varies within the final successful trials of the experimental data, indicating potential redundancy in the policy that allows task completion. Another explanation is that by the end of

the instructed number of trials, subjects have achieved partial -as opposed to complete- learning of the task dynamics. They may therefore engage in further exploratory behaviour that affects essentially the final part of the object trajectories.

3.3.4 Predicting human behaviour and the scenario of partial learning

We examined the correlation between updated human and simulated behaviour. In particular, we compared the evolution of the average object trajectories as measured in the experiment and the average simulated learning process (see Section 3.2.5). The average was estimated across subjects and target sets. The human learning process was determined for each individual by the updated root-mean-squared error between each experimental object trajectory and the experimental end-performance object trajectories. We considered as end-performance the human data acquired from the ten last successful trials, in which subjects had reached a plateau in their achieved scores (Fig. 3.4). The progression of the simulated behaviour was displayed as the updated root-mean-squared error between each predicted object trajectory and the experimental end-performance object trajectories. In Figure 3.6 we noted a strong correlation between the human adaptation rule and the learning process as reflected by PLM (mean R^2=0.899). In fact, our motor learning approach followed human adaptation on a trial-by-trial basis, despite the more disorderly evolution of the human trace, which can be attributed to noise that is not captured by our motor learning model.

The end-performance predictions of our motor learning model were also compared to the predictions of an ideal actor model (IAM). The latter is build upon a formalization of the structure of task dynamics and control policies identical to the one proposed by PLM. However, it assumes complete initial knowledge of the real body and object dynamics and can thus predict final learning outcomes without accounting for any experimentally observed changes in human performance. Essentially, the ideal actor model predicts end-performance by sidestepping the learning process. We noticed that the ideal actor end-performance prediction displayed a larger error to the human end-performance than the motor learning model (Fig. 3.6 A). Conclusively, the policy learning model performs better in capturing the behaviour of subjects at the end of the experiment than the ideal actor model. This triggers a reasonable hypothesis that the brain requires long training periods to fully learn the dynamics of complex motor tasks. For shorter exposure to training, the brain manages to acquire a partial knowledge of the task dynamics. This partial knowledge can therefore not be succesfully captured by the commonly used IAM (Nagengast et al. 2009), since the latter determines control policies assuming full knowledge of the world dynamics.

We tested the generalisation of our conclusions to a long training paradigm, similar to the design of the main study (simple hand-reaching Experiment A and object manipulation Experiment

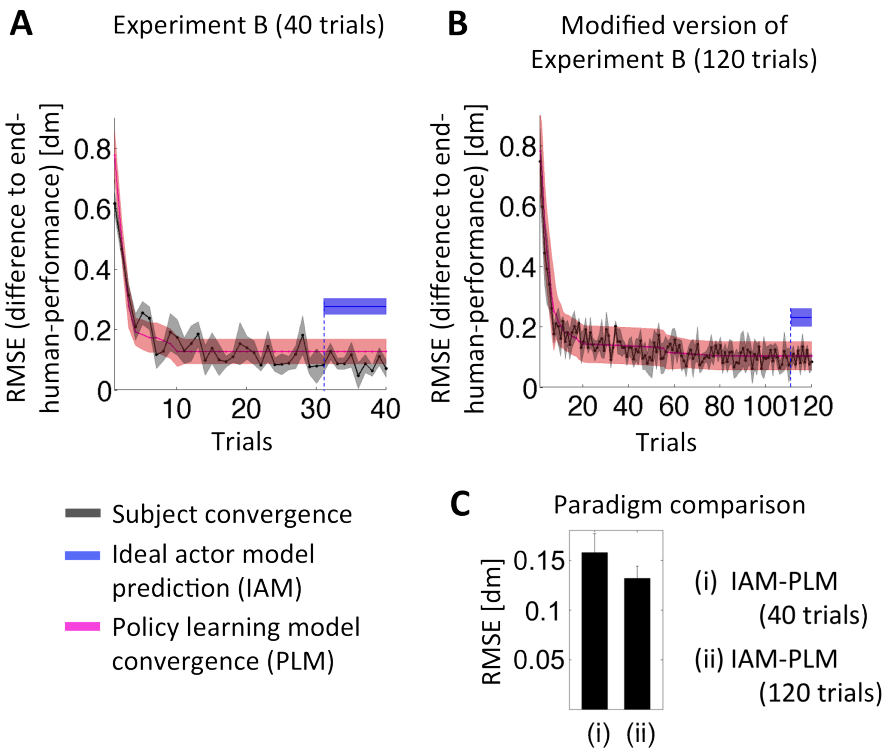

Figure 3.6: (**A**) The proposed learning rule (PLM) captures human learning, namely the temporal dynamics of learning the assumed task parameters on a trial-by-trial basis. The RMSE in the plot captures the difference between (i) PLM's updated object trajectories and the experimental end-performance object trajectories [pink] as well as (ii) the updated experimental object trajectories and the experimental end-performance object trajectories [black] (see Section 3.2.5). Interestingly the difference between the end-performance prediction of an ideal actor model (based on complete knowledge of the world dynamics) and the human end-performance trajectories is higher than the end-performance prediction error of our learning model (PLM). Our model thus captures the aspect of partial learning within the given time scales of experience. (**B**) Learning model performance for the overtrained subjects. The figure reflects convergence of the human data and the learning model to a plateau closer to the end-performance prediction of the ideal actor model. (**C**) The convergence is verified by the observed decrease in the difference between PLM's and IAM's end-performance prediction errors at the long training paradigm. In all cases end-performance is captured by the mean motor output (object trajectories) of the last ten trials of each session and target set.

B), in which the object manipulation session was extended to 120 trials (from the original 40) for each target hitting set. Also here, PLM predicted succesfully the temporal dynamics of human learning on a trial-by-trial basis (Fig. 3.6 B). We additionally observed a convergence of the human data and PLM's predictions to a plateau closer to the end-performance prediction of IAM (Fig. 3.6 C). This convergence validates our interpretation of PLM's predictive superiority in the short training paradigm as the model's success in accuratetly capturing the aspect of partial learning during the familiarisation with a complex motor task. As PLM predicts in the

long training paradigm at the level of individual motor outputs, human subjects eventually learn object trajectories, which are slightly different to the end-performance trajectories of the short training sessions (Fig. 3.7). The close-to-linear shape of these end-performance trajectories is the result of a further tuning of control signals linked to bimodal speed profiles such as the ones observed in the previous experiments (Fig. 3.3 B). This suggests that in the long training paradigm too, subjects learn non-trivial motor control, as opposed to the control patterns employed in movements with normally distributed speed profiles (e.g. simple hand reaching movements).

3.3.5 The policy learning framework outperforms an ideal actor and a naive model of motor learning

Aside from evaluating our motor learning framework's predictions against human data, we also tested PLM's performance by comparing it to two alternative frameworks, which are commonly

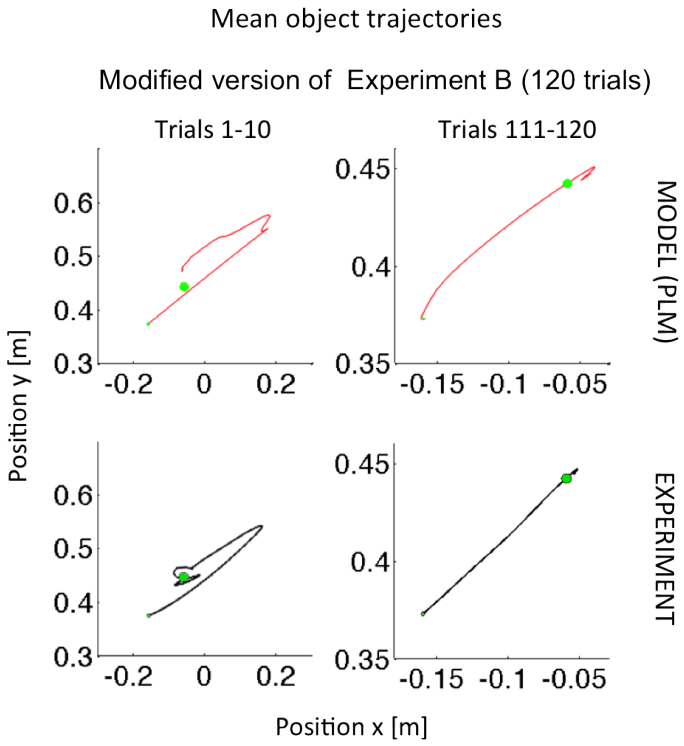

Figure 3.7: The mean object trajectories for the first and last ten trials in the long training paradigm reveal consistency between experimental and simulated data (trajectories displayed for one representative subject and target hitting set). In fact, the motor learning model captures an evolution of object manipulation, which differs from the end performance of our shorter training paradigm of 40 trials per target. This suggests that our motor learning framework successfully captures an ongoing optimization process in the brain which in the given task context continues to evolve with longer exposure to task execution.

employed to predict human behaviour during complex motor tasks. As already noted in Figure 3.6, we compared the predictions of PLM to IAM, which assumes full knowledge of task dynamics (see Section 3.2.7). The predictive inferiority of the ideal actor model, which we verify in Figure 3.8 C, can be explained if humans converge to partial knowledge of the actual task dynamics within the given experimental timescales. The aspect of partial learning can thus be recorded more successfully by a learning mechanism, which follows the update of body and object dynamics as well as the control signal throughout the instructed experimental trials.

Furthermore, we compared PLM to a model-free framework of learning (naïve model, NM), whose learning progression is defined by the motor output error between consecutive trials. Thus NM's estimates were assigned to half-trial timepoints and subsequently interpolated to determine its values at the exact trial timepoints for which PLM prediction errors are approximated. Our hypothesis is that PLM's evolution reflects both convergence to the experimental data and the model's inherent learning mechanism. In contrast to this, NM simply reflects the trend of the experimental data throughout trials. In Figure 3.8 A,B we notice that the evolution of PLM's prediction errors converges faster to a learning plateau than NM's progression during the primary learning period of the training session (first ten trials). This verifies our hypothesis that the learning mechanism posed by PLM describes the brain's learning principles successfully.

3.4 Discussion

We have shown that in unfamiliar task conditions the brain makes continuous decisions for the generation of complex motor outputs by learning optimal feedback control based on the steady identification of unknown parameters of both body and object dynamics. This process manifests itself as an error-based learning mechanism that exploits the brain's inherent link between forward models and feedback control to compute dynamically updated policies. We presented experiments, which verified that our theoretical motor learning framework successfully predicts the temporal dynamics of human motor learning in an unintuitive object manipulation task on a trial-by-trial basis. Importantly, our findings show how the brain simultaneously learns control policies and task representations, thereby expanding previous perspectives of motor learning as a pure adaptation of task related parameters (Berniker and Kording 2008).

Our approach also departs from previous views on control policies in motor behaviour, which assume that the goal of the motor system is to produce straight, minimum-jerk trajectories (Hogan 1984; Berniker and Kording 2008). We argue that rather than aiming at the preservation of a desired trajectory, the nervous system appears to act so as to satisfy a number of performance criteria, which can be even optimised under conditions of behavioural variability (Todorov 2004; Wu et al. 2014). In fact, variability appears to accumulate in task-irrelevant dimensions in

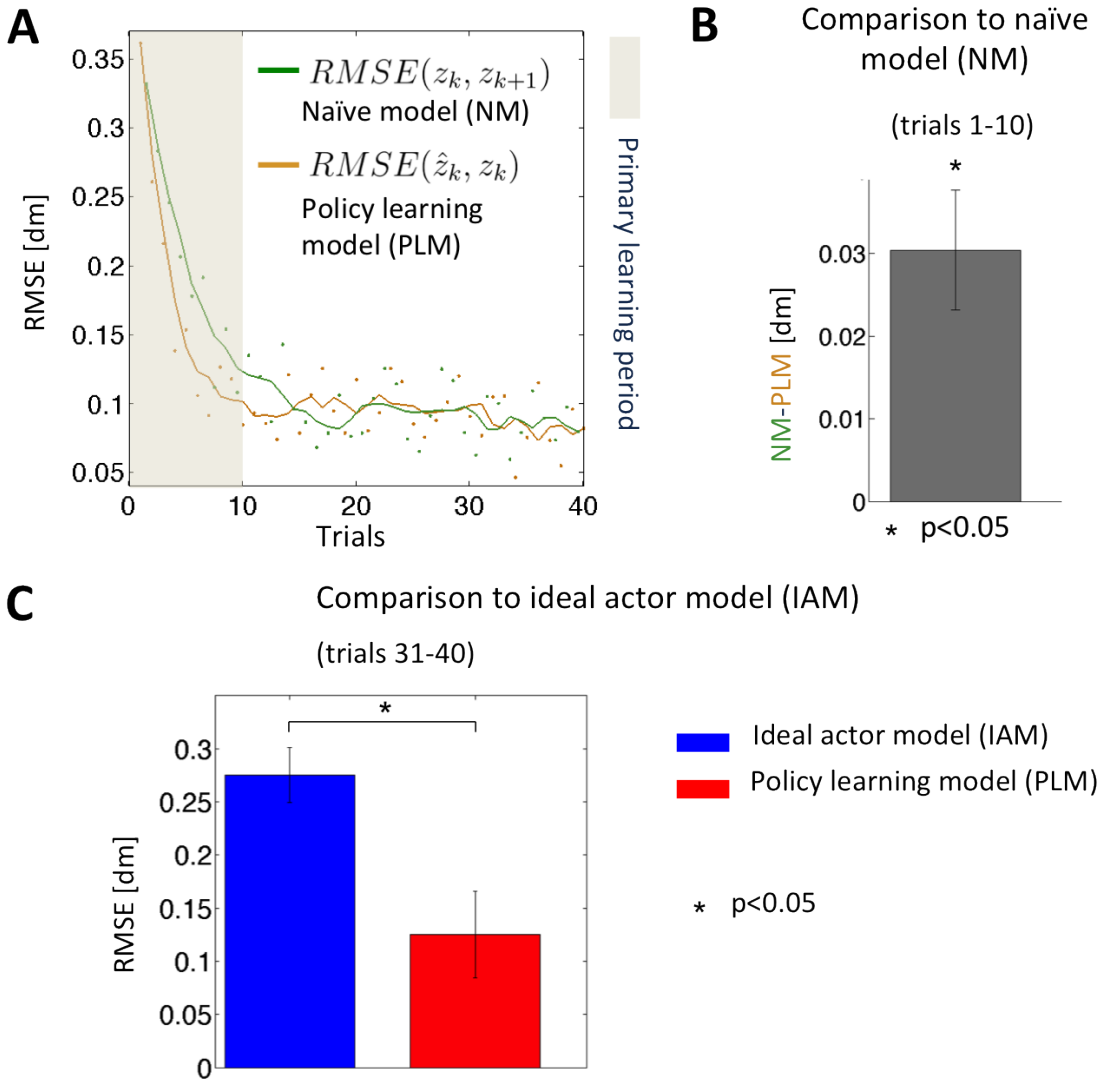

Figure 3.8: (**A**) The evolution of PLM's prediction errors, estimated as $RMSE(\hat{z}_k, \hat{z}_{k+1})$ throughout trials, is compared to the evolution of a naive model's learning progression, estimated as $RMSE(\hat{z}_k, z_k)$ (see Section 3.2.7). For the estimation of PLM's and NM's updated $RMSE$s we only use the object position from the system state, z or $\hat{z}$. (**B**) For the first 10 trials, which represent the primary learning period based on cost minimisation (Fig. 3.4), we estimated the difference between PLM's and NM's progression in Figure 3.8 A and verified that it is significantly different than zero across subjects and target hitting sets. (**C**) PLM produces a significantly smaller prediction error than IAC for the end-performance (last 10 trials).

agreement with the 'minimal intervention' principle, that is interpreted as underlying mechanism for motor synergies (Todorov and Jordan 2002; D'Avella and Bizzi 1998; Santello and Soechting 2000). Our policy learning model (PLM) is consistent with this phenomenon, in that instead of

a target trajectory, which penalises relative deviations, it uses a set of task objectives, based on which the brain supports motor outputs.

We are furthermore expanding past studies of optimality principles in sensorimotor control, which focused on predicting performance at the end of experimental sessions (Nagengast et al. 2009). Such an approach defines a version of optimal control policies that is based on full knowledge of task and world dynamics. It thereby sidesteps the aspect of motor learning. This ideal actor model assumes that complete learning has been achieved at the end of the investigated paradigm. While in some simple task contexts (e.g. hand reaching movements) this might be the case, in many complex motor tasks encountered in naturalistic environments, learning the unknown dynamics evolves over longer periods of time, which for practical reasons cannot be covered in one experimental session.

Our framework thus examines aspects of motor learning that do not exclude the possibility of achieving partial knowledge of task dynamics in some given experimental timescale. This scenario can in fact be expected for our experimental design, which imposes on human subjects an initially unknown and unintuitive task context. In this context we observed an extended phase of familiarization until the naïve participants managed successful trial completion. Arguably, the subsequent phase of optimizing the completion process might require a considerably large amount of trials, which is only partially covered in the present study. Our model predicts accurately the gradual progression of human learning for the provided amount of trials and outruns an ideal actor model (IAM) in capturing end-performance. We also verified our anticipation that both human behaviour and the policy learning model (PLM) converge to the predictions of an ideal actor model (IAM) for longer timescales of training.

At present our approach has proposed an underlying mechanism for the gradual progression of performance in an unfamiliar task environment. An interesting extension of the study would be to test our model in experiments that introduce sudden perturbations in the learned task dynamics. In fact, recent work has examined the principles based on which the brain reacts to unexpected disturbances during hand reaching movements (Braun et al. 2009a). Yet, although this study explains policy switches to such changes in task dynamics, it focuses on interpreting adaptation at the end-point of learning. It thus does not account for the evolution of control strategies up until the moment that the brain establishes an optimal way of altering feedback gains when the target or force-field within the trial are altered.

We used a linear dynamical system description to determine the control problem in our examined paradigm. This approach stands in agreement with our task design (e.g. constrained workspace), but was also largely inspired by recent theoretical work in optimal control theory. Attempts have been namely made to describe the control of systems with complex, non-linear

dynamics by decomposing it into a selection of locally linear control strategies (Abramova et al. 2011). Such a hierarchical learning infrastructure enables efficient solutions to the complexity and dimensionality problems encountered when a non-linear control problem is learned monolithically. In fact, this algorithm poses in our view, a candidate for the action selection mechanism that leads to the formalization of the linear controller proposed and updated by our model. Our framework can thereby be regarded as the optimisation of a control policy, the form of which has already been decided upon during an initial structure learning phase. We considered this phase to be mainly represented by the unsuccessful trials at the beginning of the experiment, but did not account for its root principles, since we only investigated successful trials in our simulations.

In line with the above, our framework is based on the narrative that the brain extracts through experience the common underlying structure of variable tasks (Braun et al. 2010). This drives, in turn, structure-specific facilitation and further tuning of the motor control process in future trials of similar tasks (Braun et al. 2009b; Acun 2010). This principle appears to be more suggestive of a mechanism of model-based learning, consistent with the model-based approach to learning that we designed and validated empirically.

A future extension of our approach can thus merge our current version of policy learning at the level of system identification and parameter optimisation with a symbolic action selection process that will determine the structure of controllers (Abramova et al. 2011) and body/world representations. This direction renders a synthesis of cognitive and execution-level decision making processes in motor behaviour. Its implementation can incorporate the role of motor variability as an aspect of behaviour that enables exploration and consequently enhances learning. The concept of learning through exploratory variability simultaneously to exploiting already existing, albeit imperfect, knowledge holds a key position in reinforcement learning theory. Our behavioural findings appear to be consistent with it, in that they display decreasing, but persistent motor variability in human behaviour throughout the completion of the experiment.

Interesting links can be drawn between our learning formalization and potential neurophysiological correlates. In particular, our approach describes motor learning as gradient descent steps in the space of unknown task dynamics parameters. This mechanism is driven by an error mapped by the sensory system and a candidate principle for its implementation at the neuronal level is the modification of synaptic weights via Hebbian learning rules.

Crucially, the present study proposes a framework that predicts accurately the gradual progression of human learning under unfamiliar task conditions. This process expands previous work on motor adaptation in task dynamics by merging it with a newly proposed mechanism of policy updates. Our results suggest that the brain employs simple learning rules to support decisions implemented by near optimal control in complex object manipulation tasks. The policy learning

framework (PLM) provides thereby an algorithmic formalization, which can guide further experimental investigations on the neural foundation of cortical action selection and motor learning rules.

4 Decoding human motor control strategies in whole-body manipulation tasks

4.1 Introduction

In Chapter 3 we formalized and validated a framework that describes the temporal dynamics of motor learning during a complex motor task. This framework captures how the brain produces complex motor outputs by learning optimal feedback control based on the system identification of composite world representations (see Chapter 2 and Sylaidi and Faisal 2012). This approach thereby predicts the gradual improvement of a control policy, the form of which has been selected at the initial phase of the experiment. An interesting question raised is how the brain manages switches to new forms of control policies during learning. In other words when does the brain support exploratory behaviour. The mechanism that underlies the latter process in a motor control context remains unclear.

There is to date substantial work in neuroscience that provides strong support for the hypothesis that the CNS learns and maintains internal models of sensorimotor control (Morasso et al. 1999; Wolpert et al. 1995; Mehta and Schaal 2002; Kawato and Wolpert 1998; Wagner and Smith 2008; Cooper 2010). Such models constitute neural systems which imitate the behaviour of the sensorimotor system and objects in the external environment. They enable the brain to predict the consequences of action and produce the motor commands required to perform specific tasks. Wolpert et al. 1995 have investigated the existence of such a model of motor control in a sensorimotor integration task in which human subjects estimated the positioning of one of their hands at the end of movements made in the dark and under conditions of externally imposed forces. In particular, the framework used to examine and predict the produced errors in this task was a combination of two processes. The first one reflected the properties of a forward model, which employs the motor command and current state estimate to predict the next state estimate. The second one constituted a simulation of the sensory processing system to determine the sensory feedback from the current state estimate. The correction of the state estimate was achieved based on the error between predicted and actual sensory feedback. The framework was proven able to parsimoniously reproduce the human end-performance, thereby providing direct

support for the existence of an internal model. Subsequent work has advanced the hypothesis of internal motor control models by proposing computational frameworks of multiple internal models (Kawato and Wolpert 1998).

Further studies have employed internal models to explain motion planning and control. Flanagan and Wing 1997 examined the ability of human subjects to predict variable hand-held loads by testing grip force adjustments used to steady the load in the hand during arm movements. They found evidence that the CNS achieves motor control by predicting the kinematics of hand movement and the corresponding load force and thus based their interpretation of behavioural data on an internal model of the motor system and the external load. Milner and Franklin 2005 looked into the neuromuscular apparatus supporting the initial phase of adaptation to unfamiliar dynamics. Subjects were exposed to destabilizing velocity-dependent force field conditions and their responses to the encountered disturbances were measured. The findings provided evidence for the modification of muscle activation patterns in anticipation of the force field effects. These anticipatory steps suggest that the CNS responds to novel disturbances in the task dynamics by a synergy of viscoelastic impedance control and an internal model of dynamics.

Nevertheless, despite the extensive work on formalizing and predicting motor control processes, motor neuroscience is still missing a general framework that explains how the brain shifts to new control strategies during the learning process. The mechanism of exploring new control policies has on the other hand received more attention in cognitive neuroscience; namely how adaptive agents manage a trade-off between exploration and exploitation. In fact, it is assumed that the brain uses sets of different mechanisms to address families of problems spanning over different timescales (Cohen et al. 2007). Still, there is a lot of speculation on the emergence of exploration. Exploration can be either random (in unknown environments) or structured on sophisticated heuristics and thus based on advanced cognitive control (in static/well structured environments). However, it can be assumed that even for realistically complex, known environments (recreated in experimental conditions), strategies of structured exploration, which appear to be optimal, are likely to be computationally intractable for biological systems. Thus, stochastic (random) exploration may be more biologically realistic.

Mechanisms of exploring control policies have been examined based on tools derived from Reinforcement Learning (RL) theory (Kaelbling et al. 1996). Traditionally, models in RL act by strengthening associations between stimuli and actions when these are related to a reward (Sutton and Barto 1998). In fact, experimental research provides evidence that this strategy might be supported by the dopaminergic (DA) system (Montague et al. 1996). Simulations on the other hand, show that this progressive strengthening of associations makes agents robust and efficient in stationary environments, but resistant to change in non-stationary environments. For instance, when an initial association is learned rapidly and strongly (with large weights), it is

then difficult for the agent to adjust to a significant change in the received stimuli. A solution to this problem is to introduce an annealing mechanism, based on which noise is added to the system when new learning is required. This noise allows a random exploration of new associations. As the agent starts strengthening (and thus learning) newly rewarded associations, noise is progressively reduced. Models such as the latter have not been tested in motor behaviour, where the role and emergence of exploration in general, as well as stochastic exploration in particular, remain unclear.

4.1.1 Balance as a motor control paradigm

Here, we set out to examine the control process in a complex motor task, to evaluate the stability status of the learned controllers and to spot potential evidence for exploratory behaviour. We selected and designed a balance control task, since human balance can serve as a good context for applying modeling to specific problems in motor control (Kuo 1995). In fact, balance control primarily in human upright standing has been a popular experimental paradigm (Nashner and McCollum 1985). Furthermore, the regulation and stabilization posture has been featured as a complex control problem that requires the multi-modal integration of sensory data and the compensation of transmission and processing delays (Morasso et al. 1999). This highlights balance as a control mechanism, at least as complex as a number of other motor paradigms that involve coordinated movements with many degrees of freedom (e.g. eyes, arms and hands). It also contradicts the notion that standing posture can be perceived merely as the initial state of walking and can therefore be considered a 'simple' task.

4.1.2 Core contributions

In the present chapter we develop a generic computational approach to formalise the control process in a balancing task. Additionally, we design a methodological platform for investigating whether behaviour in the examined paradigm converges to a stable motor control regime or instead displays switches to new control policies. This tool allows us to link partial learning to the potential persistence of exploratory behaviour.

4.2 Aims and Methods

The scope of this chapter is to investigate a body balance task as a motor control paradigm. Within this context we aim at elucidating the stability of the learning process; that is whether the brain switches at certain points to new control policy forms instead of persisting on exploiting and improving already selected motor control strategies. We thereby move towards a direction of expanding our investigation of the motor control process beyond the formalisation of the latter

in Chapter 3. To this end we designed and carried out an experimental study on human balance control inside a virtual reality setting.

4.2.1 Experimental setup and data acquisition

Participants stood on a force plate (Nintendo Wii-board) with four pressure sensors capturing a user's center of weight. We designed and implemented a software and hardware interface between the force-plate and a processing/visualization unit (Fig. 4.1). Measurements from the four pressure sensors where acquired at a frequency of 500 Hz. They were used to capture the shifts in a subject's center of weight which were subsequently translated into the 2D displacement of a cursor on a computer monitor. The cursor was a black sphere moving on a blue background while the target was presented as a red sphere. The monitor was located 95 cm in front of the participant. The design of the experiment and the necessary visual feedback were programmed using a custom made motorlib C code library and opengl. All acquired motion data was analysed using MATLAB (The Mathworks Inc., Natick, Massachusetts, USA).

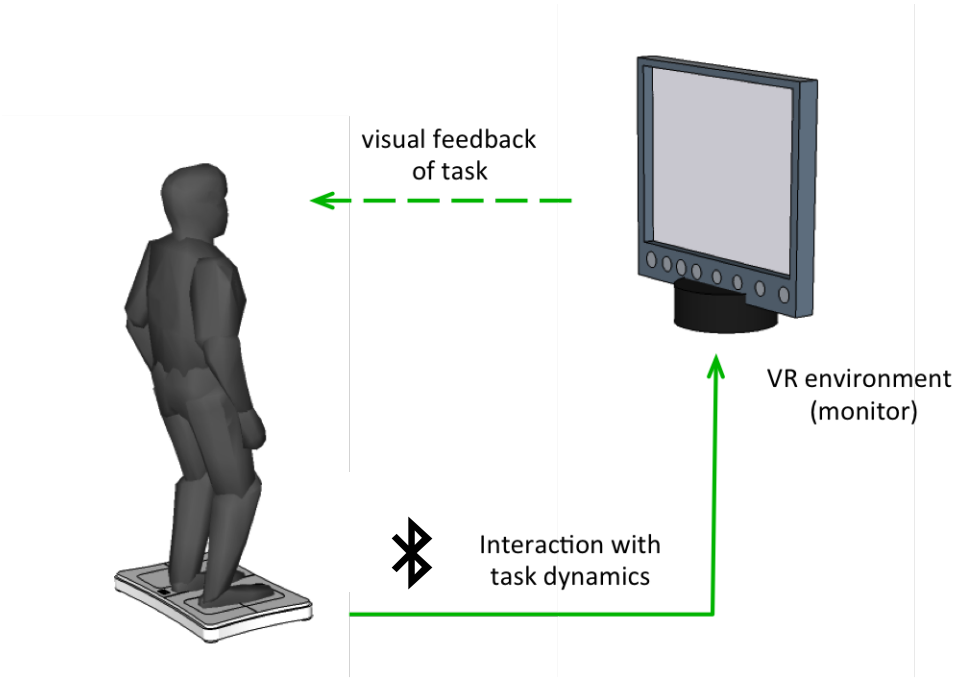

Figure 4.1: Human subjects stand on a force plate with four pressure sensors. Their balance shifts are tracked by the sensors and mapped onto the movement of a cursor on the visualising/processing unit. The instructed virtual task involves moving the cursor to a specifc target location. Participants receive visual feedback of their performance which is processed by the sensory system and in turn updates the world state estimation (see Fig. 4.2). The feedack fulfills the real-time closed-loop interaction between subjects and experimental setup.

4.2.2 Experimental protocol

Six adults volunteered for this study (2 women; 4 men). The participants' mean age and mass were 25 years and 78 kg respectively. The subjects had no known disabilities. They were naive to the purpose of the experiment and all provided informed consent consistent with the policies of the Imperial College London Ethics Committee.

Participants were instructed to perform balancing body movements on the force-plate in order to control the motion of a cursor on a PC screen. In specific, they were asked to make the cursor hit a displayed red target within a given timeframe and as accurately as possible. During the experiment subjects were allowed to slowly lean forward and backward, left and right as far as possible while keeping both feet completely on the ground.

Both the selection of a force-plate and the definition of the dynamics, that mapped weight balance to cursor motion, were intentionally aimed at facilitating non-intuitive yet solvable target hitting tasks for the subjects. The non-intuitive nature of the task was assumed to encourage an augmented use of random exploration strategies. One target position was prescribed at 8.5 cm distance from the home location and 14° counter-clockwise orientation with respect to the cartesian plane defined by the VR screen (on which home location coincides with the x and y axis origin). Throughout the whole experiment a white sphere on the monitor indicated the home position of the cursor. Before the initiation of each trial subjects were asked to balance the cursor (black sphere, 0.4 cm radius) on the indicated home position for 0.8 sec. As soon as they balanced, a target (red sphere, 0.5 cm radius) was displayed at the prescribed target position and the trial was initiated. The subject had to move the cursor to the target and balance it there for 0.6 s (within a given accuracy range). The time-limit within which subjects had to hit the target was 10 s. This limit was defined based on prelimilary feasibility measurements on test subjects, which provided estimates for the average trial duration range. Participants received a positive or negative visual feedback (written message) depending on whether they respected this time-window or not. After hitting a target, subjects had to move the cursor back to the home position of the screen so as to satisfy the balancing requirement for the initiation of the following trial. One complete experiment included 75 succesful trials. The dynamics that determined the motion of the cursor on the virtual reality screen had the following form:

$$\dot{s} = A \cdot s + B \cdot u \tag{4.1}$$

where s corresponds to the states of the system (here to a vector including the 2D position of the cursor in x, y coordinates) and u represents the control signal that updates the state of the system. The mapping between balance and cursor motion was normalized based on the weight

of each subject individually. In Equation (4.1) the control variable is a four-valued vector that includes the four weight values tracked by the force plate (w_i)

$$u \rightarrow \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} \tag{4.2}$$

where w_1, w_2, w_3, w_4 correspond to the bottom left, top left, bottom right and top right force-plate sensor respectively. A and B represent the state and input matrices respectively which in the task design are arbitrarily defined as

$$A = 0.07 \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{4.3}$$

$$B = 0.35 \cdot \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \tag{4.4}$$

The system dynamics can thus be determined in a discrete space as

$$\frac{s_{n-1} - s_n}{dt} = A \cdot s_n + B \cdot u_n \tag{4.5}$$

$$\Leftrightarrow s_{n+1} = \frac{1}{F_{rec}} \cdot (A \cdot s_n + B \cdot u_n) + s_n \tag{4.6}$$

where $F_{rec} = 1/dt$ is provided by the sampling frequency in our experimental setup.

4.2.3 Formalising balance control and analyzing behaviour

Based on the control signal u_n and state values s_n, we examined the learning process that the subjects followed. We assumed a closed-loop control system, inspired by Todorov 2004 and Li 2006 and adjusted to the balance paradigm (Fig. 4.2). This generic architecture captures a continuous update of control policies and system state estimates produced by the brain in response to its experience in the actual task dynamics. In this sense, the architecture is conceptually consistent with our hypothesis in Chapter 3 on how task dynamics and control policies are represented and learned in the brain (see Fig. 3.2). In the present study the experienced task dynamics depend on the balance and posture biomechanical system, whose state is arbitrarily

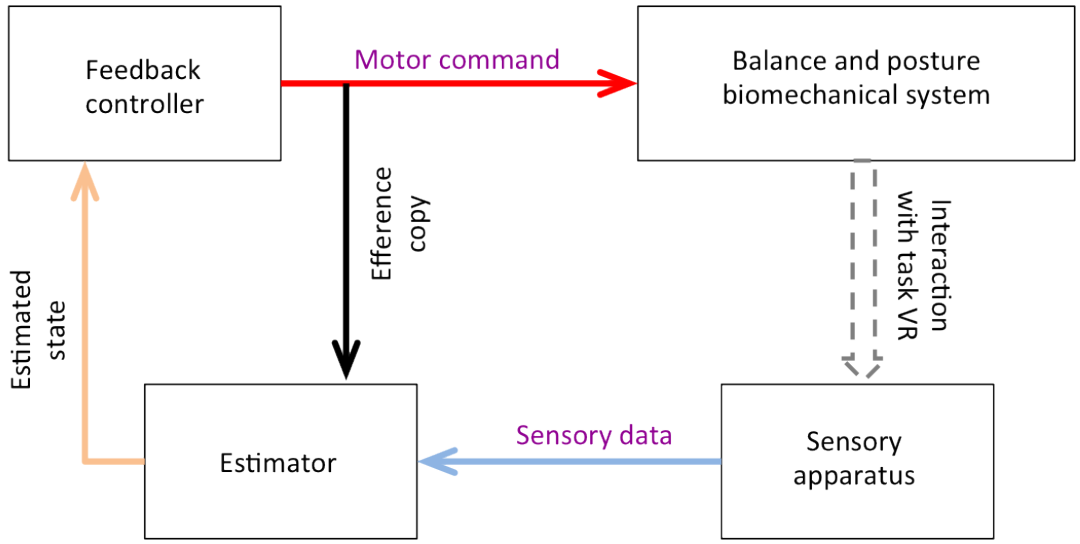

Figure 4.2: The model (inspired by Todorov 2004 and Li 2006 and adjusted to the balance paradigm) assumes closed-loop interaction between the brain and the world. The brain produces motor commands based on its state estimates of the task dynamics. It thereby interacts with the actual task dynamics that depend on the balance and posture biomechanical system. The visual feedback of the produced action is processed by the sensory apparatus and in turn drives further updates on the brain's system state estimates. The estimator component is consistent with the existence of an IFM (see Fig. 3.2), since it receives an efference copy of the motor command, in accordance to which it can predict a system state change.

translated to the motion of a cursor on the virtual reality screen.

Our aim was to determine whether subjects converge to a stable motor control strategy throughout their learning experience. The first part of the analysis focuses on specifying an expression of motor control which can successfully capture how subjects learn across trials. Assuming locally linear optimal feedback control, a feedback-gain function was used to link cursor motion and control values:

$$u = -L \cdot s \tag{4.7}$$

We used a linear approximation based on pseudoninverse matrices of the state values to estimate the feedback gain L. The latter was thus estimated as:

$$L_i = -u_i \cdot s_i^+ \tag{4.8}$$

where i denotes each recording and s_i^+ the pseudoinverse matrix of the state vector:

$$s^+ = \left(s^T \cdot s\right)^{-1} \cdot s^T \tag{4.9}$$

Based on the above, the estimated gain L_{2x4} (controller) for each recording constitutes an eight-valued matrix. We subsequently analysed the stability of the system dynamics in order to investigate whether the subjects preserve a stable motor control regime. To this end we examined the eigenvalues of the first order differential equation that determined the system dynamics in our experiment. Based on equation (4.7), equation (4.1) becomes:

$$\dot{s} = A \cdot s - B \cdot L \cdot s \Leftrightarrow \dot{s} = (A - B \cdot L) \cdot s \tag{4.10}$$

The solution of system equation (4.10) can be considered as asymptotically stable when $t \to \infty$ or in our discrete space as $i \to \infty$ ('in the future'), if and only if for all eigenvalues λ of $(A - B \cdot L)$:

$$\mathrm{Re}(\lambda) < 0 \tag{4.11}$$

In order to normalise the process of analysis, all cursor trajectories produced by a subject within one experimental session were rescaled to the maximum trial duration for that subject. The rescaling was performed based on cubic spline interpolation and applied to both the state estimations and the corresponding control estimations (acquired for each trial).

4.2.4 Control policy clustering

In a coplementary approach to examine the course of learning and motor control, we investigated whether subjects do not complete the experiment by progressively optimising a single controller, but instead switch to different controllers at different stages of the experiment. In order to address the question, we focused on specific sections of the subjects' trajectories and probed whether these sections reveal information about control policy exploration. In particular, we arbitrarily considered the first 700 recordings of each trial as the corresponding trajectory's initial phase and estimated for them a *common feedback gain* matrix, L_{common}, based on equation (4.8). This was done based on the assumption that this part of the trajectory is produced by a single controller. By generalisation, the whole trajectory emerges as a synthesis of controllers, each of which is used for different trajectory segments. The *common feedback gain* for the initial trajectory segment was estimated for all trials. We subsequently estimated the mean of these gain matrices across the last trials of the experiment, assuming that the latter represent the end-

point of learning within the given training time scale. We quantified the deviation of each trial's *common feedback gain* matrix, L_{common}, from this *mean final feedback gain*, L_{final}, using their root mean squared error ($RMSE$). A threshold $2 * th$ was set to distinguish which trajectories are products of controllers similar to L_{final}:

$$th = \frac{\sum_{k=1}^{k_{all}} RMSE(L_{common,k}, L_{final})}{k_{all}} \tag{4.12}$$

Trajectories k with $RMSE(L_{common,k}, L_{final})$ below $2*th$ can thus interpreted as motor outputs produced by controllers similar to the one responsible for the subject's end-performance (L_{final}). On the other hand, trajectories above the threshold can be regarded as motor outputs of different control policies. Based on the control similarity criterium

$$RMSE(L_{common,k}, L_{final}) < 2 * th \tag{4.13}$$

we thus designed a method of clustering motor control policies responsible for different types of motor outputs. This approach can be consequently employed to determine whether subjects explore new types of control policies during the motor learning process, instead of improving a single control policy until the end of their exposure to the experiment.

4.3 Results

We set out to examine the control patterns of the six subjects across all trials by first examining the cursor trajectories that they produced throughout the experiment (Fig. 4.3). These motor outputs reveal a noticable variability of adopted control policies across subjects. In Figure 4.3 we can distinguish at least three main movement patterns which subjects reach in the final phase of the experiment: (i) straight, close-to-linear trajectories (subjects 1, 4 and 5), (ii) two-step trajectories (subject 3) and (iii) multi-step trajectories (subjects 2 and 6). Furthermore, one can also distinguish variable movement patterns across trials of the same subject: straight and two-step trajectories (subject 3). We can therefore assume that the nature of the experimental task leads to the use of different motor control strategies across participants, based on the latter's individual previous experience and potentially encourages exploratory behaviour within subjects.

The selection of variable motor control policies is also apparent during the examination of the controller estimates across the timepoints of each trial. In particular, we estimated the feedback gain evolution for each trial based on equation (4.8), given the experimentally measured state s

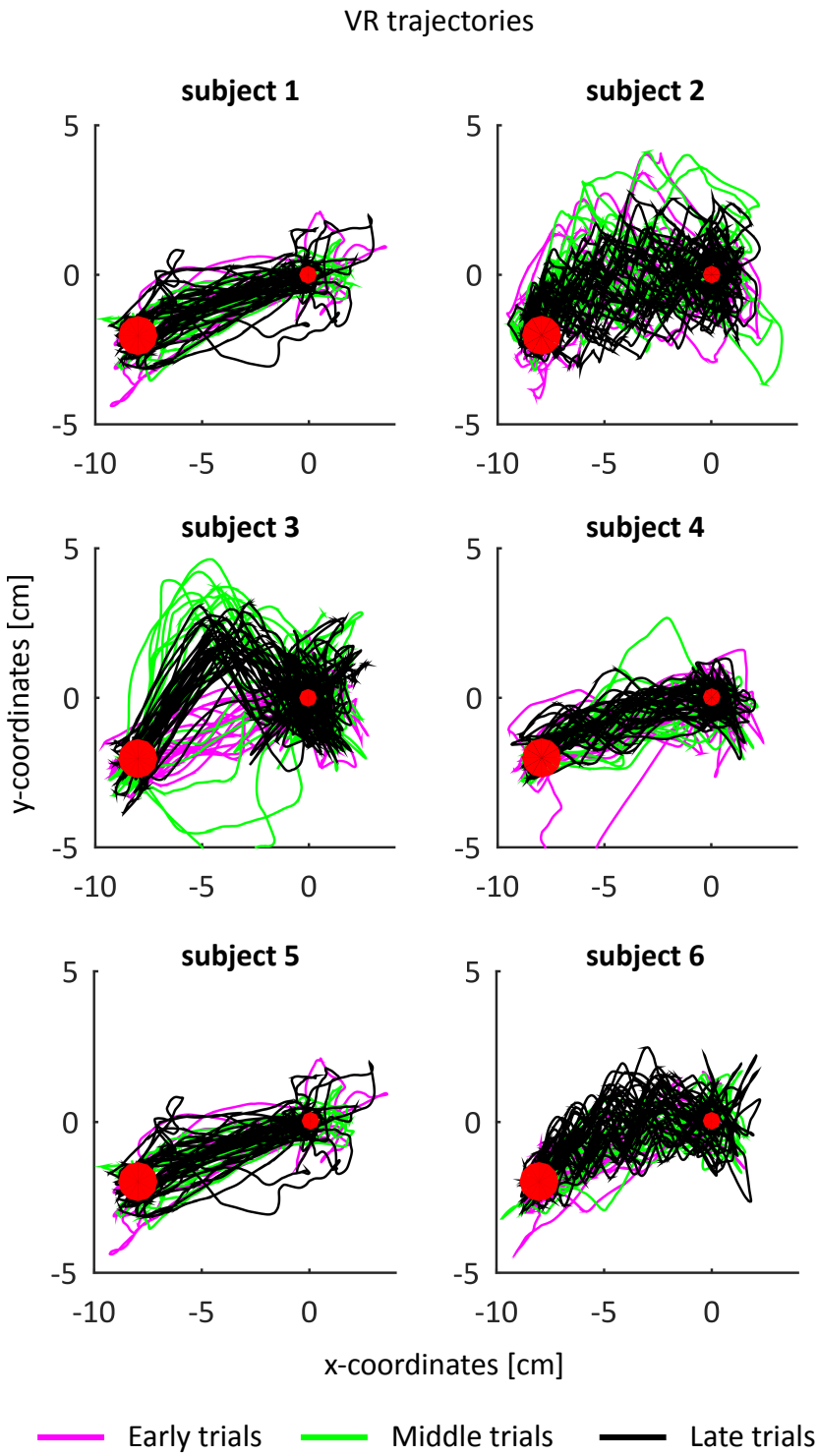

Figure 4.3: The illustrated trajectories represent the system state values, namely cursor motion on the 2D VR screen for all 75 trials per subject. Trajectories in magenta represent the early trials in an experiment (1-25), trajectories in green and black represent the middle (26-50) and late (51-75) trials respectively. The large and small red markers represent the target and home locations respectively.

and control u vectors. We realised that, as also illustrated by their actual behavioural outputs (Fig. 4.3), subjects reach different control patterns at the end of the experiment. Figure 4.5 illustrates the mean within-trial evolution of the eight-valued feedback gain matrix for subjects 1 and 3 across the last 25 trials of the experiment. We notice that subject 3 (two-step end-performance trajectory) employs a more complex motor control strategy than subject 1 (close-to-linear end-performance trajectory) to complete the task. This complexity is reflected in multiple oscillatory changes of the feedback gain matrices values, primarily in the latter stages of each trial, but also in earlier trial stages at a smaller scale.

The feedback gain's within-trial oscillatory trend -in either small or larger scale as seen in subjects 1 and 3 respectively in Figure 4.5- suggests a potential shift between controllers for the derivation of a task completion strategy. In other words, subjects might employ a combination of controllers for different segments of their motor output, in order to form a strategy that will achieve efficient trial completion. In order to further investigate this hypothesis we analysed the stability of the system dynamics during the final stage of each experimental session. The solution of this system (see equation (4.10)) can be considered as stable if and only if for all its eigenvalues λ, $\mathrm{Re}(\lambda) < 0$. We estimated the mean of the real eigenvalue component across the 25 last experimental trials (for both system eigenvalues). Figure 4.6 illustrates these

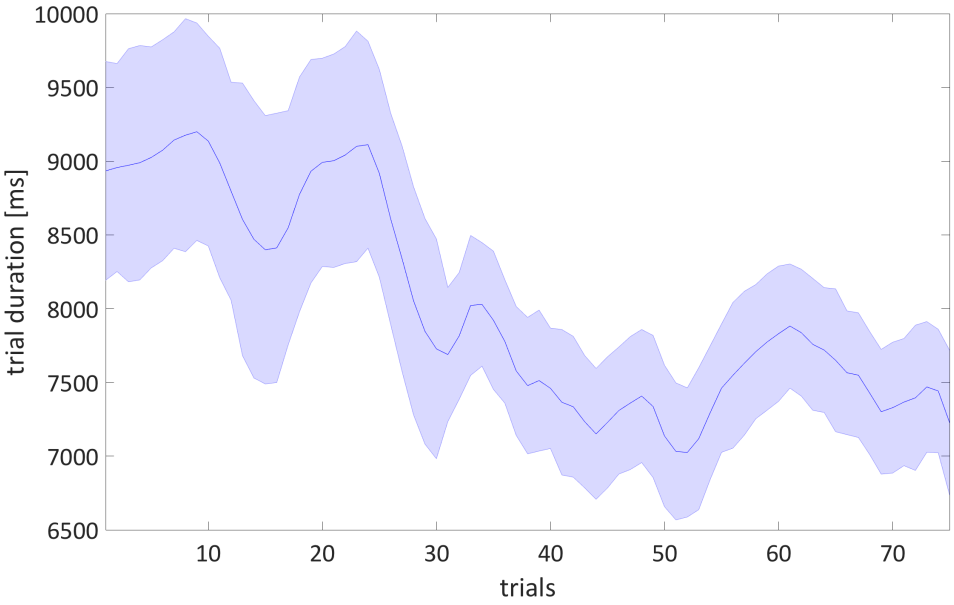

Figure 4.4: Mean group performance optimisation evaluated based on the progression of trial duration throughout the experiment.

mean estimates for the end-performance of all subjects. We notice that for the majority of the subjects, system dynamics are stable for the larger part of the trial (negative $Re(\lambda)$) apart from its last stage, during which there is an apparent instability (positive $Re(\lambda)$). The latter could be justified by the corrective movements subjects perform when they miss a target and try to re-approach it. These movements seem to require the use of new controllers; potentially different to the ones used for the earlier stages of the trial; hense the late-emerging instability. Evidently, for subject 2 this instability exists throughout a larger part of the trial, suggesting an association to this participant's very distinct multiple-step motor output. To summarise, the observed instability throughout the whole study appears to be related to the amount and form of trajectory steps produced by subjects until goal achievement: (i) minimal instability at the end of single step outputs and increased instability at the end of double step outputs potentially because of corrective movements before hitting the target, (ii) recurrent instability throughout the trial for multiple step outputs which is significantly constrained for flatter steps (smoother trajectories, e.g. subject 6).

On the whole, this method constitutes a stability analysis approach to the examined control system. As such, it links system stability shifts to motor output changes and provides indications of transition between different types of controllers that serves the purpose of trial completion.

In addition to analysing performance stability at the end of the experiment, we attempted to ex-

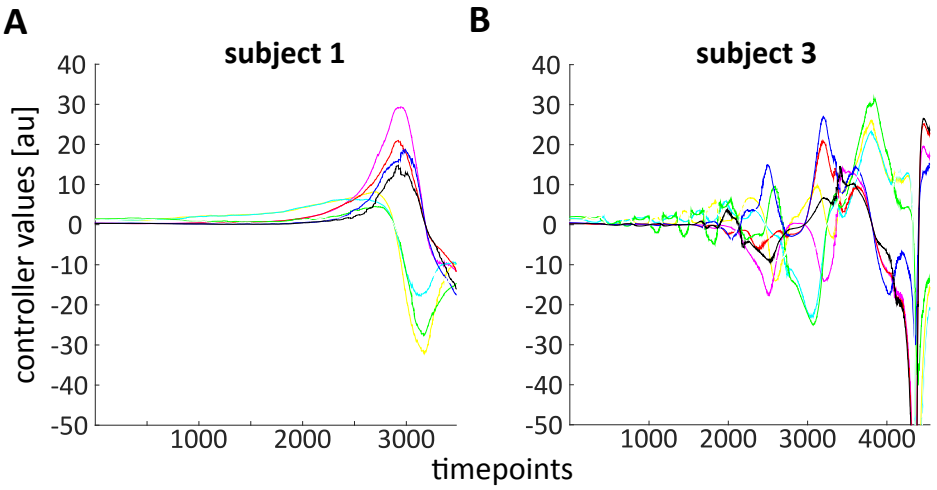

Figure 4.5: Mean within-trial evolution of the eight-valued feedback gain matrix for subjects 1 and 3 across the last 25 trials of the experiment. Each colour corresponds to a different component of the gain matrix (estimated based on equation (4.8)). Vivid oscillatory changes of the feedback gain estimates in subject 3 (**B**), primarily at the later stages of task completion, might be linked to greater motor output complexity (two-step versus one-step trajectory produced by subject 1 (**A**)), but also to corrective movements right before reaching the target.

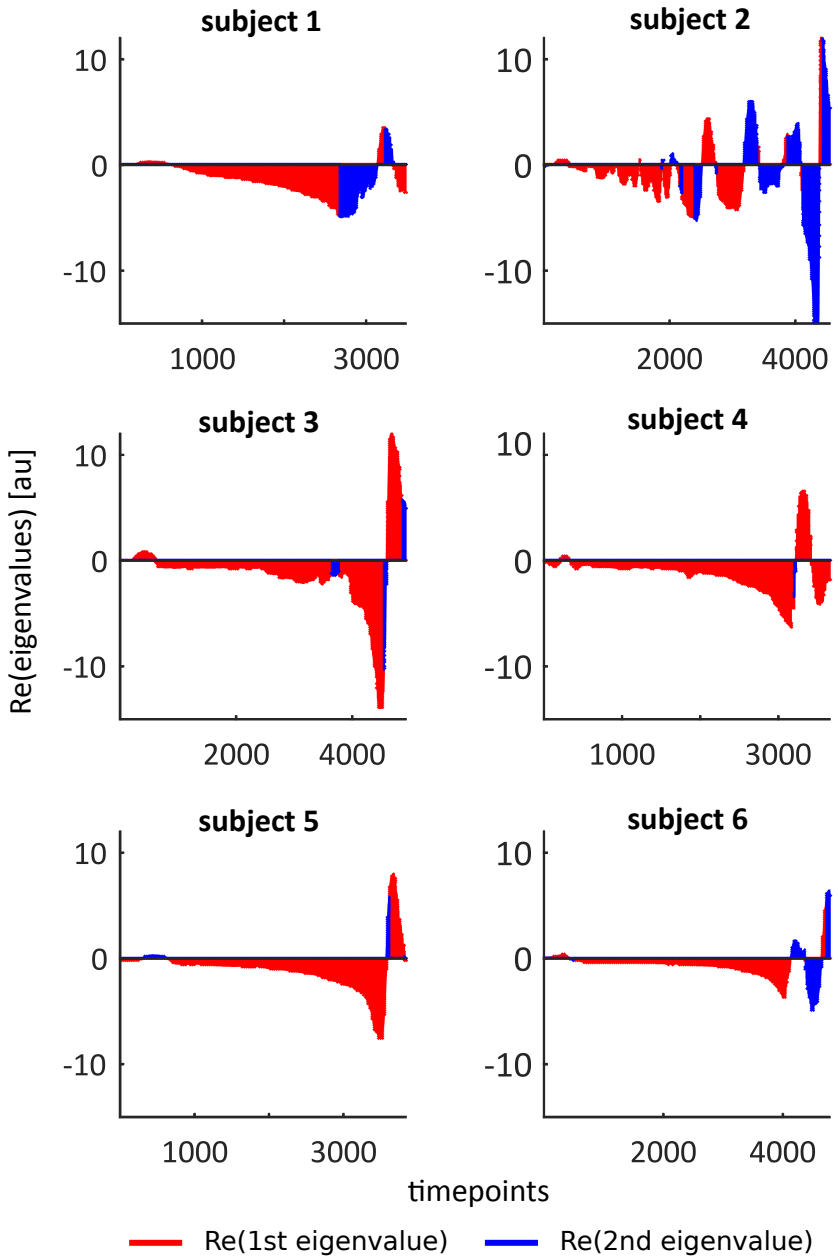

Figure 4.6: Histogram of the average real eigenvalue component across the 25 last experimental trials (for both system eigenvalues, see equation (4.10)). Essentially, the estimates illustrate the mean end-performance stability status of the experienced task dynamics, where negative $Re(\lambda)$ indicates stability and positive $Re(\lambda)$ indicates instability.

amine the course of learning and motor control throughout the experiment. We first established that subjects actually learn to optimise their performance in completing the task by progressively decreasing trial duration throughout the experiment. This decrease is evident in Figure 4.4 although its rate varies from subject to subject depending on employed motor strategies. We subsequently investigated whether subjects perform transitions to different controllers not only within a trial but also across trials. To that end we implemented our method in Section 4.2.4. Based on equation (4.8), we determined a *common feedback gain*, L_{common} for each trial's initial phase (captured arbitrarily by its first 700 recordings). We subsequently estimated the mean of these gain matrices for the final stage of the experimental session, L_{final}. The final stage was defined by trials 40-65. Trials 66-75 were not used for the estimation since signs of fatigue were evident during this interval. The deviation between L_{common} and L_{final} is illustrated in Figure 4.7 C,D across all trials for two different subjects (3 and 5 respectively). The corresponding L_{common} 8-valued matrix estimates for the two cases are also displayed in Figure 4.7 B,E.

Based on our similarity criterium in Equation (4.13), trials k with $RMSE(L_{common,k}, L_{final})$ below $2 * th$ (red, dashed threshold in Figure 4.7 C,F) were interpreted as motor outputs produced by controllers similar to the one responsible for the subject's end-performance (L_{final}). We thereby clustered motor control policies depending on their proximity to the end-performance control regime. Crucially, our clustering results group together motor outputs of similar appearance particularly in the case of subjects who appear to have produced variable behavioural outputs within their experimental session (e.g. one and two-step trajectory of subject 3 in Figure 4.7 A). At the same time our method associates these different motor outputs to the concept of different control policies. In the case of subjects, who appear to consistently produce only one type of movement pattern (e.g. close-to-linear trajectories of subject 5 in Figure 4.7 B), our method produces a much smaller cluster for the control policies that deviate from the end-performance control regime. This is visualised by the small number of red trajectories in Figure 4.7 B.

Consequently, our approach provides a tool that clusters motor control policies responsible for different types of motor outputs. It can be therefore employed to determine whether subjects explore new types of control policies during the motor learning process, instead of progressively optimising a single control policy until the end of the experiment.

4.4 Discussion

We showed that in a complex whole-body manipulation paradigm the brain allows switches between distinct motor outputs, which satisfy task completion. This suggests that the brain does not exclusively employ continuous exploitation of previously tested control policies, but

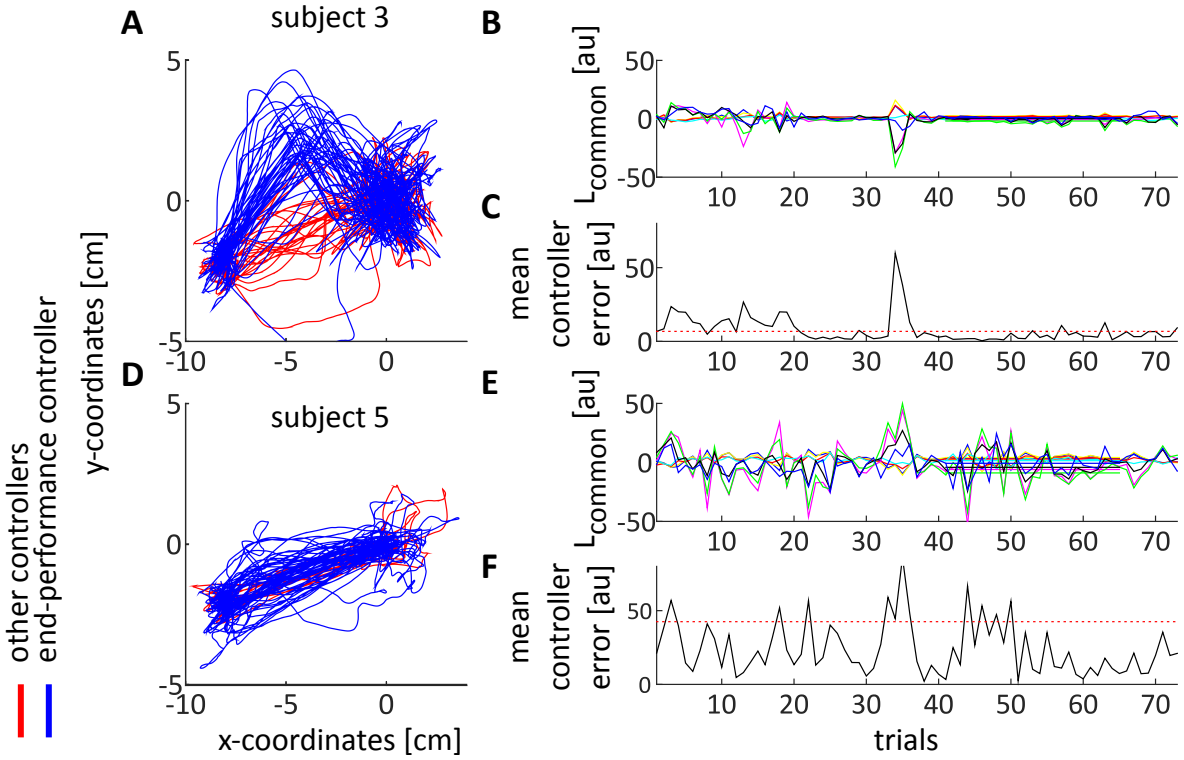

Figure 4.7: Control analysis for two representative participants is illustrated (subject 3 in upper half and subject 5 in lower half of the figure panel. (**B**),(**E**) The eight-valued *common feedback gain* for the initial phase of each trial, L_{common}, is plotted throughout the experimental session (black line). (**C**),(**F**) Based on its deviation from the *mean final feedback gain*, $RMSE(L_{common,k}, L_{final})$, we define a similarity threshold (red dashed line, see Equation (4.13)) to distinguish the *common feedback gains* which are proximal to the end-performance control regime. (**A**),(**D**) The emerging clustering provides a reasonable distinction between different types of motor patterns within the same experimental session. This distinction is particularly evident in the case of subjects who have clearly transited between variable motor outputs throughout the experiment (e.g. subject 3).

may also enable the exploration of novel control policies. The pattern of policy switches along with the selected policies themselves appear to vary amongst human subjects (e.g. single, double and multiple step trajectories), thereby highlighting the importance of previous experience in the evolution of motor learning. Our study also revealed an unstable regime of motor control within the completion of single trials at the end-point of learning. The degree and occurrence of this instability appear to be related to the amount and form of the trajectory steps produced until goal achievement: (i) minimal instability at the end of single step outputs and increased instability at the end of double step outputs arguably due to corrective movements before hitting the target, (ii) recurrent instability throughout the trial for multiple step outputs which is significantly constrained for flatter steps (smoother trajectories).

Our findings are based on a generic computational analysis approach to reverse engineer the control process in a balancing paradigm. The approach was used to capture the system dynamics in the examined task context and to estimate the control commands produced in response to sensory feedback from the world and in accordance to internal system state estimates. Furthermore, we proposed and tested a methodological platform that probes the stability of the motor control system within and across trials. This platform can also be used to investigate whether in our given task context the brain switches to new control policies while learning, or instead persists on the exploitation and improvement of a single motor control strategy.

Our approach thus contributes to the question of exploratory behaviour during the learning process, which is still an open challenge for neuroscience. This question has been addressed extensively in cognitive science and computer science, namely how adaptive agents manage a trade-off between exploration and exploitation (Sutton and Barto 1998). Such approaches, however, have not been thoroughly tested on motor behaviour, where the emergence of exploration remains unclear. Notably, motor neuroscience -despite its substantial work on describing and predicting motor control processes- has not validated a generic framework that captures the mechanism underlying the brain's exploration of new control policies during learning.

Along the direction of this challenge, our work here contributes a method that attempts to detect policy switches during a formalised motor control process. We used a balance task based on the criterium that balance and posture control serve as attractive paradigms for modelling specific problems in motor control (Kuo 1995; Morasso et al. 1999). Furthermore, interest in the motor control mechanisms of balance and posture evades the purely intellectual domain and finds great relevance to a clinical and daily life context. Balancing and postural stability skills constitute an elementary requirement of daily motor activities and are amongst the first skills to degenerate as a consequence of aging, leading to incidents of falls and bone ruptures (Coogler 1992; Gillespie et al. 2003). A comprehensive understanding of the general balance control mechanisms can thus contribute to an informed clinical treatment of aging population. According to the specialised

motivation of the present project, this understanding can be further enhanced by the study of the motor learning process and its conceivable 'plasticity', as the latter is expressed through exploration.

Moreover, setups such as the one developed for the present experiment can be used as a tool for balance assessment and neurohabilitation (Mombarg et al. 2013; Goble et al. 2014; Kennedy et al. 2011; Ding et al. 2013; Horak 2006). Based on this potential use, our balance control setup was selected and featured by the Medical Research Council, UK as a model neurotechnological platform in the 'Lab of Today' that combines gaming technology and tools to monitor healthy and pathological human motor behaviour. We specifically coordinated a broad public testing (over 2000 visiting users) of a balance control booth, which implemented a customised version of our system's task settings. The latter simulated balancing performance of participants in an assumed younger and older version of their present state in an attempt to shed light on the relation between aging and increased motor control and sensory processing variability ("Strictly science", science exhibition by Medical Research Council, UK, April 2013).

Our work here enhances our view of motor learning as the gradual improvement of one selected strategy, as we modelled it and investigated it in Chapter 3. It namely expands it to the context of switching to new motor control policies while learning, which in turn lead to evidently different behavioural outputs, that still achieve task completion. A reasonable outlook of this study would thus involve the formalisation and empirical validation of the exploratory aspect of learning, as an enrichment of our policy learning (PLM) framework. This direction can be arguably perceived as a link to a cognitive level of control, which is responsible for a high-level selection and exploration of motor strategies, that are dynamically configured at the execution level so as to optimise performance within a given task context.

Such a unifying approach can be informed by recent theoretical work on hierarchical learning algorithms for controlling non-linear dynamical systems with continuous states and actions (Abramova et al. 2011). Abramova et al propose a model that makes use of a high-level reinforcement learner, which searches and forms an optimal combination of low-level locally linear controllers and their associated quadratic costs (as actions and rewards respectively) in order to maximize performance. This algorithm can therefore be regarded as a candidate theoretical account of how the brain bridges symbolic action selection and low-level actuation control. As such, it could complement our low-level motor learning framework (see Chapter 3) so as to capture the findings of the present chapter on control policy exploration at the level of cognitive control.

4.4.1 Reviewing findings and emerging goals

With the present study we completed a line of research, which examines the process of motor learning in complex motor tasks. We first investigated how the brain represents a task with regard to object and body based coordinate systems. We subsequently integrated our findings on composite action representations into the development of PLM, an empirically validated motor learning model, which fuses control policy learning with task parameter adaptation. Lastly, we designed and tested an empirical study on whole-body manipulations to set the foundation in expanding PLM's perspective of motor learning at the level of continuous policy exploitation to a higher level of policy exploration. Admittedly, the question that emerged repeatedly throughout the completion of the afore-mentioned steps was: What are the neural constraints to our computational view and behavioural findings on motor learning? The implementation level appears to be a necessary context of testing so as to further validate but also inform/update our formalisation of sensorimotor functions. Thus, in order to address this challenge we proceeded to merging motor psychophysics with neuroimaging in an attempt to extract neurobehavioural correlates of the motor learning process.

5 f2MOVE: An fMRI-compatible object manipulation system for closed-loop motor control studies

5.1 Introduction

Our research aims at elucidating mechanisms of sensorimotor control and learning. We mainly rely on high-resolution behavioural data and a number of computational approaches that describe abstract mechanisms of the interplay between perception and action, as well as mechanisms of generalisation of learned activity (e.g. internal forward models, optimal feedback control). Such approaches -when empirically validated- provide a solid algorithmic formalisation that assists a systemic understanding of sensorimotor processes. Yet, the comprehesion of brain functions at the computational level often calls for projections to the implementation level (see Marr's three levels of analysis, Marr 1982); especially when it comes to drawing clinically relevant conclusions. This observation appears to relate to a greater central challenge in neuroscience.

Indeed, neuroscience has systematically studied sensorimotor functions for more than 100 years. Common methodological approaches have involved the use of psychophysical experiments and computational theories of inference and policy formation (Wolpert et al. 2011; Todorov 2004),(Todorov 2004; Scott 2004; Friston et al. 2009; Friston 2010; Körding and Wolpert 2006). The latter provide insight into the patterns of adaptive responses in tasks that introduce target, workspace or force-field perturbations (Shadmehr and Mussa-Ivaldi 1994; Shadmehr and Moussavi 2000) and examine multisensory integration (Vaziri et al. 2006), body and world representation (Sylaidi and Faisal 2012) or performance optimization in the face of sensory and motor noise (Faisal et al. 2008).

However, despite this substantial progress in the investigation of motor behaviour, less advancement has been achieved in associating motor psychophysics and computational models of sensorimotor control to their underlying neural foundation. A growing number of studies have attempted to address this challenge with the use of fMRI technology, which they primarily employed during lab-constrained simple hand reaching movements or non-specific open-loop

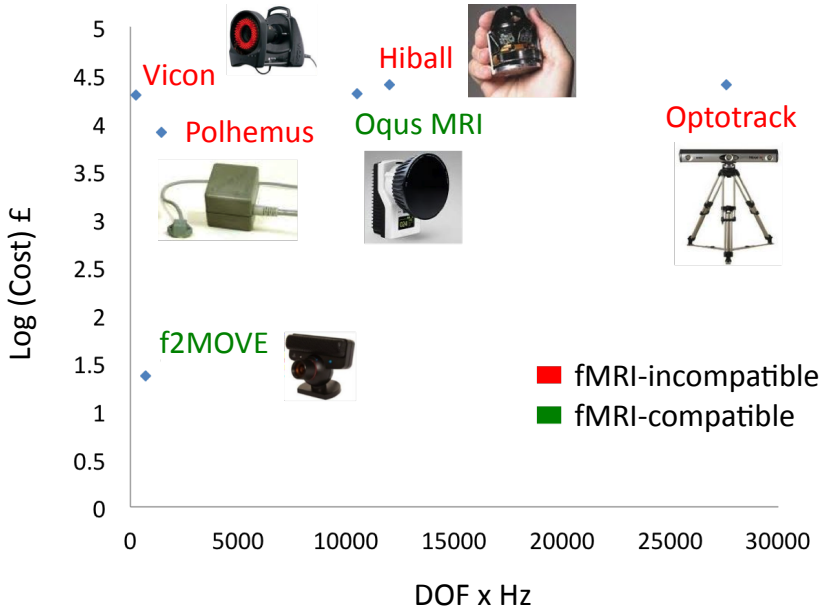

Figure 5.1: Frugal innovation for motion tracking: Cost versus efficiency of established motion tracking systems (not fMRI compatible) against which f2MOVE (fMRI compatible) is compared. f2MOVE corresponds to the lowest cost level (camera expenses) and possesses data with satisfactory motion tracking performance within the range of state-of-the-art motion trackers (e.g. Polhemus, Vicon).

manipulations (e.g. finger tapping). The main reason for this restriction lies in the technical constraints of fMRI, which is often incompatible to advanced motion tracking systems, that could monitor more complex motor behaviour.

Here we designed and developed f2MOVE, a novel 6DOF fMRI-compatible motion tracking system to support realistic object manipulation (haptic) tasks during a neuroimaging session. The development was motivated by a rapidly growing body of studies that focuses on life-like tasks, which enable movements in naturalistic settings without the usual confines of strict lab protocols. This work supports a better understanding of human natural movement statistics (Ingram et al. 2008),(Howard et al. 2009), provides insight into the structure of motor primitives and thus carries, through movement predictability, direct implications for neuroprosthetic approaches and brain-machine interfaces (Thomik et al. 2013).

f2MOVE was built upon our low-cost motion capture technology, fMOVE (Rodriguez et al. 2014), which we expanded to adjust to the fMRI environment (Clinical Imaging Facility, Hammersmith Hospital, London). We complemented our system with a methodological platform, which can support closed-loop task contexts to encourage learning based on online sensory

feedback of performance. f2MOVE's motion tracking performance lies in the same range as established fMRI-incompatible motion tracking methods (e.g. in Figure 5.1 Vicon with 1DOF motion tracking at 250 Hz, Polhemus Liberty with 6DOF at 240 Hz). Furthermore, it presents a significant cost-efficiency benefit compared to other motion-tracking technologies with advanced features (e.g. Hiball with 6DOF motion tracking at 2000 Hz, Optotrack with 6DOF at 4600 Hz, Oqus MRI with 6DOF at 1750 Hz).

5.1.1 Core contributions

The core contribution of the present project lies in developing a low-cost, fMRI-compatible haptic, object manipulation interface which is integrated inside a closed-loop motor control experimental platform. The system poses itself as a novel strong tool towards bridging high resolution naturalistic motor behavior to its neural foundation level.

5.2 Aims and methods

Our main motivation was to design and build a system that allows real-time association of motor psychophysics to the neurophysiological level of behaviour. We particularly aimed at developing a platform that can support a closed-loop interaction of subjects with physical objects of adjustable shapes and dynamics, in an attempt to encompass naturalistic tasks in the context of our studies.

5.2.1 Hardware

Our system allows a closed-loop interaction of human subjects lying inside a 3 T fMRI scanner and a PC-handled experimental paradigm (Fig. 5.2 A,C). Subjects use their dominant hand to hold and manipulate a compact object so as to move it correctly between some instructed home and target orientation. The exact task conditions are displayed to them via a mirror system built inside the scanner, which reflects the virtual progression of the task on a computer screen (e.g. 'rotate object from home to target' in Figure 5.2 B). Motor performance is tracked by a high-zoom lens, low cost camera (Playstation 3 Eye) positioned on one of the scanner room walls, at a 3.5 m outside of the scanner, facing the foot-end view of the cylinder. The camera operates at a 120 Hz frame-rate for a 320 x 240 pixel resolution with low time delays ($<$20 ms) and can track 6DOF movement of the object, based on a customized 2D marker adjusted on the latter. We mounted it on a customized platform that allows a 6DOF rotation and precice alignment of the camera orientation with the scanner opening. The motion tracking system feeds all acquired images into a PC-based software module, which monitors motor behaviour and determines the transition between consecutive experimental phases. This module communicates

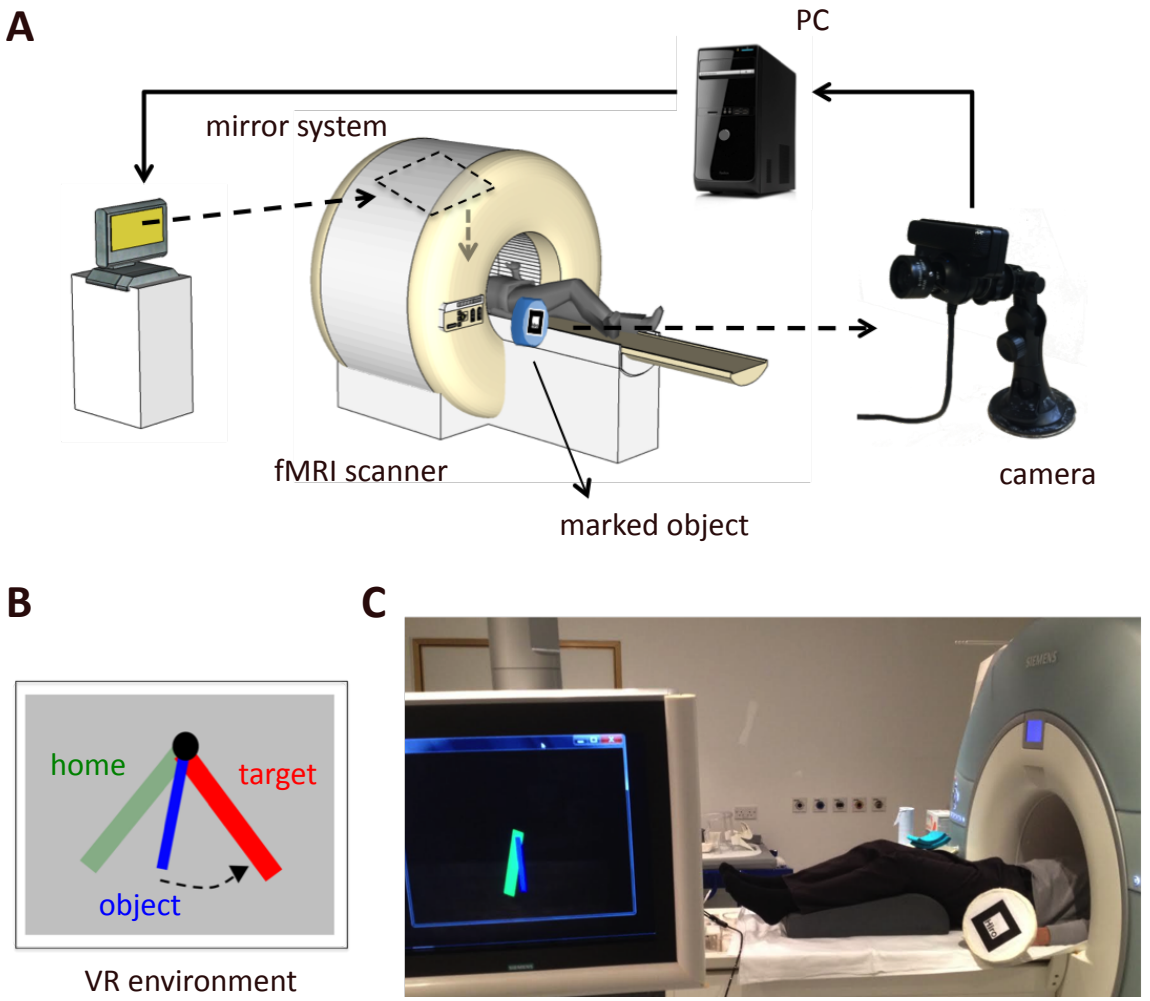

Figure 5.2: Closed-loop object manipulation platform (**A**) The platform consists of several components which interact so as to support closed-loop haptic object manipulation experiments. The object position is tracked by our customized marker-based motion tracker, which provides a PC-run software module with continuous information about the subject's movements. (**B**) The software platform displays a virtual online feedback of the monitored behaviour and performance, which is in turn presented to the subject via a computer screen and a mirror system adjusted to the scanner. (**C**) f2MOVE was adjusted, calibrated and tested in the Clinical Imaging Facility, Hammersmith Hospital, London.

with the subject via the scanner-based mirror system, on which it provides visual access to the task goals, and real-time feedback of motor performance.

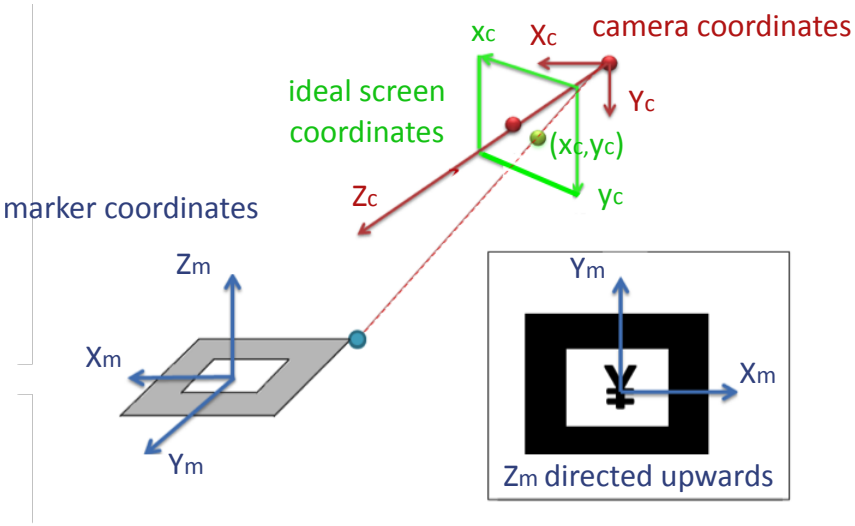

Figure 5.3: Coordinate systems used for motion tracking: The marked object position is estimated based on a transformation between a marker-based (X_m, Y_m, Z_m) and a camera-based (X_c, Y_c, Z_c) coordinate system (see Equation (5.1)). The ideal screen coordinates define a projection plane for the camera coordinates and are linked to a real screen coordinate plane via a distortion factor (Kato and Billinghurst 1999; Wagner and Schmalstieg 2007).

5.2.2 Continuous closed-loop object tracking

The camera continuously records images, which support the estimation of the object position and the visualization of its virtual analogue based on a software library for designing Augmented Reality applications (ARToolkit, Wagner and Schmalstieg 2007). In particular, f2MOVE estimates object position by localising the position of an orthogonal marker (Fig. 5.3) or of multiple markers, depending on the design of the object. The motion tracking system is based on two coordinate systems which are related via an affine transformation (Fig. 5.3); a camera-based (3D) one and a marker-based (3D) one:

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} = \begin{pmatrix} R_{11} & R_{11} & R_{11} & T_1 \\ R_{11} & R_{11} & R_{11} & T_2 \\ R_{11} & R_{11} & R_{11} & T_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{pmatrix} \tag{5.1}$$

where $R_{i,j}$ denotes rotation and T_i translation. The axes of the marker-based coordinate system, X_m and Y_m, are aligned with the horizontal and vertical marker sides respectively (Fig. 5.2 C). The marker centre is located at $(X_m, Y_m, Z_m) = (0, 0, 0)$. So, by convention, motion tracking

considers the marker location fixed and estimates the movement of the camera with regard to the marker centre. This displacement can subsequently be interpreted as the displacement of the marker/object in the physical reality. In Figure 5.3 x_c and y_c define a camera projection plane, called ideal screen coordinate plane, which in turn is related to the real screen coordinate plane via a distortion factor. The latter is estimated by calibrating the motion tracking system to markers located in the distance ranges of our clinical setting.

The software module of f2MOVE essentially includes the motion tracking unit which we interface with an algorithmic component that determines the experimental stages and the feedback to the subjects. Our implementation is based on a custom made motorlib C code library and opengl and operates by defining consecutive experimental phases as finite states which allow transition, depending on the monitored and processed motor performance.

5.2.3 Haptic interface

We developed a haptic interface that includes a light fMRI-compatible object. The design consists of a plastic handle, used as the reception for mounting multifaceted objects of variable shape and dynamics. For the needs of our first clinical sessions we designed and used an object with thin cylindrical body that hosted a single marker tractable by our motion tracking system (Fig. 5.2 A,C). Depending on the needs of the experimental study and the complexity of the examined movement, the number of markers on the tracked object can be increased to ensure that occasionally obscured markers, are substituted by more visible ones. This offers flexibility to the experimental design and allows us to investigate motor behaviour during haptic interaction with complex objects.

5.2.4 Tracking accuracy

We tested the tracking accuracy of f2MOVE against a standardized motion tracking system (Optitrack Flex-13). In particular, our reference is a state-of-the-art infrared marker-tracking system that offers millimiter resolution of 3D spatial displacements and operates with accuracy at 100 Hz. In order to compare the signal acquired by the two systems, we aligned the f2MOVE camera with one of the three used Optitrack cameras. We placed four Optitrack markers at the corners of the f2MOVE 2D marker, so as to create a rigid body with the same center to f2MOVE's marker center. We acquired motion information from a healthy female subject, based on an experimental paradigm designed for the Clinical Imaging Facility at Hammersmith Hospital, London. In this paradigm the subject continuously moves the marked object between an instructed home and target orientation.

After data acquisition, f2MOVE signal was downsampled to match the reference signal. The

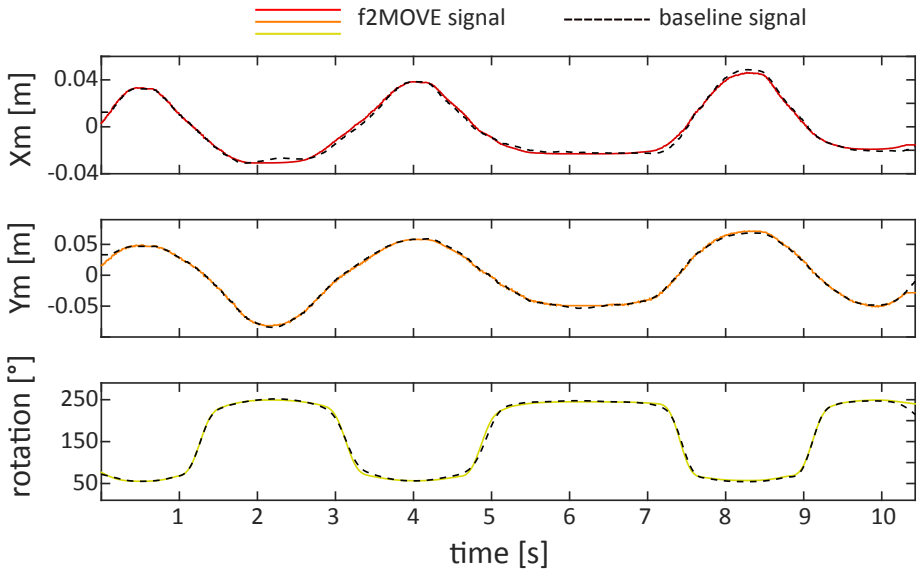

Figure 5.4: Testing the tracking accuracy: We compared the f2MOVE signal against the signal monitored by Optitrack Flex-13, which was used as reference. Both signals were acquired simultaneously in a lab setting simulating the experimental conditions of an object manipulation paradigm designed for clinical studies inside the fMRI environment. The f2MOVE signal (translation in x, y and rotation of the xy plane) demonstrates a satisfactory matching to the reference motion tracking performance (z-signal not included here since displacement in the z axis is minimal). However, motion tracking can also cover larger displacements in the z axis, based on the measured marker dimensions, which increase or decrease depending on the distance to the camera.

two signals were subsequently aligned temporally by matching the first trial initiation after a resting period. A mean affine transformation was estimated for resting period data, to match the f2MOVE signal to the reference signal, via rotation and translation. The transformed signals were compared based on $RMSE$ and R^2 to determine f2MOVE tracking accuracy.

5.3 Results

We tested the operating features and tracking accuracy of our system, f2MOVE, in lab conditions and inside the fMRI environment to establish its functionality for closed-loop object manipulation tasks. f2MOVE operates successfully at 120 Hz frame-rate with low time-delays and tracks 6DOF movement of the marked object.

f2MOVE's precise tracking accuracy was estimated against an established infrared marker-tracking system (Optitrack Flex-13) in a lab setting simulating the current experimental setting of our system at the Clinical Imaging Facility at Hammersmith Hospital, London. We measured motion information for a task instructing the manipulation of the experimental object inside a specified orientation range. We aligned the two acquired signals in the temporal and spatial

domain to achieve matching, after which we estimated their RMSE difference. Figure 5.4 reveals good consistency of movement measurements between f2MOVE and our reference system, both in the dimensions used for our experimental paradigm (execution was completed primarily on the xy plane and movement in the z dimension was minimal) and in the angle rotation domain of the xy plane. The matching between signals was estimated at $R^2 = 0.98$ and $RMSE$=2 mm, $R^2 = 0.99$ and $RMSE$=2.6 mm, $R^2 = 0.99$ and $RMSE$=1.02 mm, $R^2 = 0.99$ and $RMSE = 4.2°$ for the x,y,z translation and rotation of the xy plane around the z-axis respectively.

Our tested system was transferred and adjusted to the fMRI setting at the Clinical Imaging Facility at Hammersmith Hospital, London. It was used to support scanning sessions on healthy subjects during an object manipulation paradigm. Figure 5.5 A illustrates 3D rendered views of activation patterns for a task-versus-rest analysis design in our paradigm. These views were created based on the available software tools within FMRIB's Software Library (FSL) and can be used to investigate the enhanced role of specific brain areas during error-driven closed-loop object manipulations (e.g. cerebellum, SMA). Furthermore, the simultaneously acquired behavioural data (Fig. 5.5 B,C) can be regressed against activation patterns to examine neural correlates of behaviourally determined motor control mechanisms (see Chapter 6 and Chapter 7). Such approaches can shed further light into the foundation of sensorimotor functions in naturalistic motor tasks, such as the one examined by our system.

5.4 Discussion

We designed and developed an fMRI-compatible haptic object manipulation system (f2MOVE) for closed-loop 6 DOF motor control studies. We built upon our previously developed 3 DOF marker-based motion tracking system (Rodriguez et al. 2014) to adjust our technology to the fMRI environment and to expand it so as to accommodate motor experiments with goal-directed hand and wrist movements as well as the interaction with objects of variable dynamics.

f2MOVE poses technical benefits for high frequency data acquisition inside the fMRI environment, which is commonly incompatible to most currently established motion tracking systems. Moreover, building on our previous experience on low-cost wearable kinematic body-sensor networks (Gavriel and Faisal 2013), we aimed at developing an easily affordable neurotechnological tool. Its cost is practically limited to the customized camera price, unlike recent approaches to fMRI-compatible motor control platforms, which rely on robotic manipulanda for motion detection and/or perturbation (Diedrichsen et al. 2005; Menon et al. 2014). Furthermore, in contrast to these approaches, f2MOVE enables interaction with physical objects with familiar or unfamiliar dynamics, thereby supporting a naturalistic perspective to sensorimotor control, which provides an important adjunct to traditional laboratory studies (Ingram and Wolpert

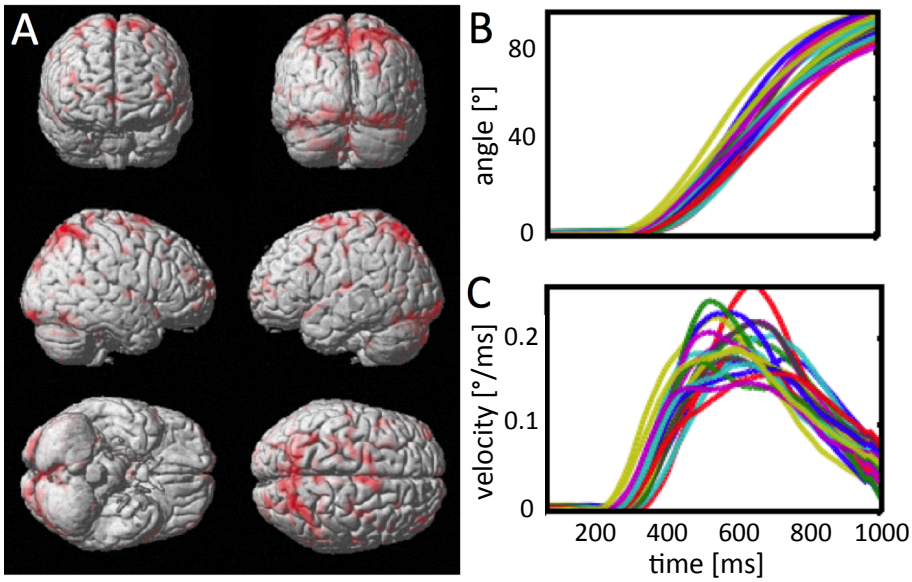

Figure 5.5: Linking neuroimaging and motor psychophysics: f2MOVE enabled fMRI sessions on a healthy subject during a closed-loop object manipulation task, equivalent to the one tested in our simulated lab conditions (moving object between home and target orientation). (**A**) Our methodological platform allowed the acquisition of cortical activation patterns in a task-versus-rest analysis design. An examination of these patterns reveals the role of the cerebellum and SMA in error-driven motor learning. Network-level analysis can support a further understanding of the neural implementation of motor behaviour in complex naturalistic manipulations. (**B**) The object orientation and (**C**) angular velocity over time (trial start at 0 ms) across trials can be examined in parallel with the cortical activation patterns, so as to determine behavioural performance measures that can be subsequently employed as regressors against the fMRI signal.

2011).

In summary, f2MOVE, offers an accessible technological and methodological platform to re-approach the objectives of motor neuroscience in examining the neural foundation of sensori-motor control and learning in naturalistic task contexts (see Chapter 6). It also supports the design of clinically-valuable behavioural and neuroimaging markers (see Chapter 7) to monitor motor coordination in healthy or pathological cases (e.g. neurodegenerative diseases).

6 Neuroimaging correlates of closed-loop motor learning during object manipulation

6.1 Introduction

Motor neuroscience has systematically examined motor behaviour via experimental and computational approaches (Wolpert et al. 2011). This research has provided valuable empirical and theoretical insight into the mechanisms of sensorimotor control and learning. However, despite hundreds of published studies, a comprehensive understanding of the neural processes that underlie these mechanisms remains an open challenge.

Past neurophysiological studies have suggested the role of different components of the CNS in planning and coordinating movements. Amongst these components, the cerebellum constitutes an area that has been highlighted as a strong candidate for the foundation of motor learning (Houk et al. 1996). From the early lesion studies and mechanistic models of the nineteenth century (Flourens 1842) to the latter electrophysiological and computational approaches to cerebellar processes of synaptic plasticity, credit assignment, and the generation of training information (Houk et al. 1996; Albus 1971; Marr 1969) a substantial body of work has been dedicated to this theory. Importantly, the cerebellum has also been proposed as the locus of internal forward and inverse models of the motor apparatus (Wolpert et al. 1998; Kawato et al. 2003; Albus 1971; Marr 1969; Kawato et al. 1987; Jordan and Rumelhart 1992) which predict consequences of actions (Wolpert et al. 2011) based on cortical representations of the world. Nevertheless, the details of implementing such models in neural tissue remain obscure.

6.1.1 Human neuroimaging in motor neuroscience: the puzzle of bridging the laboratory and naturalistic context divide

Human neuroimaging, and fMRI in particular, is currently featured as a key tool in the attempts to shed light on the neural foundation of motor behaviour. Recently, it has contributed significant progress in uncovering the functional neuroanatomy of learning. fMRI has been used to investigate processes of visuomotor learning, motor sequence learning and conditional motor

learning (Graydon et al. 2005; Simmonds et al. 2014). Numerous such approaches examine the impact of different aspects of learning on shaping/strengthening distinct cortical representations (Wiestler and Diedrichsen 2013; Graydon et al. 2005). Yet, while comparing brain activations before and after training sessions is useful for determining the overall effect of learning on neural processes, more attention is gradually being drawn to a better comprehension of the learning-related evolution of activations. This perspective targets the actual implementation of the learning mechanism. To date, work in this direction primarily aims at positing the role of different regions throughout the stages of motor learning (Wiestler and Diedrichsen 2013; Jenkins et al. 1994; Sakai et al. 1998; Grafton et al. 1994; Deiber et al. 1997; Sakai et al. 1998). However, a systemic interpretation of these findings requires a careful regression against behaviour, which represents the learning output. This regression is arguably the path to achieving an integration of neurophysiological observations in behaviourally validated computational models of motor learning.

Notably, acquiring multiple behavioural measures of motor learning during neuroimaging sessions is subject to technical constraints embedded in fMRI use, which is often incompatible to advanced motion tracking systems (see Chapter 5). Thus, instead of tracking complex movements, some studies use alternative physiological data as "behavioural metrics" to complement fMRI signal and analysis (e.g. electrodermal activity in MacIntosh et al. 2007). In recent years, new methods emerged in which fMRI-compatible robotic manipulanda allow online tracking of motor behaviour simultaneously to neuroimaging sessions (Menon et al. 2014; Diedrichsen et al. 2005). However, such methods examine predominantly simple hand reaching movements and have not yet covered a broader range of naturalistic task contexts within the fMRI environment.

En masse, the poorly understood mechanisms underlying naturalistic motor behaviour pose an enduring challenge in neuroscience: to solve the puzzle of motor learning by bridging the laboratory and real-world divide. This divide appears particularly large for the fMRI environment, which by design restricts the range of examined task contexts. One way to address this restriction would be to introduce life-like objects in fMRI task settings, since object manipulation emerges as a dominant component of naturalistic behaviour in daily life (Ingram and Wolpert 2011; Brown 1986; Parker and Gibson 1977; Piaget 1954). Yet, fMRI literature has not yet substantially covered object manipulation experiments. One study by Binkofski et al. 1999 examined the existence of a fronto-parietal circuit for learning object manipulation in the human brain, but relied on an open-loop non-specific interaction of the subjects with the provided physical objects. Namely, participants were not provided with real-time feedback of their movement since the task merely instructed them to name the objects upon recognition.

6.1.2 Core contributions

Driven by the afore-mentioned challenges, we used our fMRI-compatible object-manipulation system to achieve real-time association of motor psychophysics to neurophysiology (see Chapter 5). The system allows closed-loop interaction with life-like physical objects of potentially variable shapes and dynamics and provides online feedback of motor performance to subjects to facilitate naturalistic task conditions for the study of motor learning. In the present section, we employ this platform to specifically design and conduct an object manipulation experiment on healthy human subjects driven by the design of our empirical study in Chapter 2. Our methodological tools support the acquisition of behavioural outputs, whose evolution can be evaluated and regressed against their concurrent fMRI signal. They thereby present the opportunity of a fuller understanding of motor learning via both its neural correlates and performance measures inspired by our formerly validated computational approach (see Chapter 3). The novelty of our work thus lies in the development of technology and methodology that aims at the systematic extraction of neurobehavioral correlates of sensorimotor functions in naturalistic motor tasks. Such correlates can serve as metrics towards a fuller characterisation of the sensorimotor system, in an attempt to bridge the ever challenging divide between behavior, computation and implementation.

6.2 Aims and methods

Our aim in the present section was to expand our previous psychophysical studies on sensorimotor functions to the fMRI environment so as to encompass neurophysiological information. To this end we conducted a naturalistic object manipulation experiment on healthy human subjects that would enable a correlation of motor psychophysics to neuroimaging data. For the purposes of the study we employed the previously developed f2MOVE system; an fMRI-compatible haptic object manipulation system for closed-loop naturalistic motor control studies.

6.2.1 Experimental setup and data acquisition

Five healthy right-handed human subjects participated in the study (2 women; 3 men). The participants' mean age was 35.8 years. Based on self-reports the subjects had no known disabilities. They were naive to the purpose of the experiment and all provided informed consent consistent with the policies of the Imperial College London Ethics Committee.

The experiment was performed inside an fMRI environment, with subjects lying horizontally inside the scanner while performing an object manipulation task with their dominant hand. The control of the transition of subjects through experimental phases, the simultaneous activation

of fMRI scanning sessions and the monitoring of the subjects' movements was handled by our closed-loop object manipulation system, f2MOVE (see Chapter 5). The f2MOVE platform includes (i) a compact physical object, which subjects are instructed to manipulate based on the virtual task conditions displayed to them via (ii) a mirror system built inside the scanner, (iii) a 2D-marker-based motion tracking system that records 6DOF object movements with a high-zoom lens Playstation 3 Eye camera (120 Hz frame-rate for a 320 x 240 pixel resolution with low time delays (<20 ms) and (iv) a central processing unit with a software module, that monitors motor behaviour and determines experimental flow and online visual feedback to the subjects. For the needs of our first clinical sessions we designed and built a handle-held object with thin cylindrical body that hosted a single marker tractable by our motion tracking system (Fig. 5.2 A,C). Both the design and the configuration of the object with regard to the hand were partially inspired by naturalistic object manipulation conditions employed in previous studies (see Chapter 2).

Neuroimaging data were acquired with a 3 T magnetic resonance imaging (MRI) whole body scanner (Siemens 3 T Verio MRI Scanner, Imperial College London, Clinical Imaging Facility, Hammersmith Campus, London, United Kingdom), equipped with a radio frequency (RF), 32-channel head coil offering whole head coverage. Before the functional images were collected for each subject, a high resolution, T1-weighted, 3D anatomical volume was acquired with the same slice orientation and field of view as the functional images [MP-RAGE sequence; TI=900 ms; TR=2.3 s; TE=2.98 ms; flip angle=9; 160 slices; 1 mm thick; matrix size=240x256 pixels]. Subsequently, contiguous, multislice BOLD images [repetition time (TR)=3 s/vol; echo time (TE)=30 ms; field of view (FOV)=192 mm^2; matrix size=64x64 pixels] were collected with echo planar imaging, using an axial slice orientation. The volume acquired for each subject covered the whole brain and consisted of 52 slices, 3 mm thick (voxel size=3x3x3mm^3). Scan duration was variable, as a result of the experimental design described below. The experiment started with a 30-second resting period to allow the MRI signal to reach steady state. At the beginning of each subject's scanning session, an initial TTL signal from the scanner to the computer was used to trigger the start of the sequence. Foam padding was used to limit head motion within the RF coil.

6.2.2 Experimental protocol

Two experimental paradigms on object manipulation were used for this study; one provided online visual feedback of movement to the subjects during the completion of each trial and the other one didn't (Fig. 6.1 A). In particular, subjects were instructed to hold a provided handle-held thin cylindrical object in their dominant hand (Fig. 5.2 A,C and Section 6.2.1) and perform rotational movements between a home and target orientation. The instructed home and target

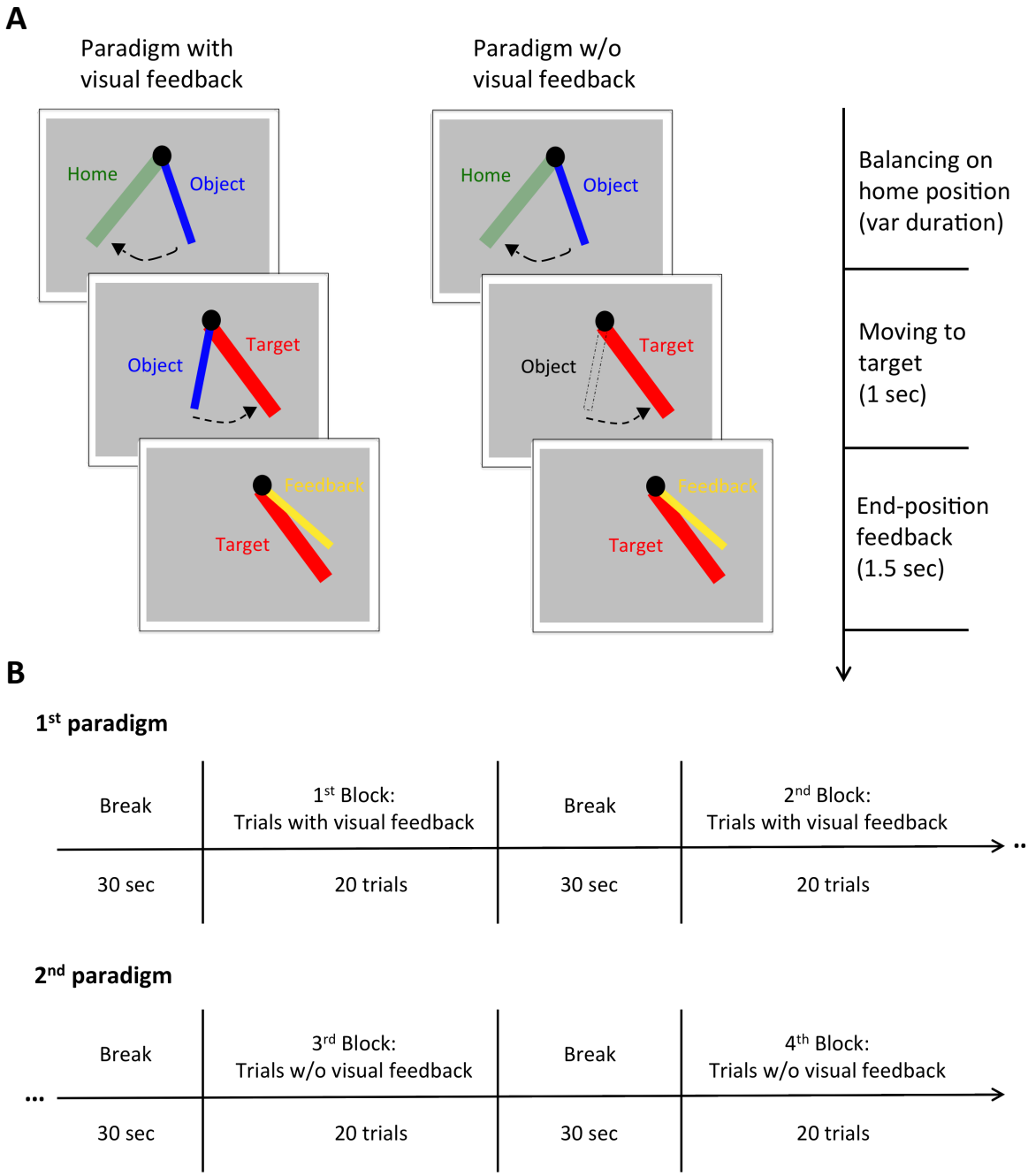

Figure 6.1: (**A**) Two object manipulation paradigms are used in the study; one that offers online visual feedback of movement during the execution of the trial (left panel) and one that doesn't (right panel). The first segment of each trial instructs the alignment of the object (blue bar) to a given home orientation (green bar) and has variable duration. The second segment instructs the alignment of the object to a given target orientation (red bar) as accurately as possible within 1 s. The third segment provides visual feedback of the alignment error (yellow bar) at the end of the target hitting segment and lasts 1.5 s. (**B**) Each paradigm contains 2 blocks of 20 trials each, which are preceded by 30 s resting periods.

orientations were illustrated as thin green and red bars respectively (Fig. 6.1 A) on the virtual reality screen displayed to the subjects via the mirror system built into the scanner. During the experiment subjects' rotating arm was extended horizontally along the body and stabilized with the help of foam padding to limit redundant motion.

The instructed movement was determined for each subject individually, by first estimating the maximum rotation range for the extended dominant arm, and then confining each trial within the +/- 25% of this maximum rotation range about the frontal part of the sagittal plane (-25% for the home orientation and +25% for the target orientation resulting in clockwise movement while reaching the target). The rotation centre was localised at the object's 2D marker centre, which in turn was controlled by the subjects' wrist rotation centre.

Each trial included three parts. During the first one subjects were presented with the home orientation bar, with which they were asked to align their object's orientation with a 1° accuracy. The duration of this task was variable and after the stabilization criterium was met, subjects were presented with the target orientation bar and instructed to hit it as accurately as possible within 1 s. After this 1 s, subjects were presented with a static feedback of their object's orientation at the end of the target hitting interval for 1.5 s (yellow bar in Fig. 6.1 A). During the target hitting interval of our first paradigm subjects were provided with visual feedback of the object orientation (blue bar in Fig. 6.1 A), whereas in the second paradigm they weren't. Consequently, in the second paradigm subjects only received information about their motor performance during the 1.5 s feedback interval. Each paradigm consisted of two 20-trial blocks, each preceded by a 30 s resting phase (Fig. 6.1 B).

6.2.3 Analysis of motor behaviour and correlation to neurophysiological signal

f2MOVE monitored the changes of object orientation and by that the positioning of the manipulated object throughout our experimental paradigms. In order to evaluate the learning process in a behavioural context we used this information to estimate and examine the evolution of angle trajectories and angular velocities throughout the trials. Due to the individualised rotation ranges instructed based on each subjects' physical skills, we normalised the instructed movement range (from home to target orientation) between [0°,100°] and adjusted angle trajectories accordingly (overshootings reach >100°, whereas small, insufficient movements reach <100°).

Further behavioural metrics were employed to assess learning based on the root mean squared error between target and object orientation at the end of the target hitting interval (also called end-point-error (EPE)), the reaction time (defined as the time from the initiation of the target hitting interval until the angular velocity surpasses a critical threshold of $0.03°/s$ (RT)) and

angle variability (defined as the mean standard deviation across angular trajectories of different experimental blocks).

End-point-error (EPE) represents a difference between the targeted (expected) and actually produced motor output. In this respect it reflects a performance measure employed by our computational approach in Chapter 3 to drive motor learning. In the present study we thus set out to investigate the neural implementation of the learning mechanism that utilises this behavioural measure. In an attempt to examine whether the progression of EPE is reflected in the evolution of whole brain activation, we formulated and tested the hypothesis presented in Figure 6.7 A.

In this hypothesis behaviour and neuroimaging signal are treated in four subsequent segments of 10 trials each; $C1$ and $C2$ for the first and last 10 trials of the first experimental block respectively, $C3$ and $C4$ for the first and last 10 trials of the second experimental block respectively. (The hypothesis was tested only for the experimental paradigm that provides online visual feedback of movement.) Based on Figure 6.7 A, if learning in behaviour -as reflected through significantly large EPE changes- is primarily observed in the first 10 trials of the experimental paradigm ($C1$), then we expect the brain activation for this segment to be significantly different to the brain activation in the following segments ($C2$, $C3$, $C4$). In contrast to that, we expect the activation differences between the three last segments to be smaller (second activation hypothesis column in Figure 6.7 A). The difference between two compared segments is approximated as the summation of the whole brain activations (see Section 6.2.4) estimated for the contrasts between these segments in both directions (e.g. for $|C1\text{-}C2|$: $C1\text{-}C2$ and $C2\text{-}C1$).

6.2.4 fMRI methodology

In the present study, in order to model the measured BOLD response as a predictable hemodynamic response function (HRF) (Huettel Scott A. 2004) in response to stimulation parameters we employed the tools within the FMRIB's software library (FLS) (Smith et al. 2004). The latter performs standard pipelines of fMRI data pre-processing and analysis (**friston1994**; Monti 2011).

Preprocessing We pre-processed fMRI data in accordance with the FSL pipeline with the aim to correct for motion artifacts, to introduce spatial and temporal filtering and to register functional and anatomical information.

As shown previously in fMRI literature even subtle head movements during a scanning session can cause the detection of false activations (Hajnal et al. 1994). In order to deal with this type of artifacts we used MCFLIRT, FSL's dedicated motion correction tool. The latter tracks and corrects both large and subtle head movements. In particular, substantial movement is

identified by video-looping through the volumes and leads to truncation of the corresponding data segments (Brammer 2001). Smaller movements are corrected by aligning each volume with a reference volume of each series (e.g. first volume of series, Friston, Frith, et al. 1995; Brammer 2001; Jenkinson et al. 2002; Huettel Scott A. 2004. Moreover, in order to regress out temporal changes in image intensity, which are associated with head movements, we can define additional confounds based on the direction-specific estimates of head displacement; these will be added to the low-level analysis as additional explanatory variables. This method enables a removal of the head displacement effects on the analysis, without biasing adversely the statistics (Brammer 2001; Jenkinson et al. 2002).

Based on the realigned and motion-corrected volumes, a mean image T_2^*-weighted image was created co-registered to the anatomical volume to secure that functional and structural information is spatially aligned for each individual subject case. To further enable comparability amongst subjects' fMRI signal, the realigned functional images were spatially normalised to a standard T_2^* template (Friston, Ashburner, et al. 1995). The result are images which are in a standardized space (MNI coordinates), thus allowing us to draw conclusions applicable to larger populations. Moreover, these standard templates are thoroughly labelled, making the assumptions on brain areas much more reliable.

Spatial filtering was implemented based on the convolution between the functional data volumes and a 3D Gaussian filter (Huettel Scott A. 2004), so as to remove high spatial frequency components and increase signal-to-noise (SNR) ratio at the cost of lower spatial resolution.

Temporal filtering is performed with a highpass filter (with cut-off at 100 s, which is embedded in the FSL pipeline. This filtering removes low frequency noise and linear trends, in order to filter out scanner-related and physiological signals, such as cardiac and respiratory cycles and thus provide a higher relevance of the brain signal.

Study design Driven by the purpose of our study, we selected a block design paradigm to describe the evolution of the BOLD response to the experimental stimuli. This paradigm captures how the BOLD signal is modeled as the convolution of the HRF with a binary representation of the presented stimuli (Friston et al. 1994). Stimuli in our experiment are presented in a continuous fashion and distinct blocks of repetitions are intermixed with periods of rest. Contrasting BOLD responses associated with stimulation periods to BOLD responses associated with periods of rest can thus lead to a statistical identification of stimulus-specific whole brain activations (Bandettini et al. 1993; Donaldson and Buckner 2001).

Low-level analysis At this level we analysed each session's (subject's) data. We specifically used the General Linear Model (GLM) hypothesis to explain the raw 4D fMRI signal as the linear combination of the weighted time courses of known explanatory variables (EVs) (Poline

and Brett 2012). EVs represent the different effects we wish to model. Based on FSL's standard implementation of the GLM (FILM, FMRIB's Improved Linear Model), the intensity of any given voxel in response to the stimulation effects we intend to investigate in our paradigm is formulated as follows:

$$Y = X \cdot \beta + \varepsilon \tag{6.1}$$

which in timeseries form is also expressed as

$$\begin{pmatrix} y(t_1) \\ \vdots \\ y(t_N) \end{pmatrix} = \begin{pmatrix} x_1(t_1) & \dots & x_n(t_1) \\ \vdots & \ddots & \vdots \\ x_1(t_N) & \dots & x_n(t_N) \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1(t_1) \\ \vdots \\ \varepsilon_n(t_N) \end{pmatrix} \tag{6.2}$$

where Y denotes the observed BOLD signal at various time points at a single voxel, X the design matrix which includes several components that explain the observed data, β the contribution of each design matrix component and ε the difference between observed (Y) and predicted ($X \cdot \beta$) data. The model's weights, β, are estimated based on a linear regression method (Ordinary Least Squares, OLS) so as to minimise the sum of the squared residuals, ε (Friston et al. 1994). The parameter estimates are divided by their standard error (SE) to provide a t-statistic and subsequently converted into a z-score via standard statistical transformation to generate a z-statistical parametric map, which reflects intensity through its voxelwise z-estimates. These maps are finaly thresholded based on a standardised significance level (2.3) (Woolrich et al. 2001; Genovese et al. 2002; Smith and Nichols 2009; Nichols 2012).

In the present study we defined four EVs -$C1$, $C2$, $C3$, $C4$ (see Figure 6.7 A)- that correspond to BOLD responses acquired during trials [1-10], [11-20], [21-30], [31-40] respectively. We thereby attempted to capture four consecutive phases of our first paradigm (block 1 and 2, with visual feedback, see Figure 6.1 B), which was the part of the experiment from which we obtained the clearest evidence for motor learning (Fig. 6.5).

In order to investigate specific questions regarding the evolution of the acquired neurophysiological signal we set up a distinct contrast design. The latter consists of contrast vectors, each of which is responsible for the creation of a z-statistic image. Thus, to convert a single EV ($C1$, $C2$, $C3$ or $C4$) into a z-statistic image, we set its contrast value to 1 and all others to 0. Aside from this simple contrast we examined pairwise the contrasts amongst all selected EVs, by setting the contrast values of the compared pairs to 1 and -1 and the remaining ones to 0 (e.g. for contrast $C1$-$C2$, $C1$ was assigned to contrast value 1, $C2$ to -1, whereas $C3$ and $C4$ to 0).

High-level analysis We performed high-level analysis to combine low-level analysis results across subjects (Holmes and Friston 1998). A basic prerequisite for analysing brain activation patterns across subjects was to spatially register the data sets to a standard anatomical template (Jenkinson and Smith 2001). We employed FLAME (FMRIB's Local Analysis of Mixed Effects) to model and estimate the variance arising from (i) within session/subject (fixed-effects, FE) measurements and (ii) across session/subject (random-effects, RE) measurements. In particular, FSL implements Markov Chain Monte Carlo (MCMC) sampling to estimate the true random-effects variance and degrees of freedom at each voxel. A mixed-effects model of high-level analysis was considered appropriate for ensuring the compliance of our results with a broader subject population (more than one subject, Beckmann et al. 2003; Mumford and Nichols 2006; Mumford and Poldrack 2007.

Intersection and unification of activation patterns In order to approximate the intersection between activation patterns (z-statistic images) associated to distinct contrast vectors, we used FSL's embedded program 'fslmaths' which allows mathematical manipulation of images. By multiplying one high-level z-statistic image with another one we were able to estimate their common areas of significant activation (z-score$>$2.3).

In order to extract the unification of two activation patterns we implemented an f-test at their corresponding low-level contrast vectors. This test determined at the individual subject level the degree by which the combination of these contrasts is significantly non-zero. Subsequently, we performed random permutations (Winkler et al. 2014), which are used for inference on statistic maps when the distributions of data and/or noise are unknown. This is the case for the f-statistic maps emerging from the low-level analysis. This method thus assumes that random permutations of the data will not produce significant difference on the distribution of values, if there is no correlation between the measured activation and the concurrent stimulation. FSL's tool for nonparametric permutation inference on neuroimaging data is 'randomise'.

We first created a 4D image of all the subject-specific thresholded z-statistic images. Based on this 4D input we performed a nonparametric one-sample t-test by using 'randomise' with a 5 mm variance smoothing option, so as to enhance the method's power, given that we have a low subject population (fewer than 20 subjects).

'Randomise', alongside performing the random permutation technique, employs a cluster correction method named Threshold-free cluster enhancement, so that clusters are enhanced, but not binarized, while keeping the image's voxels intact. It also uses the Family-wise error rate (FWE) cluster correction, in order to maximize the significance of the clusters found.

ROI-based masking of brain activation outputs For a more detailed interpretation of the functional significance of our low-level analysis results, we set out to determine the contribution

of certain cortical regions of interest (ROIs) to the latter. We used FSL's embedded toolbox 'fslmaths' which enables, amongst other operations, a local or global masking of a given input based on a given mask. Our inputs were defined as the low-level z-statistic images acquired for distinct contrast vectors and our masks were created so as to reflect the anatomical location of fourteen distinct cortical ROIs, presumably related to sensorimotor functions. We specifically isolated them from standard anatomical maps embedded in the FSL suite, the Talaraich label maps.

In order to secure the comparability of our masking results across cases (subjects and ROIs) we implemented the following normalization approach: We estimated for each masking operation the ratio between the ROI-specific cluster mass (i.e. the mean z-statistic values across all voxels in a ROI multiplied by the number of significantly activated voxels in that ROI) and the whole-brain cluster mass (the mean z-statistic values across all voxels in the brain multiplied by the number of significantly activated voxels in the brain).

6.3 Results

In the present study we aimed at expanding our previous empirical approaches to sensorimotor learning so as to encompass neurophysiological information. Our objectives included the implementation and employment of an experimental object manipulation platform that would allow the extraction and investigation of neurobehavioural markers of learning in naturalistic motor tasks.

6.3.1 Analysis of behaviour

We instructed human subjects to perform two experimental paradigms, during which they matched rotational movements from a home to a target orientation while holding in their dominant hand a compact cylindrical object. In one of the two paradigms subjects received online visual feedback during the execution of the target-hitting segment of the trials, while in the other one they didn't. In both cases, after the completion of the target-hitting segment subjects received visual feedback of their end-point-error (EPE); that is the RMSE between object and target orientation after 1 s of moving towards the target (Fig. 6.1 A).

We examined the raw measured behavioural data to obtain a first perspective of motor learning evolution in the two paradigms. The mean angle trajectory across trials and subjects for each block revealed a change in the movement pattern from the first to the second half of our first paradigm (blocks 1 and 2 in Figure 6.2 A). This evolution in movement patterns was not that evident in the second paradigm. This arguably suggests that when the brain does not receive online visual feedback of motor performance during the completion of a trial, it may rely on an

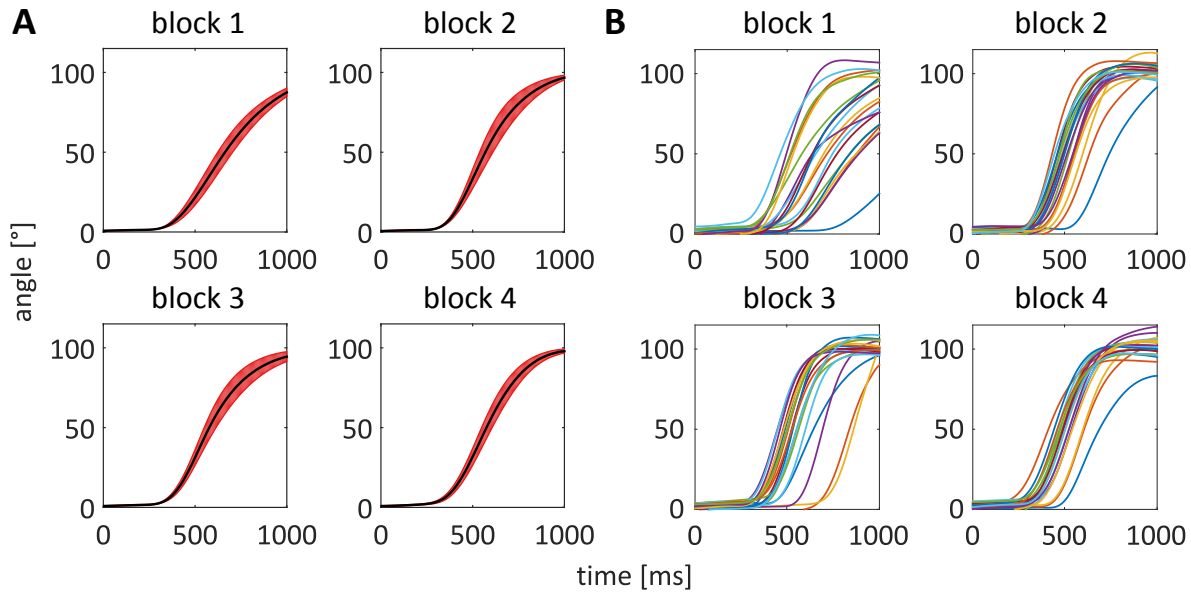

Figure 6.2: Examining raw positioning information: (**A**) Mean angle trajectories for each experimental block (Fig. 6.1 B). A change of the obtained movement pattern from initial to later experimental stages is observed only in the paradigm that provides subjects with online visual feedback of their motor performance. Standard error of mean block trajectory profiles across subjects is displayed in red. (**B**) Angle trajectory profiles at the individual subject level (representative subject selected) reveal a decrease in movement variability from initial to later experimental stages, as is illustrated for both paradigms.

already existing internal model of the task, based on which it can employ similar control policies from trial to trial in a continuous fashion.

On the other hand, in both paradigms, a closer look at the evolution of motor outputs at the individual subject level showed a decrease of movement variability in the latter experimental stages (representative subject data displayed in Figure 6.2 B). This is indicative of an adaptation process, which fosters convergence to a motor output that appears to be stereotypical across subjects.

Our observations on angle trajectories are verified by the examination of angular velocity throughout the course of the experiment. In the first paradigm subjects exhibit an evident change in their mean angular velocity profiles from initial to later experimental stages (blocks 1 and 2 respectively in Figure 6.3 A). Namely, subjects appear to evolve to a close to gaussian velocity distribution. This learning pattern can also be noticed at the individual subject level (Fig. 6.3 B) amongst the first and second block. On the other hand, the velocity profiles in

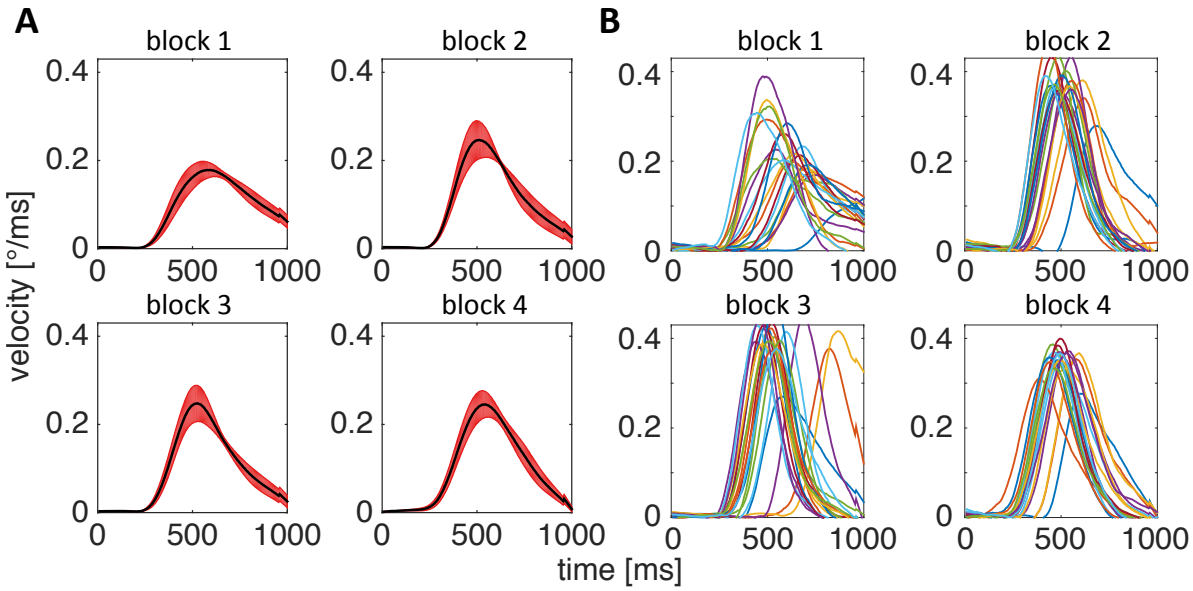

Figure 6.3: Examining raw velocity information: (**A**) Mean angular velocity profiles for each experimental block (Fig. 6.1 B). Evidence of learning displayed predominantly in the first paradigm, throughout which subjects appear to evolve closer to normally distributed velocity profiles. Standard error of mean block velocity profiles across subjects is displayed in red. (**B**) At the individual subject level (representative subject selected) we observe movement variability decrease from initial to later experimental stages for both paradigms (similarly to Figure 6.2 B).

the second paradigm -similarly to our previous findings on position- bare minimal indications of adaptation.

We further employed three behavioural metrics to assess the aspect and evolution of learning in motor behaviour (see Section 6.2.3): (i) the end-point-error (EPE) between the targeted and actually produced object orientation at the end of the target hitting segment of the trial, (ii) the reaction time (RT) and (iii) the variability across angle trajectory profiles of different blocks. We examined the progression of EPE and RT throughout the first two paradigms (Fig. 6.4 A,B) and validated our findings from the inspection of raw position and velocity data. Indeed, the first paradigm displayed a gradual minimization of EPE and RT, primarily during the first 20 trials. The second paradigm provided minimal evidence of progression in the selected behavioural measures of performance, with the RT profile sustaining a constant mean level across both blocks. This reinforces the interpretation that upon lack of online visual feedback during the trial the brain utilises a prior internal model of the task dynamics to essentially repeat motor outputs throughout trials.

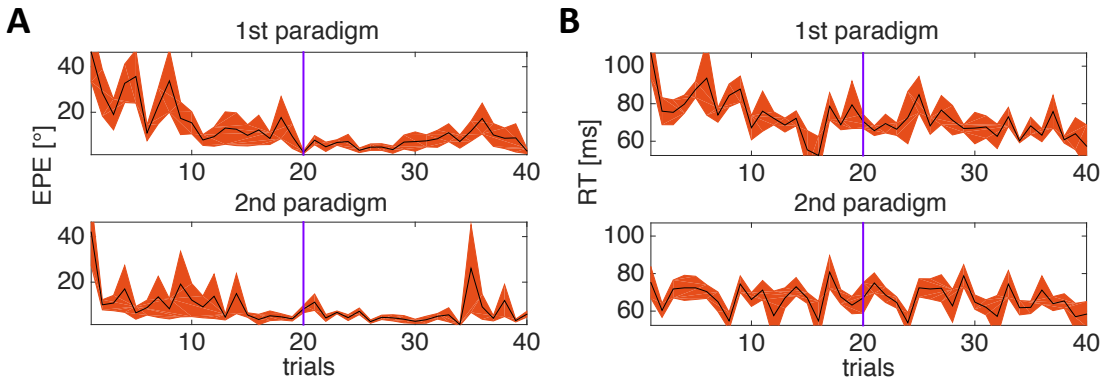

Figure 6.4: End-point-error (EPE) and reaction time (RT) as behavioural metrics: (**A**) Evolution of EPE throughout the two blocks of trials of the two examined paradigms (Fig. 6.1 A, first paradigm 'with online visual feedback': upper panel, second paradigm 'w/o online visual feedback': lower panel). Gradual minimisation of EPE evident predominantly for first paradigm. (**B**) Similar observations were made on the evolution of RT throughout trials for both paradigms.

This allows us to focus the behavioural and neurophysiological analysis of motor learning on the experimental paradigm, which provides subjects with real-time visual feedback of trial execution. A closer view at the binned plots of EPE and RT for blocks 1 and 2 (Fig. 6.5 A, B), shows a decrease of EPE variability at the later experimental stages and a slower learning rate for RT ($t_{1/2}$=7.5 trials for EPE) and ($t_{1/2}$=9.2 trials for RT). A possible explanation for the different learning rates could be that the two behavioural measures are linked to different underlying neural processes, which impose on motor outputs their distinct adaptation components. As a further indication of learning, the mean angle profile variability decreases significantly (Wilcoxon test, $p<0.05$) from the first to the second block of the first paradigm (Fig. 6.5 C). In the second paradigm the mean angle profile variability remains stable at a level that roughly corresponds to the average variability in the first paradigm. Again, this finding potentially suggests the repeated use of a pre-learned control policy without any significant revisions.

In a further attempt to examine the motor learning process based on behavioural criteria we probed a narrative that we discussed previously in Section 3.4 and Section 4.4. That is the scenario that motor learning relies on a synthesis of motor planning processes and lower level mechanisms that optimize motor execution. One possible consequence of such a synthesis would be that complex motor outputs comprise of different adaptation components (Chen-Harris et al. 2008; Körding et al. 2006). This assumption is in agreement with observations in a recent study of Chen-Harris et al. 2008 which showed that there are at least two learning processes involved in

ballistic eye movements; one responsible for the initiation of saccadic trajectories and a second one which tunes the completion of these trajectories by steering their late parts to hit a target.

In order to explore the learning components in object manipulation, we quantified the angular trajectories in the first block of our first experimental paradigm ('with online visual feedback'), which appears to constitute the primary learning period (see Figure 6.5 A): We first aligned all angle trajectories at their individual reaction time points (see criterium in Section 6.2.3) and subsequently isolated the sections expanding between this reaction time point and the end of the trial. We performed this step in order to ensure that the reaction time is not a confounding variable in the analysis of each motor output's evolution.

The isolated movement parts were then rescaled to a common mean duration. Each of the parts was divided into three temporally equal segments, which were labeled according to their corresponding trajectory slopes (Fig. 6.6 A). The evolution of these chord slopes was evaluated throughout different phases of the first block; that is the primary learning phase ('start' for trials [1-7], 'middle' for trials [8-13], 'end' for trials [14-20]). In particular, we estimated the

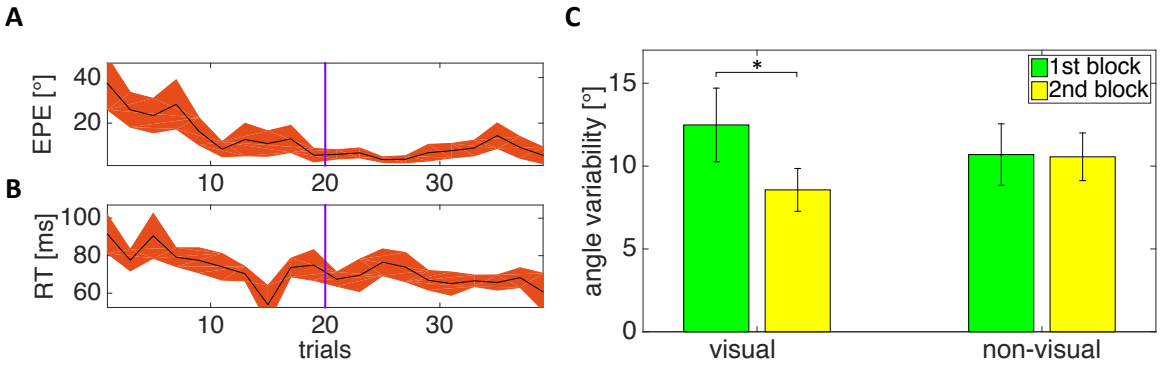

Figure 6.5: Behavioural metrics of motor learning: (**A**) Binned EPE values throughout first experimental paradigm (Fig. 6.1 A) converge to a mean end performance at a learning rate of $t_{1/2}$=7.5 trials (estimated based on the fitted exponential function). Standard error of mean EPE across subjects is displayed in red. (**B**) Binned RT values throughout first experimental paradigm converge to a mean end performance at a slower learning rate of $t_{1/2}$=9.2 trials. Standard error of mean RT across subjects is displayed in red. (**C**) Learning is also evident in the significant decrease of mean angle variablity from the initial to the later stage of the first paradigm ('with online visual feedback'), in contrast to the second paradigm ('w/o online visual feedback') which preserves a stable mean angle variability across trials. Mean angle variability is estimated as the standard deviation across all angle trajectories in one block. Changes between 1_{st} and 2_{nd} block are evaluated based on Wilcoxon rank sum test, from which we obtain a significant difference only for the first paradigm comparison ($p<0.05$).

mean absolute slope change between consecutive trials for each of the three considered block phases. Each of these mean estimates reflects a slope change rate since it is derived from a total slope change over a certain number of trials. Figure 6.6 B reveals that at the beginning of the first block there are significantly higher slope changes for trajectory segment $S2$ than for any other segment. This trend decreases significantly for the middle block phase while all other segment slopes retain stable change levels. On the other hand, in the final block phase only slope changes for trajectory segments $S1$ and $S3$ experience a significant increase. Slope change differences between different block segments and different trajectory parts were evaluated based on the Wilcoxon rank sum test with significance criterium $p<0.05$.

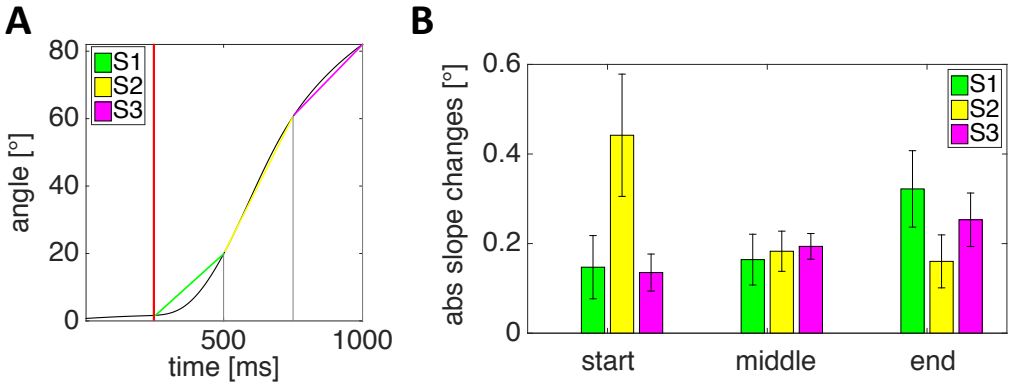

Figure 6.6: Evidence for multilple learning rates: (**A**) In order to investigate the narrative of multiple learning rates associated to the adaptation of distinct motor output components, we divided all aligned angular trajectories into three temporally equal segments ($S1$-$S3$). Alignment point illustrated by vertical red line (see criterium in Section 6.2.3). (**B**) The evolution of the segment slopes (mean slope change between consecutive trials) was estimated for the 'start' (trials [1-7]), 'middle' (trials [8-13]) and 'end' of the examined block (trials [14-20]). Each of these mean estimates reflects a slope change rate since it is derived from a total slope change over a certain number of trials. The results reveal distinct adaptation components related to different learning phases, amongst which (i) an early one that drives an increased modulation of the main middle part of the output trajectory ($S2$) and (ii) a late one that drives an increased modulation of the initiation and finalisation of the output trajectory ($S1$ and $S3$). The statistical significance of slope change differences between segments was evaluated based on Wilcoxon rank sum tests with $p<0.05$. These findings suggest the existence of multiple learning processes, which compose a complex motor output. The neural correlates of these distinct processes are investigated in Section 6.3.2.

These observations suggest that the early and advanced stages of motor learning are associated with significant changes in the adaptation rates of different motor outputs. In fact, the initiation of learning features an increased modulation of the main middle part of the trajectory, whereas the more advanced period of learning is linked to an increased modulation of the start and the ending of the trajectory. Our findings validate our hypothesis by broadcasting the existence of

multiple learning processes, which synthesize a complex motor output.

6.3.2 Behaviour-driven analysis of neurophysiological signal

As a measure of performance, EPE is compliant with the driving force of motor learning in our computational approach in Chapter 3. We therefore selected it in the present study as a metric, whose evolution can be investigated in correlation to the obtained neurophysiological signal (see Section 6.2.3. Based on our fMRI methodology, four z-statistic images are generated to reflect the task-versus-rest brain activation throughout the first experimental paradigm. These images correspond to the explanatory variables (EV) $C1$, $C2$, $C3$ and $C4$ which are associated to BOLD responses acquired during trials [1-10], [11-20], [21-30] and [31-40] respectively. Our hypothesis is that if EPE progression (Fig. 6.5 A) is reflected in the evolution of brain activation, then the contrast between $C1$ and every other EV will capture large activation levels, whereas contrasts measured amongst the later segments of the paradigm ($C2$, $C3$ and $C4$) will yield smaller activation levels (Fig. 6.7 A). For a full account of the differences between different paradigm segments we estimated their cumulative contrasts; that is the summation of the contrast specific activation levels in both directions of comparison (e.g. for $|C1\text{-}C2|$: $C1\text{-}C2$ and $C2\text{-}C1$). Our results in Figure 6.7 B -for average activation levels across subjects- validate our hypothesis on EPE-analogous activation progression, with significantly smaller activation levels observed for the contrasts amongst the last three segments of the paradigm (Wilcoxon test, $p<0.05$).

For the purposes of this study the activation levels in Figure 6.7 B were assessed based on whole brain masking of all contrast specific z-statistic images. We namely estimated the whole brain cluster mass, by multiplying the mean z-statistic values across all voxels in the brain multiplied by the number of significantly activated voxels in the brain (see Section 6.2.4). Future directions can step on our result to attempt a localisation of the detected correlation between EPE and fMRI signal via ROI-based analysis.

In the present work, we undertook an approach to further analyse the neurophysiological correlate of EPE evolution. We focused on activation within the first 20 trials of the examined paradigm ('with online visual feedback'), which appear to constitute the primary learning period in terms of EPE optimization (Fig. 6.5 A). We performed ROI-based masking on the activation patterns (z-statistical images) obtained for contrasts $C1\text{-}C2$ and $C2\text{-}C1$. This operation was implemented by approximating the ratio between each ROI-specific cluster mass and the whole-brain cluster mass (see Section 6.2.4). The relative activation results are illustrated in Figure 6.8 for fourteen different cortical regions which are presumably related to sensorimotor functions.

Our findings show that the brain activation associated to the fast, early-stage rate of EPE

learning (trials 1-10) is formed -comparatively to a later phase of motor learning (trials 11-20)- by a significantly larger contribution of regions such as the parietal, occipital and temporal lobe, frontal areas (such as the PFC and frontal lobe) as well as M1 and premotor cortex ($C1$-$C2$, left panel of Figure 6.8 A). On the other hand, the brain activation detected for the contrast of the plateau-level, later-stage EPE learning to the earlier phase of learning ($C2$-$C1$, left panel of Figure 6.8 B) is formed predominantly by cerebellar contribution. This is also reflected in the slice view of the contrast-specific functional information (z-statistic images, thresholded based on z>2.3 significance criterium) in the right panels of Figure 6.8 A and B.

The interpretation of these results can be that the cerebellum's role in more advanced stages of learning becomes more significant compared to other brain regions that support motor behaviour. Particularly, based on Figure 6.8 B, the cerebellum appears to bare the primary underlying mechanism for the later EPE evolution. This is amenable to the narrative that associates the cerebellum to motor execution, by featuring its role as the locus of internal models of the task and world. It is these models that the cerebellum arguably exploits to generate predicted consequences of action, which in turn form an error (e.g. EPE) when perceived against the real consequences of action (see Chapter 3). In the very early stages of learning, before these

A **Activation hypothesis based on EPE learning evolution**

contrast	activation hypothesis	contrast	activation hypothesis
|C1 - C2|	large	|C2 - C3|	small
|C1 - C3|	large	|C2 - C4|	small
|C1 - C4|	large	|C3 - C4|	small

B

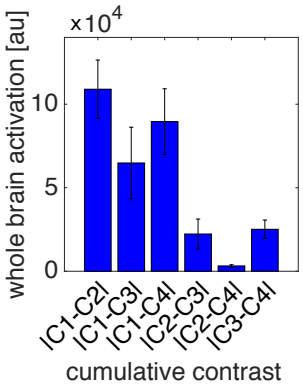

Figure 6.7: A neural correlate for EPE learning: (**A**) We fomulated a hypothesis that brain activation in the examined task (first paradigm) may reflect correlation to the EPE learning evolution. This hypothesis predicts large activation levels for the contrasts between $C1$ (EV associated to first paradigm segment, trials [1-10]) and $C2$, $C3$, $C4$ (EVs associated to second, third and forth paradigm segments, trials [11-20], [21-30] and [31-40] respectively). It also predicts smaller activation levels for the contrasts amongst the last three segments of the paradigm. (**A**) The estimated activation levels for all cummulative contrasts (Section 6.2.4) across subjects validate our hypothesis on EPE-analogous activation progression.

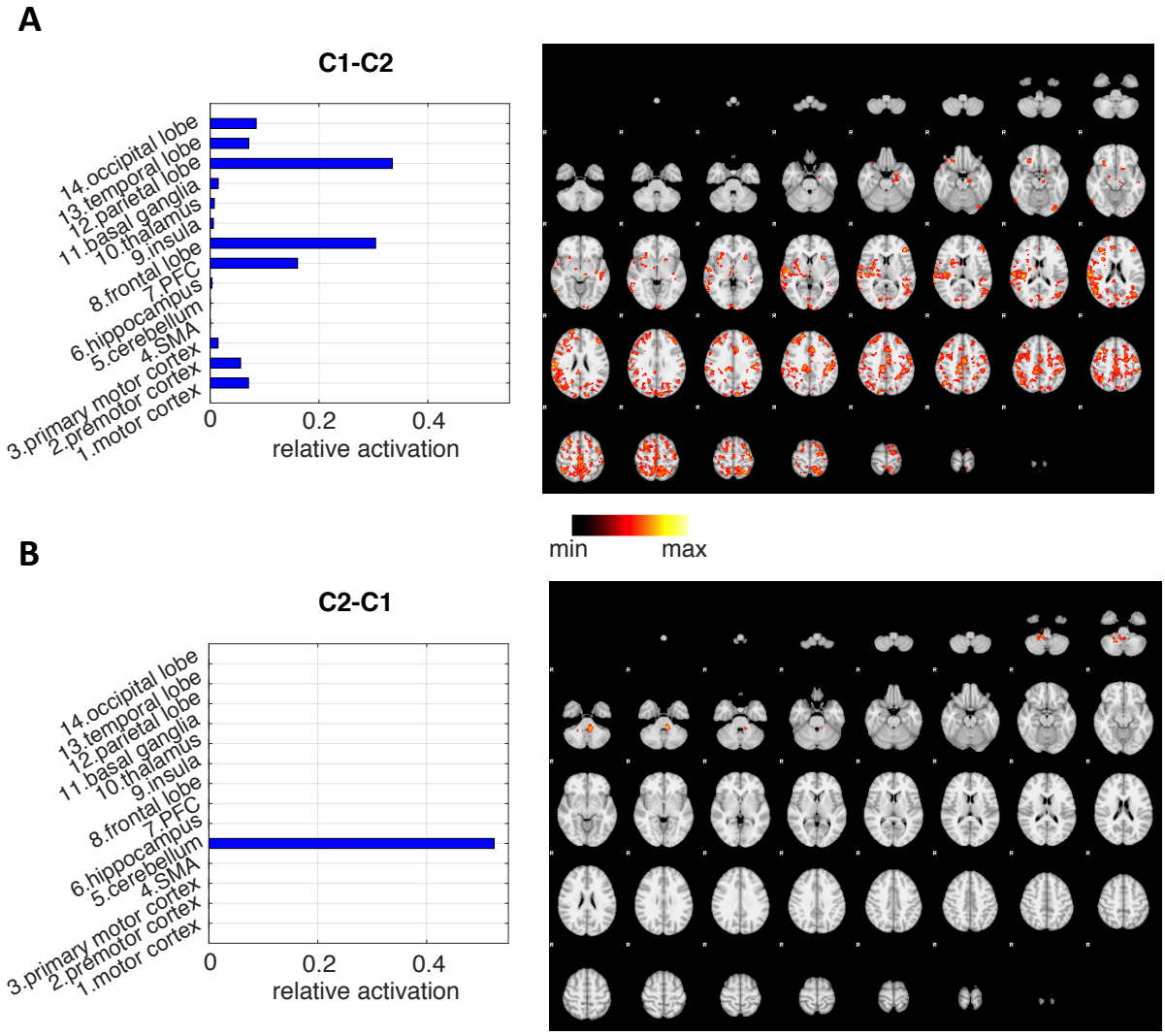

Figure 6.8: Neural correlates of fast, early-stage vs slow, later-stage rates of motor learning (**A**) Left panel: ROI-based masking of the activation patterns obtained for contrasts $C1$-$C2$ reveal a significant contribution of frontal areas, the parietal, occipital and temporal lobe, as well as the motor cortex. (**B**) Left panel: $C2$-$C1$ reveals a predominant cerebellar contribution at the later stages of motor learning as compared to its early stages. Right panels for (**A**) and (**B**): Our findings are validated by the visualisation of the thresholded z-statistic images for the two examined contrasts, with activation patterns resembling attention and planning specific networks (Smith et al. 2009) for the early learning and a cerebellar monopoly for the advanced learning period.

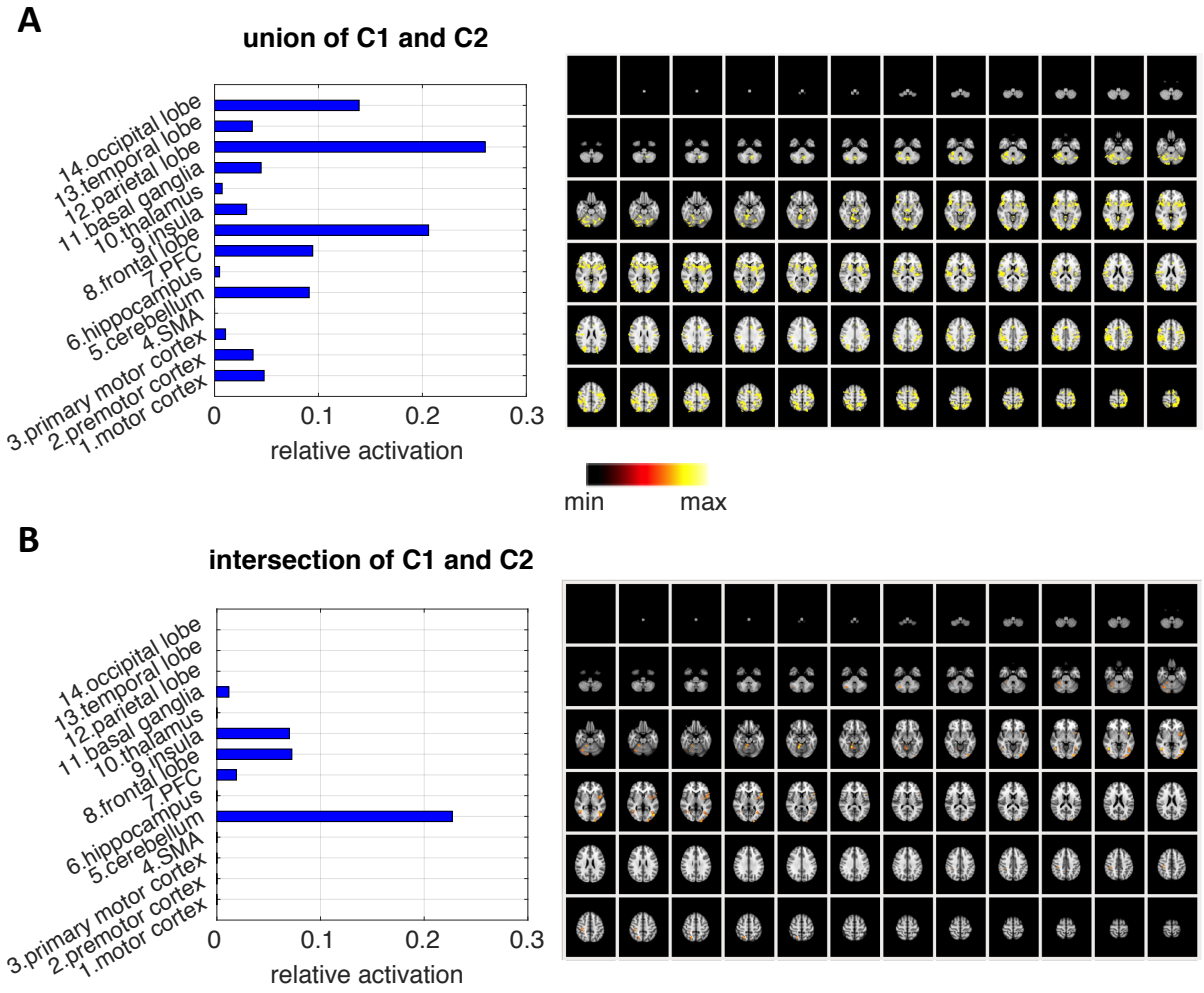

Figure 6.9: (**A**) The union of $C1$ and $C2$ specific contrasts (for trials [1-10] and [11-20] respectively) yields significant relative activations for almost all selected ROIs (left panel). The largest ones are observed for frontal areas, the temporal, occipital and parietal lobe, which have been previously associated to early stage learning (Fig. 6.8 A). Left panel of (**B**): The intersection of $C1$ and $C2$ specific contrasts reveals a primary role of the cerebellum (larger than its relative contribution in the union of $C1$ and $C2$) and a smaller relative role of the frontal areas, PFC, insula and basal ganglia. This verifies that the cerebellum's function is not limited to its evidently primary contribution to the advanced stages of motor learning (Fig. 6.8 B), but extends to all stages of motor learning. Right panels of (**A**) and (**B**): The masking results are validated by a full slice view of the thresholded z-statistic images that we acquired for our merged contrasts ($C1 \cup C2$ and $C1 \cap C2$ estimated based on Section 6.2.4). The activation pattern of $C1 \cup C2$ appears to include mostly uniform and high intensity values. This can be explained by the nature of its derivation technique (random permutations, (Section 6.2.4)) which for a small number of subjects produces a relatively coarse distribution of statistical values. This means that if some areas are consistently highly activated, then they will all be assigned to the top end of the distribution and lumped together with a single statistical value.

internal models are appropriately tuned to task and world dynamics, the cerebellum's role is complemented by the enhanced contribution of neural processes underlying cognitive planning and attention (prefrontal areas), the processing of visual information and formation of visual memories (occipital and temporal lobe), multisensory integration (parietal lobe), formation of motor representations and sensory guidance of movement (motor cortex areas).

In light of our findings on the comparison between neural correlates of early-stage and late-stage learning, we can revisit some of our behavioural analysis results. Specifically, we are able to evaluate from a neurophysiological perspective the obtained evidence for the generation of a complex motor output based on multiple learning processes. As noticed in Figure 6.6, an increased adaptation component of a substantial middle trajectory segment is associated to the early learning stage, whereas a second increased adaptation component of the trajectory start and ending is associated to a later learning stage. In correlation to Figure 6.8 A, we can argue that during the initial learning phase, attention and planning specific neural processes underly the formation of the main part of the trajectory ($S2$); potentialy by selecting the structure of the strategy (e.g. shape, direction) that can steer the motor output towards the target. During a more advanced phase of learning, the cerebellum is the primary locus of processes that refine the initiation ($S1$) and finalisation ($S3$) of the trajectory so as to further optimise accuracy and motor performance.

Additional investigation of the cumulative activation measured for the whole first paradigm block ($C1$ and $C2$ specific contrasts for trials [1-10] and [11-20] respectively) yielded the results in Figure 6.9 A. We notice that across both paradigm segments the majority of the fourteen examined ROIs display a significant degree of activation relatively to the whole brain activation. The largest relative activation levels are associated to frontal cortical areas as well as the temporal, occipital and parietal lobe, which are as seen previously the main activation components linked to early learning stages. On the other hand, an inspection of the common activation between the early and later learning period (intersection of $C1$ and $C2$ specific contrasts, Figure 6.9 B) highlights a principal role of the cerebellum (larger relative contribution to the overall activation pattern than in Figure 6.9 A) and a smaller relative role of the frontal areas, PFC, insula and basal ganglia. These findings connect neural processes that support planning, decision making and error-driven motor execution as central elements of motor learning throughout the whole spectrum of its progression. They also suggest the important part that the cerebellum plays throughout the whole process of motor learning.

6.4 Discussion

In this section we extracted behavioural measures and neural substrates of motor learning from an fMRI study on object manipulation. In particular, we showed how motor learning can be reflected by the evolution of end-point-error (EPE), reaction time (RT) and motor variability; that is metrics which were also shown to be related to the progression of learning in our previous empirical studies (Chapter 3). Furthermore, we provided evidence that throughout our paradigm the progression of brain activation is correlated to the pattern of EPE progression, thereby reinforcing the pivotal role of EPE in our computational perspective of motor learning in Chapter 3. Importantly, the study reveals that the fast, early-stage rate of EPE learning is linked to an enhanced contribution of the frontal and prefrontal cortex, motor cortical areas as well as the occipital, temporal and parietal lobe. On the other hand, the plateau-level, later-stage EPE learning is linked to an enhanced contribution of the cerebellum. We argue that this shift is consistent with the transition from early motor planning and the formation of visuomotor task representations to an eventual automatisation of the motor learning process.

On the technological front of this study we selected fMRI as a powerful neuroimaging tool and worked on enriching its contemporary use for the systemic understanding of motor behaviour. This attempt focused on sidestepping the technical constraints of modern fMRI facilities in hosting advanced motion tracking systems and lead to the development of an fMRI-compatible haptic object manipulation system for closed-loop motor control studies. Our system was initially built in laboratory settings and for the purposes of the present study was customized for a clinical setting (Imperial College London, Clinical Imaging Facility, Hammersmith Campus, London, United Kingdom).

On the front of study design and methodology, we ensured closed-loop task conditions (with and w/o online feedback of motor performance), which encouraged and controlled (via continuous 6 DOF movement monitoring) for motor learning. In order to broaden the applicability of our experimental design we ensured that our haptic interface is adjustable to hosting a wide range of physical real-world objects with variable shapes and dynamics. This enables a flexible approach to naturalistic motor behaviour, a central component of which is object manipulation (Ingram and Wolpert 2011). In terms of data processing, we focused on the development of a set of behavioural metrics, inspired by our computational and empirical insights into the sensorimotor mechanisms underlying object manipulation (Chapter 3). In turn, a part of these metrics was used to guide the design of neural activity metrics.

The outlook of our work would benefit from ROI-based analysis of the acquired raw BOLD response time-series (Poldrack 2007). This method would facilitate a more in-depth approach to determining the level of correlation between a set of behavioural measures of learning and

the signal from distinct cortical areas. The behavioural and neuroimaging markers we extracted here can guide and constrain this otherwise open-ended technique. One interesting direction would thus be to further investigate the neural foundation of our behavioural evidence for the existence of multiple distinct learning processes, responsible for the generation of different parts of the complex motor output (Fig. 6.6), since this result lies in agreement with past evidence for multiple adaptation rates in other motor paradigms (i.e. saccades, see Chen-Harris et al. 2008; Smith et al. 2006). Such an attempt could support our understanding of the subdivision of motor learning into multiple distinct adaptation components and advance our knowledge of the timescales and organisation of these components.

A logical progression of the afore-mentioned technique is to encompass our study's objectives in a network modelling approach to fMRI. Recently, an increasing number of studies has been analysing fMRI datasets in the context of functional connectivity so as to inspect the covariance of signals from different cortical regions (Smith et al. 2011; Smith 2012; Friston 2011; Scott et al. 2015). Network analysis can also be driven and shaped by the results of our present work, particularly via using the latter to identify a set of functional nodes, amongst which we can subsequently conduct connectivity analysis. This direction could lead to a more systemic comprehension of the neural correlates we traced for the fast, early-stage and the plateau-level, later-stage learning pattern of EPE.

In summary, our work determines a closed-loop approach to linking psychophysical motor studies to the neural foundation of behaviour: Our previously established computational insight to motor learning supports the design and testing of learning-specific behavioural measures, which guide the extraction of neural substrates, which in turn update our knowledge and theoretical formalisations of motor learning. This strategy essentially targets the unravelling of sensorimotor functions at both a computational and implementation level and as such, bares significant clinical and technological implications (i.e. guiding contemporary BMI design).

7 Neurobehavioural markers of sensorimotor processes as a diagnostic tool in Friedreich's Ataxia

7.1 Introduction

In the previous section we developed an approach to elucidate the neural foundations of complex motor tasks using neuroimaging. This approach can be extended beyond the examination of healthy human subjects. We can namely employ our clinically customised neurobehavioural markers to reflect and monitor sensorimotor functions in pathological cases. Indeed, understanding the neurophysiological background of neurodegenerative diseases and discovering universal biomarkers capable of diagnosing, monitoring or even predicting their progression is an open challenge of rising urgency (Skovronsky et al. 2006; Gotovac et al. 2014).

Our clinical paradigm in the present work investigates Friedreich's ataxia (FRDA), an autosomal recessive inherited disease, which causes progressive damage to the nervous system. It is the most common inherited ataxia, affecting equally males and females with an occurrence rate of 1 in 40,000 (Koeppen 2013, 2011). It manifests itself in symptoms that range from muscle weakness in arms and legs, coordination impairment, scoliosis to vision and hearing impairment as well as speech problems, heart disease and diabetes (Dürr et al. 1996; Gibilisco and Vogel 2013).

There is currently no cure for FRDA, but there are conservative treatment approaches to manage muscle weakness (e.g. physiotherapy) and treatments for cardiac symptoms. Our understanding of FRDA causes to date involves the erosion of cortical and spinal structures that control motor coordination and some sensory functions. However, the integration of these causes in a fully informed mechanistic model of the disorder's pathogenesis is still pending (Gibilisco and Vogel 2013). Furthermore, diagnosis and monitoring of the disease can be particularly challenging, since progression varies from case to case, along with the very emergence of the first symptoms, often noted between ages 5 and 15 years, but occasionally occurring even in late adulthood.

These issues highlight the need of individualised approaches to diagnosis and treatment of neurodegenerative diseases. The application of personalised medicine represents a prospect gaining increasingly more ground and is expected to yield new insights into the mechanisms of disease pathophysiology, which can be in turn integrated into clinical neurology and guide patient management (Gotovac et al. 2014). Personalised medicine implies the design and use of markers that objectively reflect diverse aspects of the individualised expression of the disease; from behaviour to genetic/biological and environmental factors.

In the case of FRDA, a number of clinical assessments have been developed and used to date to record and evaluate the disease state and progression (Fahey et al. 2007; Schmitz-Hübsch et al. 2006; Montcel et al. 2008; Jacobi et al. 2013; Schmitz-Hübsch et al. 2008). These assessments are intended to test the patients' (i) physiological information (e.g. walking gait, sitting posture etc.), (ii) neurological parameters (e.g. muscle atrophy, spontaneous speech etc.) and (iii) performance in standardised tasks (e.g. finger tapping, finger-nose test, 8 meter walk, 9-hole peg test etc.). Despite the systematic approach to quantify the neurodegenerative disorder based on combinations of such metrics, there are several factors which compromise the validity of the results. One legitimate concern relates to the inherent variability of several of these score-based exercises, which are evaluated based on subjective estimates by clinicians (Fillyaw et al. 1989; Subramony et al. 2005). Due to this variability and the usually slow progression of the disease, such clinical tests employ large study groups and long training paradigms that include numerous task repetitions within sessions and numerous sessions throughout extended periods of time. These elements overload study completion and can even compromise feasibility due to physical and psychological fatigue. All the afore-mentioned considerations reflect the need for new technological means and methodologies to develop markers that objectively quantify different aspects of patients' condition in longitudinal studies. Here, we embark on such an approach to design and clinically test a platform that evaluates and monitors FRDA progression on 7 patients based on high-resolution neurobehavioural markers. The approach is founded on a baseline clinical paradigm on healthy subjects tested on object manipulation (see Chapter 6). It employs our fMRI-compatible haptic system for closed-loop motor control studies, which enables high frequency motion tracking and finely tuned experimental design to support naturalistic motor tasks. Our work thus correlates objective behavioural metrics to neuroimaging data, which are an important tool to assessing FRDA-related pathological brain changes (Stefanescu et al. 2015; Risacher and Saykin 2013; Bonilha da Silva et al. 2014; Mantovan et al. 2006; Szymański et al. 2010).

7.1.1 Core contributions

The primary contribution of our project lies in the attempt to use neurobehavioral metrics of neurodegeneration as tools that characterise the perturbation of sensorimotor functions under pathological conditions. We design and test a systematic approach towards a foundation of personalised disease characterisation that incorporates a computational, behavioral and neurophysiological understanding of diagnostics. This approach is developed as part of a broader study on neurobehavioural biomarkers for longitudinal monitoring of neurodegeneration in FRDA, which acquires additional patient data from daily movement activities (including sleep) based on low-cost, non-invasive wearable body sensor networks (BSN) and embedded high-resolution whole-body motion trackers (Gavriel et al. 2015) [1].

7.2 Aims and methods

In this section our objective was to gain insight into the sensorimotor mechanisms underlying the completion of complex motor tasks in ataxia patients (FRDA) as well as the changes these mechanisms undergo throughout the progression of the neurodegenerative disease. To this end we employed our previously developed methodology to acquiring neurobehavioural markers from a naturalistic object manipulation experiment. In fact, we expanded our approach to investigate the subtle longitudinal behavioural and neurophysiological metrics of neurodegeneration both on a group and individual subject basis. This project was thus designed to set the foundation of advanced techniques to optimizing diagnostics and patient management in FRDA. Work is being conducted as part of a larger currently ongoing study on developing personalised biomarkers to elucidate the progression of neurodegeneration [1].

7.2.1 Experimental setup and data acquisition

For the needs of the study we employed the experimental setup and data acquisition methods used in our previously completed baseline study on healthy human subjects (see Section 6.2.1). We namely completed the study within our fMRI-compatible object manipulation system for motor control studies, as the latter was adjusted to the environment of the Clinical Imaging Facility, Imperial College London, Hammersmith Campus, London, UK.

1. The broader study is undertaken by a collaboration of the Brain and behaviour Lab at the Department of Bioengineering of ICL, the Gene control mechanisms and disease group at the Division of Brain Sciences, Faculty of Medicine, ICL, the Imperial Clinical Research Facility at the Imperial Centre for Translational and Experimental Medicine and the MRC Clinical Sciences Centre, Hammersmith Hospital, London, UK.

7.2.2 Experimental protocol

Seven subjects (4 women; 3 men) participated in the present study; aged 24-56 years. All participants had been diagnosed with Friedreich's ataxia based on clinical criteria and a genetically confirmed GAA- repeat expansion on both alleles of the FXN gene (Libri et al. 2014). All patients provided written informed consent to their participation in the study prior to its initiation. Both patient participation and clinical trial design were approved by the UK Medicines and Healthcare Products Regulatory Agency (MHRA; EudraCT 2011-002744-27), the Riverside Research Ethics Committee (11/LO/0998), and the Imperial College London Joint Research & Compliance Office. All patients were age and sex matched to the healthy subjects in our baseline clinical study (Chapter 6) with two double matches for two of the healthy female participants. Participants completed the object manipulation paradigm instructed by the experimental protocol used in our previously completed baseline study on healthy human subjects (see Section 6.2.2). Similarly to the baseline study, patients' instructed movements were adjusted to their individual maximum rotation range for the extended dominant arm. After the first completion of the experiment, the same paradigm was repeated for three additional times (a baseline check after 3 weeks, and two subsequent sessions after 3 months and 9 months) amounting to four visits altogether. The repetition enabled the collection of longitudinal neurobehavioural data, later analysed to elucidate the progression of ataxia.

7.2.3 Developing longitudinal behavioural markers

In order to monitor the progression of FRDA and its impact on sensorimotor functions of patients we designed a set of behavioural markers, which reflected different aspects of patients' movement patterns and motor performance. These behavioural measures included the mean end-point-error (EPE), the reaction time (RT) and the angle variability (defined as the mean standard deviation across angular trajectories), the way these were determined for Section 6.2.3. They also included the mean trial angular orientation (in [°]), the mean velocity (in [°/ms]) and the mean bimodality coefficient estimated based on Sarle's formula for a finite sample:

$$b = \frac{g^2 + 1}{k + \frac{3 \cdot (n-1)^2}{(n-1) \cdot (n-1)}} \tag{7.1}$$

where n denotes the number of timepoints, g the sample skewness and k the sample excess kurtosis, which are estimated as follows:

$$g = \frac{E(x - \mu)^3}{\sigma^3} \tag{7.2}$$

$$k = \frac{E(x-\mu)^4}{\sigma^4} \tag{7.3}$$

where x denotes the discrete velocity values for each timepoint, μ denotes the mean of the sample velocity, σ its standard deviation and $E(t)$ represents the expected value of the quantity t. For the purposes of the present work, our behavioural measures were estimated as means for the 2nd block of our first experimental paradigm ('with online visual feedback', see Section 6.2.2). We selected only the second part of this paradigm to avoid accounting for the effects of learning which vary from session to session as they are presumably strongest at the very first visit.

Each of the afore-mentioned behavioural metrics was further analysed to provide a deeper insight into the disease-related progression of task performance. In particular, the mean values of EPE, RT, angle variability and bimodality coefficient per trial are gathered for all trials of the second block paradigm to form a sample distribution. Subsequently to this, we applied a regression method that attempted to fit multiple standardised parametric probability distributions to the sample distribution (i.e. Continuous, Beta, Birnbaum-Saunders, Exponential, Extreme value, Gamma, Generalized extreme value, Generalized Pareto, Inverse Gaussian, Logistic, Log-logistic, Lognormal, Nakagami, Normal, Rayleigh, Rician, t location-scale, Weibull, Discrete, Binomial, Negative binomial, Poisson). The latter are sorted according to best fit based on the Bayesian information criterion (BIC) (Bhat and Kumar 2010):

$$BIC = -2 \cdot \ln \hat{L} + k \cdot \ln(n) \tag{7.4}$$

where n denotes the number of data observations, k the number of free parameters to be estimated for each standardised distribution and $\hat{L}$ the maximised value of the likelihood function of the examined model distribution M:

$$\hat{L} = p(x|\hat{\theta}, M) \tag{7.5}$$

where $\hat{\theta}$ are the parameter values that maximise the likelihood function.

Using this regression method, we evaluated for each behavioural measure which standardised distribution provides the majority of best fits (subjects) and selected its fitted parameter values as indices that can be monitored throughout the four clinical visits.

Finally, we also employed this regression method to assess the angle and velocity profiles of each trial. That is we determined a sample trial distribution with the within-trial observations and evaluate which model distribution provides the majority of best fits to the sample distributions across both trials and subjects. This process was repeated for each behavioural measure and

similarly to before, each time the dominant model distribution's fitted parameter values were tracked throughout visits.

Behavioural markers regressed against clinically established SARA scores

Alongside the afore-mentioned behavioural markers, a conventional clinical assessment was also carried out for each patient. This test is called Scale for the Assessment and Rating of Ataxia (SARA) (Schmitz-Hübsch et al. 2006) and represents one of the most common scores used to evaluate the state of the disease. Since SARA is acquired based on subjective estimates of the clinical staff, we aimed at resolving to what degree it can be explained by our objective behavioural metrics, which are acquired with the help of high-resolution motion tracking technology. To this end we applied a regression method that describes SARA scores as a linear combination of our behavioural measures.

$$F(w, \Delta B) = \sum_{i=1}^{n} w_i \cdot \Delta B_i + w_{i+1} \rightarrow \Delta S \tag{7.6}$$

where ΔS denotes the vector of SARA score changes from visit to visit, n the number of behavioural measures, ΔB_i the change of each behavioural measure from visit to visit and w_i the weight assigned to each behavioural measure. The free parameters w_i that provide the best fit of behaviour to SARA were estimated based on a least-square nonlinear curve-fitting approach, which solves the problem:

$$\min_{w} ||F(w, \Delta B) - \Delta S||_2^2 = \min_{w} \sum_{i} (F(w, \Delta B_i) - \Delta S_i)^2 \tag{7.7}$$

where F denotes a vector-valued function (equation (7.6)) of the same size as ΔS. Based on the fitted weights w_i we evaluated which behavioural measures are best reflected by SARA score.

In order to ensure that the fitted weights are not biased by outliers, we performed a leave-one-visit-out cross-validation, based on a Gaussian Process regression that was trained iteratively on all but one patient visits and predicted (was tested on) the one visit left out. The objective function optimised for each training iteration was equation (7.7).

7.2.4 fMRI methodology

We used the same general lines of fMRI methodology implemented for Section 6.2.4 based on pre-processing and data analysis tools within the FMRIB's software library (FLS) (Smith et al. 2004). In this study, however, the contrast design set to investigate the evolution of the neurophysiological signal was defined with the help of two EVs $B1$ and $B2$, representing the

BOLD responses acquired for the first and second block of the first paradigm ('with online visual feedback') respectively. In order to capture the task-versus-rest brain activation throughout each paradigm block, we set the contrast value of the corresponding EV to 1 and the other one to 0. This operation allows the conversion of each EV ($B1$, $B2$) into a z-statistic image on a lower-level of analysis performed for each subject individually. We subsequently performed two steps of high-level analysis: (i) The first one (also called middle-level analysis) aimed at combining low-level analysis results within each subject to perform comparisons across visits. The latter were implemented in a pairwise fashion within a contrast design that set the contrast values of the compared pairs to 1 and -1 and the remaining ones to 0 (e.g. for contrast $B2_{V2}$-$B2_{V1}$, $B2_{V2}$ was assigned to contrast value 1, $B2_{V1}$ to -1, whereas $B2_{V3}$ and $B2_{V4}$ to 0; where V_i denotes the indexed visit). (ii) The second step aimed at combining the middle-level analysis results across subjects (Holmes and Friston 1998).

We selected a fixed-effects model of high-level analysis to account for the variance arising from the within subject measurements and a mixed-effects model to account for the variance arising from both the within and across subject measurements (see Section 6.2.4).

7.2.5 Progression of neurophysiological signal and regression to behaviour

The z-statistical images that reveal task-versus-rest activation for each clinical visit underwent ROI-based structural masking to provide information about the role of selected areas in the progression of disease-related neurophysiological changes. These images were estimated both on a group and individual patient level. The group level results were used to evaluate the overall neurophysiological differences between ataxia patients and healthy subjects (see our baseline study Chapter 6) as well as the neurophysiological differences between consecutive visits.

The individual patient level results were used to implement a regression against our previously developed behavioural metrics. We specifically attempted to describe each behavioural measure j as a linear combination of ROI-specific activations:

$$F'(p_j, \Delta A) = \sum_{i=1}^{m} p_{i,j} \cdot \Delta A_i + p_{i+1,j} \rightarrow \Delta B_j \tag{7.8}$$

where j denotes the behavioural measure index, ΔB denotes the vector of changes from visit to visit for each behavioural measure, m the number of behavioural measures, ΔA_i the brain activation (z statistic) changes from visit to visit and p_i the weight assigned to each ROI-specific activation. The free parameters p_i that provide the best fit of ROI-specific activations to behaviour were estimated based on a least-square nonlinear curve-fitting approach (see Equation (7.7)).

Like before, to ensure that fitted weights are not biased by outliers, we performed a leave-one-

visit-out cross-validation, based on a Gaussian Process regression that was trained iteratively on all but one patient visits with objective function equation (7.8) and was tested on the one visit left out. The regression results were evaluated based on goodness of fit (R^2 between $F'(p_j, \Delta A)$ and ΔB_j for all iterations) to support conclusions on which behavioural measures are best explained by our acquired neurophysiological signal and which cortical regions represent the strongest correlates to these aspects of behaviour.

7.3 Results

In the present study we embarked on an approach to gain insight into the neurophysiological processes involved in FRDA-related neurodegeneration. To this end we employed our previously developed baseline clinical paradigm on object manipulation that was performed by healthy subjects in Chapter 6. The paradigm was repeated four times (the initial session and three revisions after 3 weeks, 3 months and 9 months). Patients were instructed to perform both paradigms of the original experiment ('with online visual feedback' and 'w/o online visual feedback', Section 6.2.2) and their instructed rotation movements between home and target orientation were adjusted to their individual maximum rotation range for the extended dominant arm.

In Figure 7.1 A, B we observe their angle profiles for all blocks of the experiment in the 1^{st} visit. We notice that both the movement range and movement pattern of ataxia patients are different to the behavioural data acquired from healthy subjects (Fig. 6.2 A, B). Patients cover smaller rotations while moving to the target and there is only little evidence of progression in the angle trajectories from the first to the second block; that is within the first paradigm. Throughout the second paradigm ('w/o online visual feedback') on the other hand, patients seem to only repeat similar movement patterns. This arguably suggests that since the brain does not receive online visual feedback of motor performance during the completion of the trial, it may rely on an existing internal model of the task, based on which practically the same control policies are selected from trial to trial in a continuous fashion. These repeated policies appear to be less efficient than the ones observed for healthy subjects (evident in the higher EPE, Fig. 7.2 A), presumably because ataxia is known to affect the cerebellum, whose role is linked to internal models of the task and the world (Wolpert et al. 1998; Kawato et al. 2003; Albus 1971; Marr 1969; Kawato et al. 1987; Jordan and Rumelhart 1992).

Equivalent observations can be made for the velocity profiles of ataxia patients (Fig. 7.1 C, D). Patients generally exhibit smaller peak velocities than healthy subjects (Fig. 6.3 A, B). The behavioural outputs in Figure 7.1 C, D reveal little evidence of learning, primarily noticed in the first paradigm ('with online visual feedback'), where velocities experience a decrease in bimodality while progressing from the first to the second block of the session. Furthermore, this

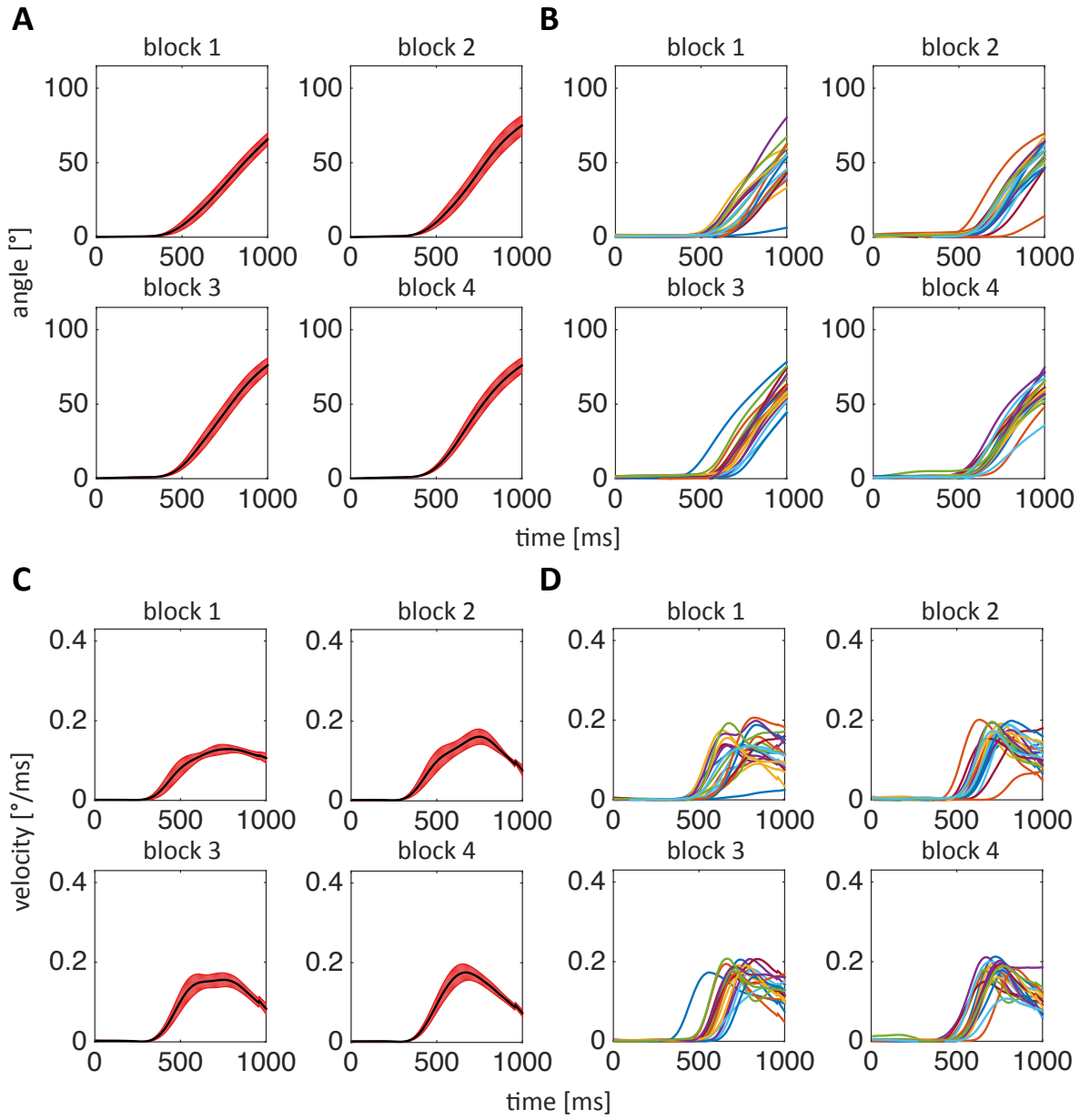

Figure 7.1: Angle and velocity profiles for first clinical visit: (**A**) Mean angle trajectories across patients for the first and second block of the two experimental paradigms ('with online visual feedback' and 'w/o online visual feedback') deviates from baseline behaviour of healthy subjects (Fig. 6.2 A). Small evidence of behavioural changes is only observed throughout the first paradigm. (**B**) Angle trajectories for exemplary individual patient also exhibiting the small rotation range noticed in mean group performance.(**C**) Mean velocity profiles across patients also deviate from healthy subjects' performance: smaller peak velocity, larger bimodality. (**D**) Velocity profiles for exemplary individual patient reveal little evidence of learning, primarily exhibited by the decrease in behavioural variability while transitioning from first to second block.

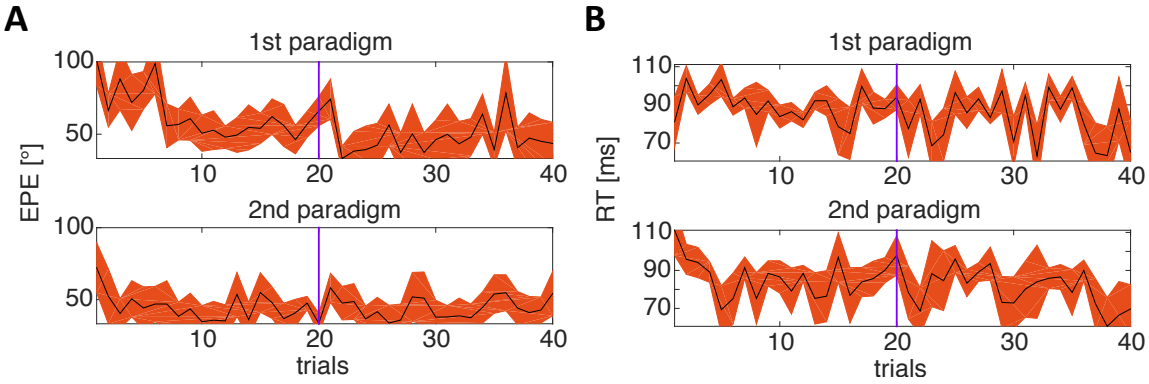

Figure 7.2: (**A**) End point error evolution throughout first and second paradigm of first clinical visit. Patients display significantly higher and noisier EPE levels than healthy subjects (Fig. 6.4 A). Moreover, there is limited evidence of progression to improved performance, noticed mainly for the first paradigm ($t_{1/2}$=5.4 trials estimated based on the fitted exponential function). (**B**) Similarly to EPE, RT also presents larger and noisier values than the equivalent measurements in our baseline study. In fact, here RT cannot be used as a behavioural metric of learning, since it doesn't progress to stable lower values for neither of the two paradigms.

transition between first and second block is linked to a small decrease in behavioural variability, as is mainly evident in the individual patient data (Fig. 7.1 D). Patients' velocity profiles also deviate from the close-to normally distributed velocity shapes extracted for healthy subjects in our baseline study (Fig. 6.3 A, B).

The evolution of EPE in Figure 7.2 A verifies that there is little progression in patients' performance throughout the experiment and this is mainly observed throughout the first paradigm. In fact, in the second paradigm ('w/o online visual feedback') patients seem to preserve noisy but stable EPE levels; something which lies in agreement with the corresponding angle profile evolution. This absence of performance progression reinforces the scenario that without visual feedback the brain relies on an existing internal model of the task, based on which it repeats the same control policies from trial to trial. In general, EPE-specific performance carries larger noise than the equivalent performance of healthy subjects (Fig. 6.4 A), which is preserved in the second block of the two paradigms, suggesting repetition-related fatigue. Additionally, the overall EPE levels of patients are significantly higher than the EPE of healthy subjects.

Similarly to EPE, the RT levels for patients are higher than for healthy participants (Fig. 6.4 B), as seen in Figure 7.2 B with the patients exhibiting great variability from trial to trial until the end of both paradigms. In fact, in contrast to our baseline study, RT here does not seem to be a candidate index of learning, since it doesn't display a systematic progression to stable values

throughout any of the paradigms. In this respect RT also differs from the evolution of EPE, which at least in the first paradigm converges to lower values with $t_{1/2}$=5.4 trials (estimated based on the fitted exponential function).

7.3.1 Tracking neurodegeneration with behavioural markers

In order to monitor the behavioural effects of neurodegeneration throughout visits we first tracked the evolution of the average EPE, RT and bimodality coefficient based on Section 7.2.3, which captures the velocity patterns and thereby the control policies employed by subjects to complete the object manipulation task. For the purposes of the present analysis, our behavioural measures were estimated as means across subjects for the 2_{nd} block of our first experimental paradigm ('with online visual feedback', see Section 6.2.2) so as to avoid accounting for the effects of learning which were expected to be different amongst visits (stronger in the first one and less evident in the subsequent ones).

Figure 7.3 A, B, C illustrates in box plots the progression of the EPE, RT and bimodality coefficient as the latter is estimated based on Equation (7.1). For all measures there is a significant difference between healthy subject and patient performance (Wilcoxon test, p<0.05). Healthy

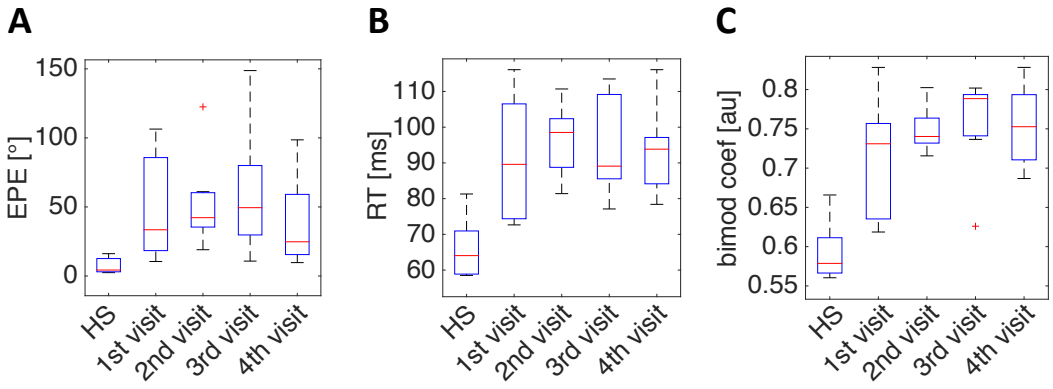

Figure 7.3: Group level progression of three behavioural metrics estimated based on Section 7.2.3: (**A**) End-point-error (EPE), (**B**) Reaction time (RT), (**C**) Bimodality coefficient b estimated based on Equation (7.1). All three measures display a significant difference between heatlhy and patient performance (Wilcoxon test, p<0.05), with healthy participants displaying close to normally distributed velocity patterns (median b close to 0.56), whereas patients display bimodal or multimodal velocities (median b significantly larger than 0.56). Furthermore, pairwise comparisons between the first and any subsequent visit show a decrease in data variability, as the latter is determined by the difference between first and third quartile. A deeper understanding of the behavioural impact of neurodegeneration appears to call for analysis at the individual level, since group level results do not reveal clear trends, arguably due to varying disease progression patterns amongst patients.

participants perform better in terms of error and reaction time criteria and employ close to normally distributed velocities (median b close to 0.56 which is the uniform distribution value), whereas ataxia patients display evidence of bimodal or multimodal velocity patterns (with median b significantly larger than 0.56). For all three behavioural measures, pairwise comparisons between the first and any subsequent visit show a decrease in data variability, as the latter is determined by the difference between first and third quartile. However, variability is not consistently decreasing throughout the last three visits. Similarly, there is no evident trend of increase or decrease in the medians. These findings suggest that besides a basic evalution of behavioural differences between ataxia patients and healthy participants, it appears difficult to infer a more detailed impact of neurodegeneration on our behavioural metrics while still treating the latter on the group level.

In order to account for the potentially varying disease progression patterns amongst patients we proceeded with behavioural analysis on the individual participant level. Additionally, for a fuller understanding of the behavioural impact of neurodegeneration we performed further feature analysis on the all afore mentioned behavioural metrics. This was implemented based on Section 7.2.3 by using the mean values of EPE, RT, angle variability and bimodality coefficient for each trial in the second block of the first paradigm. The emerging block sample distributions for each behavioural measure were fitted by a set of standardised parametric probability distributions in a regression method, which sorted all fitting results and selected the fitted parameter values for the model distribution that provided the majority of best fits throughout subjects. A similar operation was implemented to assess the angle and velocity profiles of each trial. Here, the sample distributions were created based on within trial observations and thus the fitted parameter values were estimated based on our regression for each different trial. Consequently, the average of these trial-specific estimations throughout the second paradigm block allowed us to extract mean behavioural parameter values for each visit. Altogether, EPE, RT, bimodality and angular velocity were best fitted by normal distributions, whereas angle variability and angular position were best fitted by logistic distributions.

In order to evaluate which of the twelve derived behavioural metrics are mostly relevant to neurodegeneration based on established clinical criteria, we regressed them against the previously measured Scale for the Assessment and Rating of Ataxia (SARA, see left panel of Figure 7.4 C), a common FRDA-specific test that examines gait, stance, sitting, speech disturbance, finger chase, nose-finger test, fast alternating movements and heel-shin slide. In particular, we derived the linear combination of behavioural metrics that provided the best fit to the SARA scores for all patients and visits (Section 7.2.3) and subsequently validated its goodness of fit. Indeed, as illustrated in the right panel of Figure 7.4 C, there is a good correlation between the actual Δ SARA changes from visit to visit and the estimated Δ SARA changes, which are based on the

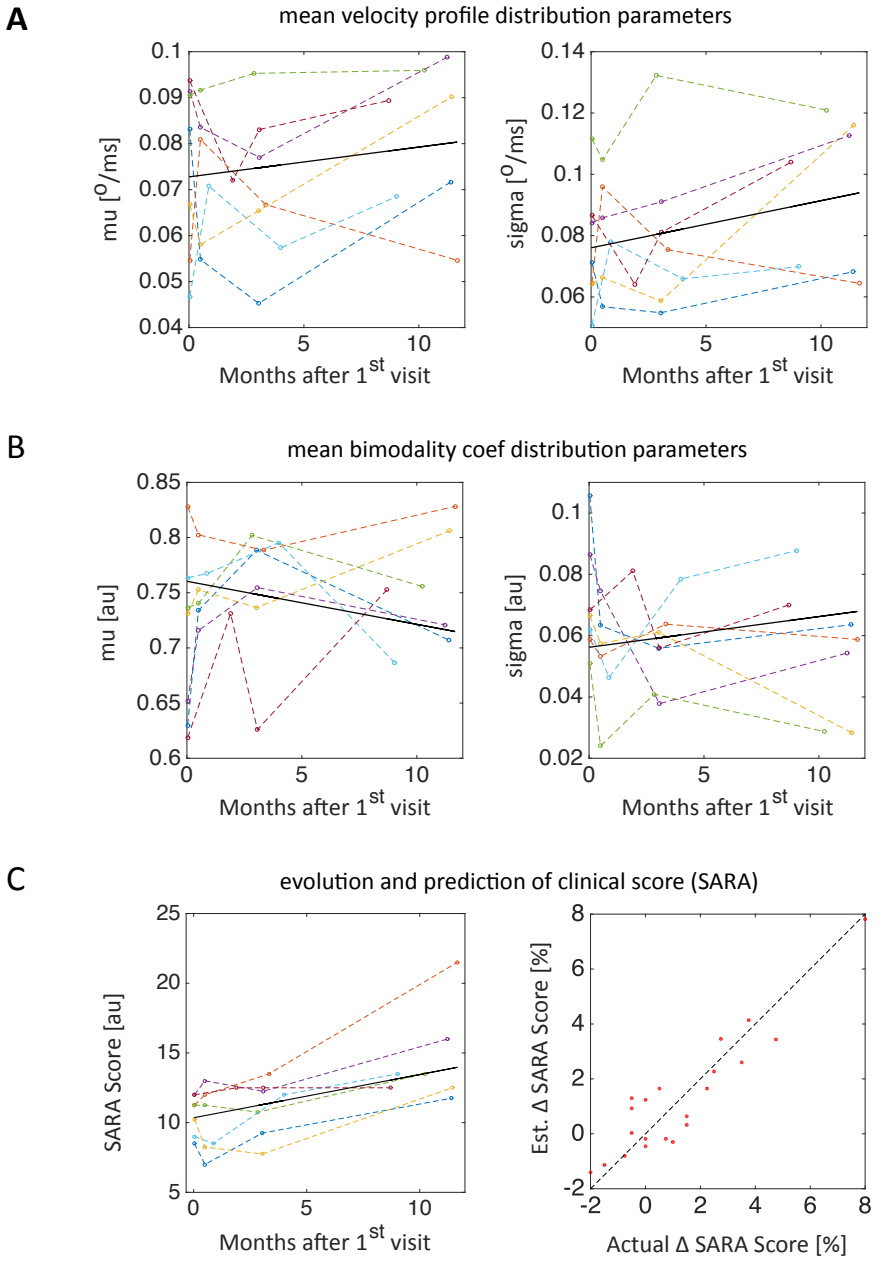

Figure 7.4: Tracking neurodegeneration with behavioural markers: We illustrate behavioural metrics with enhanced contribution in explaining the clinically established SARA score (see regression details in Section 7.2.3 and Section 7.3.1). (**A**) Mean and standard deviation of the velocity profile distribution and the (**B**) bimodality coefficient distribution. Interpretation of (A) and (B) trends included in Section 7.3.1. (**C**) Individual SARA scores progress with less noisy trends than our behavioural measures (left panel), possibly due to the inherent bias in the subjective evaluations of clinicians.The process that allows us to determine which ones of our behavioural measures best explain the evolution of SARA scores is described in Section 7.2.3. This operation involves a regression of a linear combination of all our behavioural measure changes from visit to visit against the SARA score changes from visit to visit. We secure the statistical validity of our method with a leave-one-visit-out cross-validation, which is iteratively trained on all but one patient visits and tested on the visit left out. For each regression iteration the actual Δ SARA change is thus plotted against the estimated Δ SARA change for the tested patient visit. The cumulative plots provide evidence of high correlation between actual and predicted clinical scores with R^2=0.8435 (right panel).

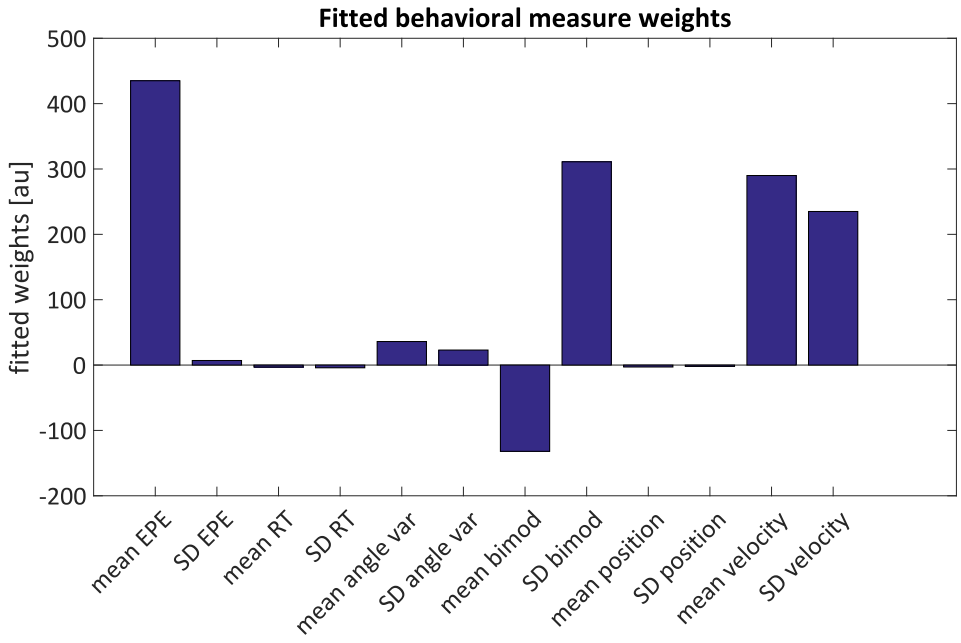

Figure 7.5: Fitted behavioral measure weights. The parameter estimates were derived based on a leave-one-visit-out cross-validation method.

fitted combination of behavioural metrics (R^2=0.8435).

In particular, our regression results based on a leave-one-visit-out cross-validation showed that SARA scores can be explained by an enhanced contribution (large fitted weight values) of the mean EPE as well as the mean and standard deviation of the velocity profiles and bimodality coefficients (Figure 7.5). We observe their evolution throughout time in Figure 7.4 A, B (the mean EPE evolution was already illustrated in Figure 7.3 A).

We notice that in contrast to SARA scores which display a clearer evolution trend (left panel of Figure 7.4 C), the parameters of velocity profile distributions and bimodality coefficient distributions experience a noisier progression with patient-specific outliers. The evolution trends of these metrics are illustrated by the robust fits (black linear plots in Figure 7.4 A, B). In the majority of cases, the mean velocity undergoes a slight increase while the mean bimodality coefficient a slight decrease throughout visits (left panels of Figure 7.4 A and B respectively). This can arguably be justified as a subtle improvement in patients' control strategies because of long-term learning, since ataxia does not necessarily affect velocity specific performance and motor strategy optimisation, but instead accuracy and variability in motor coordination. Indeed, the effect of neurodegeneration on accuracy and performance repeatability is illustrated

on the right panels of Figure 7.4 A and B, which show an increase in the standard deviations of velocity and bimodality coefficient distributions.

These findings suggest that our estimated metrics represent more realistic behavioural indices of neurodegeneration, which -contrary to the potential bias in SARA scores due to subjective clinical evaluations- reflect considerable intra- and inter-subject variability, but can still capture the subtle behavioural changes imposed by the progression of ataxia.

7.3.2 Neurophysiological findings

Based on our fMRI methodology in Section 7.2.4 and Section 6.2.4 we examined the contrast in task-versus-rest activation during the second paradigm block between healthy subjects and ataxia patients. The former activation pattern was derived based on high-level analysis across all healthy participants, and the latter based on high-level analysis across all patients within each visit and subsequently via the combination of all visit-specific results via a random permutation method (see Section 6.2.4).

The bidirectional contrasts between the two z-statistical images are visualized in Figure 7.6 B and C and further analysed in ROI-specific relative activations (Section 6.2.4) in Figure 7.6 A. We observe that the region primarily relevant to healthy subject specific activation is the parietal lobe. In fact the relative contribution of the parietal lobe to the healthy subject specific versus patient specific activation (blue bars) is significantly higher than its contribution to the inverse contrast activation (yellow bars). Other regions related to healthy subject specific activation are the frontal lobe, PFC, cerebellum, occipital lobe and motor cortex, for which the healthy subjects versus patients specific contrasts are at comparable values levels to their inverse contrasts.

However, we need to note the fact that the patient versus healthy subject specific whole brain activation amounts to a much more extended pattern than the equivalent inverse contrast activation. Consequently the absolute contributions of the frontal lobe, PFC, cerebellum and occipital lobe to the patients versus healthy subjects contrast is expected to be larger than these regions' absolute contribution to the inverse contrast. These findings suggest that sensorimotor processes underlying object manipulation in healthy subjects reveal a larger contribution of ROIs supporting multisensory integration (parietal lobe). Areas responsible for the formation of motor representations and sensory guidance of movement (motor cortex), areas underlying cognitive planning, attention (frontal and prefrontal cortex) and visual processing (occipital lobe) as well as the area commonly featured as the locus of internal models of the task and world (cerebellum) appear to be involved on a proportionally comparable degree in the brain activation of both healthy subjects and patients across all visits.

On the other hand, the area, which appears to contribute relatively more to the patients versus

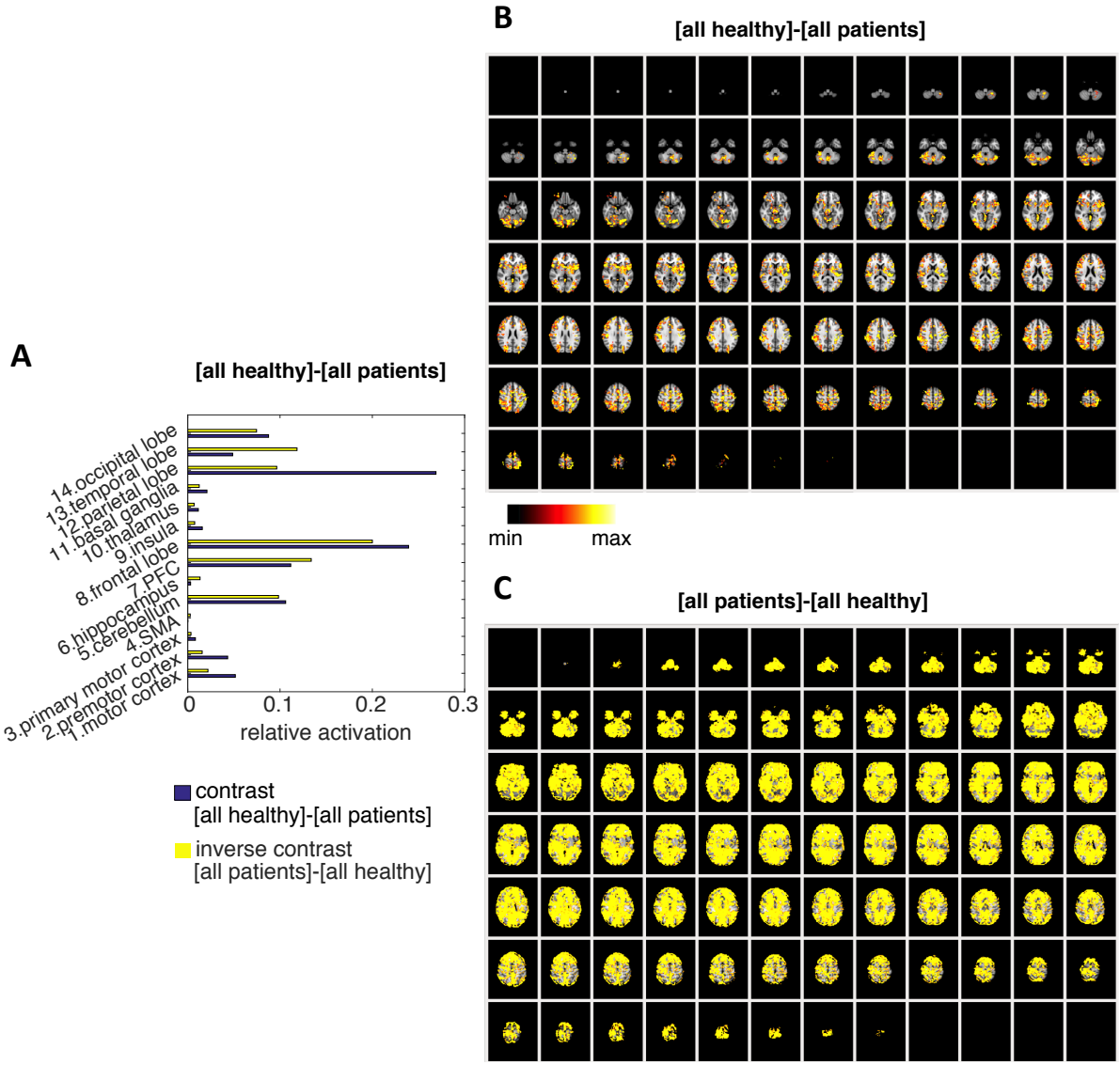

Figure 7.6: Healthy subject versus patient specific contrast activations: (**A**) The contrast activation is displayed in both directions of comparison ('healthy subject versus patient' in blue bar, 'patient versus healthy subject' in yellow bar). Among other observations, we note a larger contribution of the parietal lobe (linked to multisensory integration) to the sensorimotor functions of healthy subjects (compared to patients) and a larger contribution of the temporal lobe (linked to the formation of visual memories) to the activation patterns of patients (compared to healthy subjects). (**B**) Visualisation of healthy subject versus patient specific contrast activation pattern (thresholded z-statistic images). (**C**) Visualisation of patient versus healthy subject specific contrast activation pattern (thresholded z-statistic images). This activation pattern is evidently extended and uniform in intensity, presumably because it represents the unification of all patient visit results and also due to the nature of its derivation via random permutations (Section 6.2.4). In fact, 'randomise' produces for a small number of subjects a relatively coarse distribution of statistical values; that is a small number of possible significant results. This means that if some areas are consistently highly activated then they will all be at the top end of the distribution and lumped together with a single statistical value.

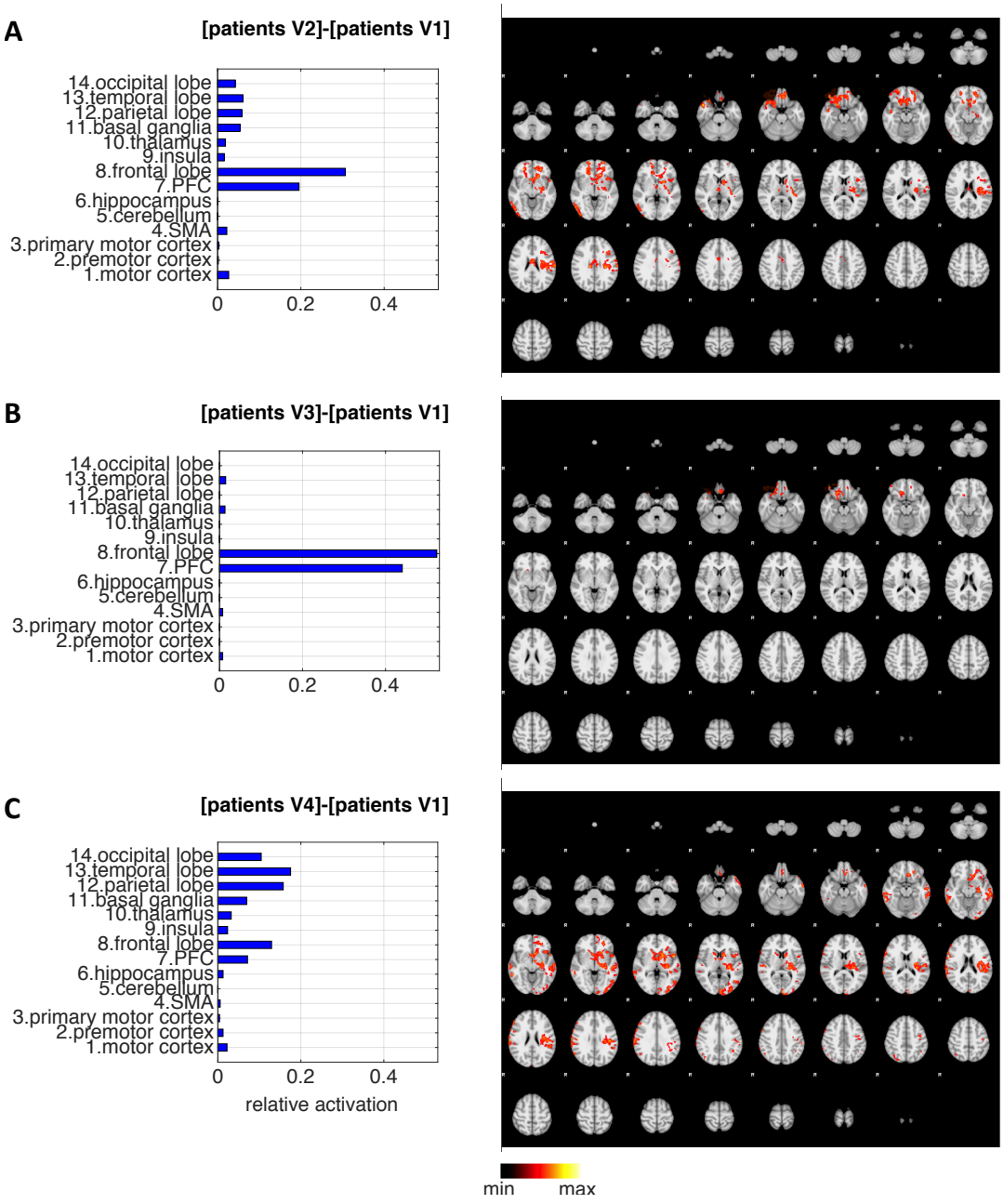

Figure 7.7: Neurophysiological indices of ataxia progression: Right panels display the contrasts in whole-brain activation patterns between consecutive patient visits and left panels the ROI-based analysis thereof (estimated as relative ROI activation with regard to whole brain activation (Section 6.2.4)). (**A**) Second versus first visit contrast activation. (**B**) Third versus first visit contrast activation.(**C**) Fourth versus first visit contrast activation. Overall, the initial increase of prefrontal areas activation eventually subsides and is replaced by an increased activation in areas involved in the formation and processing of visual information as well as areas involved in multisensory integration.

healthy subjects specific contrast than to the inverse contrast is the temporal lobe. This result arguably suggests that the damage experienced by patients in processes supporting motor control and learning through sensorimotor integration and the update of internal task models, may be compensated by an enhanced role of visual memories of the world and task.

In order to gain a deeper insight into these cumulative comparison results, we subsequently analysed the inter-visit activation contrasts to examine the neurophysiological changes occurring within our patient group throughout the progression of the disease. Figure 7.7 displays the contrasts in whole-brain activation patterns (z-statistic images) between consecutive patient visits (right panels) and the ROI-based analysis thereof (left panels). The ROI-based analysis reveals a significant increase of frontal lobe and PFC activation from first to second (Figure 7.7 A) and first to third (Figure 7.7 B) visit. The first contrast (Figure 7.7 A) may simply reflect motor learning effects since the first two visits are only 3 weeks apart. Our findings also reveal that the frontal lobe and PFC dominance in the third visit eventually subsides in favor of a significant relative activation increase of the occipital lobe, temporal lobe and parietal lobe in the fourth visit. Activation contrasts of visits n to $n+1$ were noted but were filtered out by FSL's high level analysis as not significant enough.

Our observations suggest that throughout the patient visits in this study, an initially enhanced relative contribution of areas that support motor planning, decision making and attention (frontal and prefrontal) is gradually replaced by an enhanced contribution of areas commonly supporting the formation and processing of visual information/memories (temporal and occipital lobe) as well as of areas involved in multisensory integration which in turn drives motor control (parietal lobe). This transition may imply a switch from a initially dominant cognitive component in motor control to a more automatised execution of the instructed task. However, interpretations cannot be safely formed at the group analysis level due to the differences of disease state and progression across patients.

7.3.3 fMRI regression against behavioural markers

The ROI-based masking results on individual patient activations were regressed against our previously developed behavioural measures (Section 7.2.3). Our aim was to investigate whether some of our behavioural markers can be described as a combination of ROI-specific neurophysiological signals (Section 7.2.5). The regression results provided the highest fits of fMRI activation to behaviour for EPE with R^2=0.7882 (Fig. 7.8 A), followed by RT with R^2=0.7231 and angle variability with R^2=0.7104. All other behavioural measures were fitted with R<0.7.

We notice that EPE is a metric strongly correlated both to a clinically established ataxia test (SARA) and the evolution of ROI specific neurophysiological signal. However, the fact that

other behavioural measures lead to weakest fitting results, does not support conclusive conclusions about their relevance to the disease. In fact, several of these behavioural metrics have already been proven strongly correlated to the progression of clinical SARA scores (right panel in Figure 7.4 C) and therefore still pose a valid role as disease-related biomarkers. This means that the extraction of their definite neural correlates possibly calls for a further ROI-based processing of fMRI time series and regression involving a network level (Stam 2014) formalisation of neurophysiological activation.

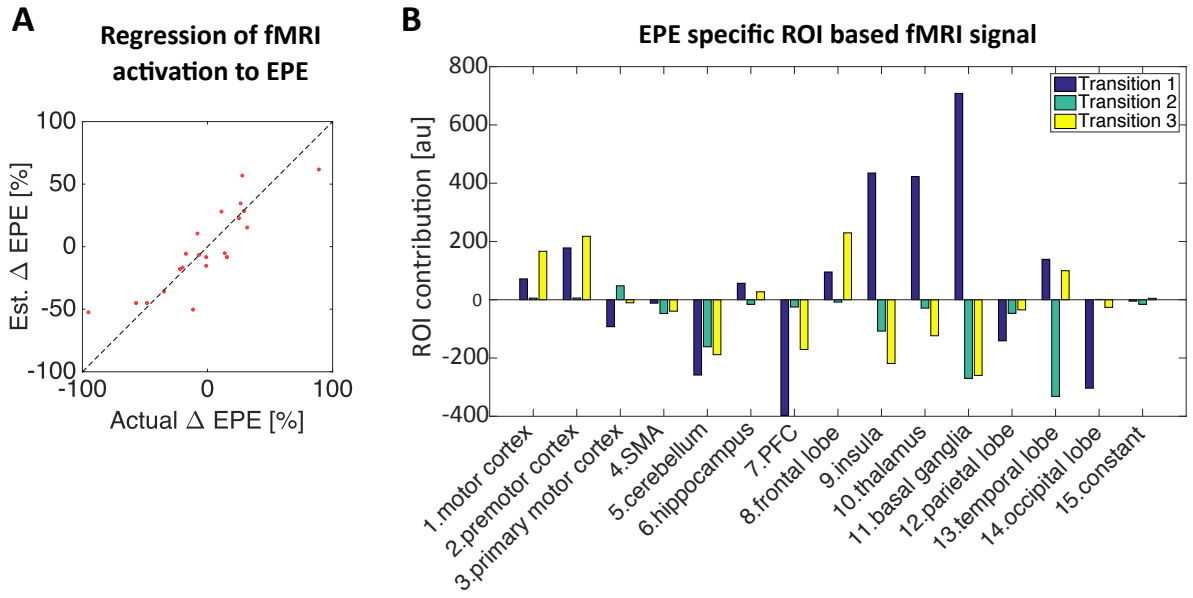

Figure 7.8: fMRI regression against dominant behavioural marker: (**A**) Our regression (Section 7.2.5) yields strong correlation between EPE and the evolution of ROI specific neurophysiological signal (R^2=0.7882, best fit amongst all tested behavioural metrics). (**B**) EPE-related ROI-specific brain signal is illustrated as changes in relative activation between consecutive patient visits ([visit 2]-[visit 1] as 'transition 1', [visit 3]-[visit 2] as 'transition 2', [visit 4]-[visit 3] as 'transition 3'). Interpretation of signal change progression is provided in Section 7.3.3.

In Figure 7.8 B we examine the ROI-specific fMRI signal (ROI-specific relative activation with regard to whole brain activation) acquired from the regression against EPE. This signal is displayed as changes in relative activation between consecutive patient visits ([visit 2]-[visit 1] as 'transition 1', [visit 3]-[visit 2] as 'transition 2', [visit 4]-[visit 3] as 'transition 3'). We observe a systematic decrease of the cerebellum related activation throughout all transitions. From the first to the second visit, which are only 3 weeks apart, this decrease may reflect the impact of a motor learning related process. This finding may at a first sight seem to be contradictory to the

baseline (healthy subjects) increase of relative cerebellar contribution to whole brain activation in late learning stages (Fig. 6.8 B). Yet, we note that here the derived activation reflects an EPE-specific progression of the fMRI signal in patients, which might indeed be moderately linked to cerebellar function in advanced learning stages.

Throughout the following visits (2-4), the decrease of relative cerebellar contribution can be treated as compliant to ataxia's gradual effects on cerebellar function. Interestingly, for the first-to-second-visit transition we also observe a decrease in PFC activation which is seemingly in contrast to the signal changes in Figure 7.7. However, it can be justified if seen only as an EPE-related change in activation. This interpretation would suggest that prefrontal areas' contribution to the optimisation of EPE is completed mainly within the first visit, whereas their contribution to other aspects of motor control continues based on Figure 7.7 A and B until the third visit.

Based on Figure 7.8 B, EPE related activation in the second visit increases in areas such as the insula, thalamus and basal ganglia which are involved in motor learning, in relaying sensory and motor signals to the cerebral cortex and in decision making respectively. This observation may suggest a progression of the motor learning process in ataxia patients, which is governed by the cognitive component of PFC at the first encounter with task conditions and subsequently is more supported by voluntary motor coordination processes and action selection that maps the already formed abstract representations of the task to specific motor control policies. After the second visit, the involvement of such processes in task completion decreases, suggesting a further automatisation of motor execution (potentially supported by the later stage increase of motor cortex activation).

In summary, these results contribute to our investigation of EPE-specific neural correlates of neurodegeneration. They also represent an exemplary method of combining behavioural and neural biomarkers to elucidate the mechanisms underlying the progression of ataxia. This method offers a foundation for further network level analysis (Stam 2014) of neurophysiological signal and regression against the behavioural symptoms of ataxia, which in turn will support disease monitoring and individual patient profiling and management.

7.4 Discussion

In this part of our work, we developed an approach towards gaining insight into the sensorimotor mechanisms underlying the completion of naturalistic complex motor tasks in ataxia patients (FRDA). The approach was founded on the experimental paradigm and analysis tools we developed in Chapter 6 to examine sensorimotor functions of healthy subjects. Crucially, we expanded our analysis tools to encompass a broader range of behavioural metrics in order to describe the

subtle disease progression on the individual patient level. We showed that neurodegeneration -as manifested by a commonly used clinical test (SARA)- was successfully captured by a linear combination of our behavioural measures, amongst which the mean EPE, velocity and bimodality coefficient displayed the strongest contribution. Furthermore, on the implementation level, our analysis revealed neural activity differences between healthy subjects and patients, as well as EPE-specific neural correlates of disease progression, which displayed a systematic reduction of cerebellar contribution.

The behavioural component of our analysis was developed based on a set of high-resolution behavioural metrics, which represented on an individual patient level different aspects of movement patterns and motor performance via EPE, RT, control policies and motor variability. The metrics thereby constitute an objective analysis toolbox for monitoring neurodegeneration, in contrast to conventional clinical ataxia assessments (Fahey et al. 2007; Schmitz-Hübsch et al. 2006; Montcel et al. 2008; Jacobi et al. 2013; Schmitz-Hübsch et al. 2008), which bare an inherent bias due to the partially subjective evaluations of clinicians. At the same time -in order to ensure the clinical relevance of our toolbox- we proposed a technique, which derived the linear combination of our metrics that best describes a common FRDA-specific test (SARA). The combination of these markers poses a strong platform for characterising neurodegeneration at the behavioural level despite ataxia's noisy and subtle progression and its state variability across patients.

We translated our approach to the neural implementation level by developing additional markers featuring the evolution of ROI-specific fMRI signal. This complementary step first revealed a larger contribution of the parietal lobe (linked to multisensory integration) to the sensorimotor functions of healthy subjects (compared to patients) and a larger contribution of the temporal lobe (linked to the formation of visual memories) to the activation patterns of patients (compared to healthy subjects). A further group level analysis of brain activation changes throughout consecutive patient visits, revealed a gradual shift from an enhanced activation of frontal and prefrontal cortical areas to the increased activation of areas supporting the formation and processing of visual information/memories (temporal and occipital lobe) as well as areas involved in multisensory integration which in turn drives motor control (parietal lobe). However, these results cannot be safely interpreted at the group level due to the differences of disease state and progression across patients.

Therefore, in an attempt to optimize our approach to neurodegeneration, we regressed fMRI signal for individual patients against our behavioural measures, in a technique that correlated behavioural changes to a linear combination of ROI-specific activations. The method yielded the strongest correlation results for EPE progression, which plays a pivotal role in our computational perspective of motor control and learning. In particular, the EPE-specific evolution

of fMRI changes throughout the visits revealed a continuously shrinking cerebellar and PFC activity. It also revealed a switch from an initially enhanced contribution to the eventually decreasing contribution of the insula, thalamus and basal ganglia (Fig. 7.8). A potential interpretation of this switch can be that the initial motor-learning-related involvement of regions that support voluntary motor coordination processes and action selection eventually declines due to neurodegeneration which might bring about a further automatisation of task execution.

These results can be further investigated through an expansion of our methodology to host ROI-based processing of fMRI time series and a network level (Stam 2014) formalisation of neural activation. Such an analysis constitutes a promising future direction since it targets the neural correlates of our behavioural metrics which have already been linked to the progression of clinical SARA scores (right panel in Figure 7.4 C). Furthermore, utilising the full information of brain functional networks would offer a more systemic understanding of neuropathology. At the same time, determining these networks into a low-dimensional space in which standardised regression analyses can be employed, would also enable the prediction of disease progression based on the estimation of future biomarker states (Qiu et al. 2015).

In summary, the neurobehavioural modality that we developed for our metrics renders them the potential of fuller disease characterization. This fosters progress in two main directions: (i) the unravelling of sensorimotor mechanisms linked to neurodegeneration and (ii) the biomarker-based profiling of patients. In the latter case, our work serves, as part of a larger ongoing study on FRDA (Gavriel et al. 2015), the purposes of a personalised medicine framework which can advance diagnostics, guide patient management and even unleash new streams of evaluating drug treatment impact.

8 Conclusions and future directions

8.1 Reviewing motivation

Eliciting a unified computational and neurophysiological account of action, perception and learning remains to date an open challenge in neuroscience. There is a rising interest in addressing this challenge, motivated by noteworthy progress in motor neuroscience and BMI technology over the past decade, as well as the re-emergence of AI objectives with a focus on neuro-inspired, embodied approaches to how intelligent systems can learn to interact with and generalise knowledge to unfamiliar task contexts.

Despite the substantial progress in computational sensorimotor control, a key point that remains unclear is *how the brain learns control policies while updating its internal model of the world.* This process is meant to support the completion of complex motor tasks in a world of unknown or partially known dynamics. In fact, a substantial body of studies examines learning only as a gradual adaptation of task related parameters (Berniker and Kording 2008). Other experimental (Nagengast et al. 2009) or purely theoretical studies (Mordatch et al. n.d.) describe how optimal control policies efficiently support complex motor tasks, assuming however full knowledge of task and world dynamics.

Aside from computational concerns, there are also certain methodological considerations, which revisit the experimental conditions employed to investigate motor control and learning mechanisms. Although so far the majority of experimental designs use a small repertoire of laboratory-constrained tasks (e.g. hand reaching movements), studies conclude to generic inferences on sensorimotor principles. How well do these reductionist inferences apply to real-world motor behaviours, which often involve the complex manipulation of objects/tools?

Driven by this debate, in the present work we followed a naturalistic approach to designing experimental conditions for testing motor control and motor learning assumptions in complex motor behaviour. In most of our experimental paradigms we instructed interaction (i) with objects in intuitive life-like tasks (e.g. rotating a bottle-like object in a pouring-water-like fashion) or (ii) objects of unknown dynamics which were intended to elucidate the brain's mechanism of familiarization with genuinely newly encountered (and therefore unintuitive) task dynamics.

8.2 Findings

8.2.1 Action representation

Our first step in unraveling the principles of sensorimotor control and learning in complex motor behaviour, was to examine how the brain represents action in naturalistic object manipulation tasks (Chapter 2). That is how it represents features of the body and environment by mapping experienced sensory states to motor commands. Our findings support the existence of a motor learning model, in which naturalistic object manipulation tasks are represented simultaneously both with regard to a body-based and to an object-based reference frame. These results are in agreement with recent evidence on the use of mixed extrinsic and intrinsic variables to support simple hand-reaching movements (Brayanov et al. 2012; Berniker et al. 2014; Diedrichsen 2007). They thereby expand the validity of the composite action representations narrative to the naturalistic complex motor task context.

8.2.2 Motor learning

Building on our action representation findings, we proceeded to an inclusive formalisation of a motor learning hypothesis for complex motor behaviour. The latter predicts that in unfamiliar task conditions the brain makes continuous decisions for the generation of complex trajectories by learning optimal feedback control based on the steady identification of unknown parameters of both body and object dynamics. The process employs error-based learning that exploits the brain's inherent link between forward models and feedback control to compute dynamically updated policies. We presented experiments, which validated that our learning model successfully predicts the transient dynamics of human motor learning in an unintuitive object manipulation task on a trial-by-trial basis.

Importantly, our approach expands previous work on adaptation of task related parameters (Berniker and Kording 2008) by merging it with a newly proposed mechanism of policy learning (PLM). It also outperforms a common ideal-actor-model (IAM), which makes use of OFC and assumes full knowledge of the world and task dynamics.

PLM's ground concept that the brain uses simple error-based gradient-descent learning rules to reach near optimal control decisions in complex motor tasks provokes interesting neurobiological parallels, since this mechanism can be arguably implemented by the Hebbian-like modification of synaptic weights.

An interesting expansion of our approach to motor learning could account for the appreciable amounts of noise encountered in the streams of sensory processing and motor command for-

mation. Such an approach would be expected to more realistically capture the variability in human performance measured in our paradigm throughout the course of learning. It would also be expected to be consistent with a Bayesian inference framework in which an internal model of the task and world given some sensory data is updated in an optimal fashion based on the likelihood of sensory inputs given their causes and prior knowledge on those causes (Körding and Wolpert 2004).

Another exciting future direction of the present work is to merge our version of policy learning at the level of system identification and parameter optimisation with a symbolic action selection process that will determine the structure of controllers and body/world representations. Such an extension will thus account for the emergence of the internal forward model postulated in our framework. Recent theoretical work on hierarchical motor control proposes a model that makes use of a high-level reinforcement learner, which searches and forms an optimal combination of low-level locally linear controllers and their associated quadratic costs (as actions and rewards respectively) in order to solve non-linear motor control problems (Abramova et al. 2011). This algorithm poses a promising candidate for bridging our low-level motor learning narrative to high-level decision making.

Along the lines of this outlook we developed a methodological platform to track and examine exploratory behaviour during the motor learning process, which remains an obscure topic in neuroscience (Section 4.1, Cohen et al. 2007). Our approach employed computational analysis to reverse engineer and subsequently assess the control process in a whole-body manipulation paradigm. We showed that during learning the brain decides upon transitions between different motor outputs. This enhances our view of motor learning beyond the gradual improvement of one selected strategy, as we modelled it and investigated it in Chapter 3. It namely expands it to the context of selecting new motor control policies while learning, which in turn lead to evidently different behavioural outputs, that still achieve task completion. Our study also revealed an unstable regime of motor control within the completion of single trials at the end-point of learning. The degree and occurrence of this instability appear to be related to the amount and form of the trajectory steps produced until goal achievement. Future work can employ this approach to empirically validate or inform any modeling attempts to fusing low-level motor learning with high-level action selection and exploration.

8.2.3 Technology and methodology to extract neurobehavioural correlates for healthy and pathological sensorimotor processes

Crucially, our work complemented computational and psychophysical frameworks with a neuroimaging investigation of our questions. Our motivation arose mainly from the current absence

of a solid association of psychophysically validated computational models of sensorimotor control and learning to their underlying neural foundation (Wolpert et al. 2011). This association places our hypotheses on sensorimotor processes in a context of physiological constraints and can thereby inform and update any proposed theoretical model.

We therefore built and tested an fMRI-compatible haptic object manipulation system to achieve real-time association of motor psychophysics to neurophysiology (Chapter 5). The system bares technical benefits with a high-frequency of data acquisition (120 Hz) for a 320 x 240 pixel resolution with low time delays (<20 ms) and can track 6DOF movement. It thereby allows closed-loop interaction with physical objects of potentially variable shapes and dynamics and provides online feedback of motor performance to subjects to facilitate naturalistic task conditions for the study of motor learning. f2MOVE poses an efficient low-cost neurotechnological alternative to established fMRI-incompatible motion tracking systems and to fMRI-compatible robotic manipulanda which enable motion within the scanner, but commonly do not facilitate life-like tasks.

We adjusted f2MOVE in a clinical setting (Clinical Imaging Facility, Imperial College London, Hammersmith Hospital, London) and employed it for an object manipulation paradigm on healthy human subjects (Chapter 5). The instructed task was translation invariant and resembled naturalistic rotational activities (e.g. turning a key, operating a door knob). We showed that a set of behavioural metrics inspired by our previous empirical observations (Chapters 2 and 3) -including end-point-error (EPE), reaction time (RT) and movement variability- capture performance progression and can therefore be treated as measures of motor learning. In particular, EPE in light of its role as driving force of learning in PLM (Chapter 3), was further investigated as a behavioural regressor against fMRI signal. Our findings provide insight into the neural correlates of motor learning dynamics and particularly elucidate the neural foundation of fast, early-stage and plateau-level, late-stage EPE learning; with the former primarily governed by frontal and prefrontal cortical areas, motor cortical areas and parietal lobe and the latter essentially supported by cerebellar contribution. A possible interpretation of this shift is the transition from early motor planning and the formation of task representations to an eventual automatisation of the motor learning process.

The same object manipulation paradigm and methodological platform were employed on Friedreich's ataxia patients (FRDA) to gain insight into the longitudinal neurodegeneration-specific changes of sensorimotor functions. We showed that by expanding our set of behavioural markers and employing a distinct linear combination thereof, we were able to capture the noisy and subtle behavioural changes that characterised the progression of the disease on an individual patient level. Amongst the combined measures, mean EPE, velocity and bimodality coefficient appear to be the ones with the strongest correlation to a clinically established FRDA test (SARA). They therefore pose objective diagnostic tools that can complement conventional

inherently subjective clinical evaluations.

We further tracked the neural underpinnings of disease progression throughout clinical visits. The overall comparison of brain activation between patients across all visits and our baseline healthy subject group yielded an enhanced contribution of the parietal lobe (linked to multisensory integration) to the sensorimotor functions of healthy subjects and an enhanced contribution of the temporal lobe (linked to the formation of visual memories) to the activation patterns of patients. However, group level analysis does not support safe inferences, due to the variability across patients' disease states and progression rates. Thus, our main attempt to examine neurodegeneration was to regress for all individual patients a ROI-specific combination of fMRI signal changes throughout visits to each of our behavioural measures. The method yielded the strongest correlation results for EPE progression, thereby highlighting EPE as a dominant behavioural signature of neurodegeneration. Notably, the EPE-specific evolution of fMRI changes throughout visits revealed a continuously shrinking cerebellar and PFC activity and a transition from an initially enhanced contribution to the eventually decreasing contribution of the insula, thalamus and basal ganglia (Fig. 7.8). We argue that this transition could be explained by the initial motor-learning-related involvement of regions supporting voluntary motor coordination processes and action selection which eventually declines due to neurodegeneration that might cause a further automatisation of task execution.

In summary, the extracted neurobehavioural metrics constitute a powerful biomarker toolbox to illuminate not only sensorimotor principles but also the perturbation thereof during the course of neurodegeneration. They thus encourage and can guide future investigations on their evolution time-scales and on their systemic (network-level) organisation.

8.3 Applications and future directions

8.3.1 Artificial systems and brain-machine interfaces

Progress in sensorimotor control and learning research becomes increasingly relevant to successful implementations in robotics and artificial systems in general. Nowadays AI has directed its objectives towards algorithms that allow artificial agents to learn how to excel at a diverse array of challenging and potentially newly encountered tasks (e.g. Mnih et al. 2015). These algorithms are increasingly founded on psychological (Thorndike 1911) and neuroscientific (Schultz 1997) approaches to behaviour, which allows an agent to learn the control of an environment. Consequently, our human-level validated motor learning formalisation (PLM, Chapter 3) offers a promising basis to expand such work.

Furthermore, exciting prospects of application suggest themselves in brain machine interfaces

(BMI). BMI research is namely a fast growing field, but a series of important challenges remain to be tackled to allow a significant impact on patients (Vaadia and Birbaumer 2009). A major one amongst those is the improvement of decoding algorithms. Contemporary BMIs mainly support an advanced yet arbitrary mapping between brain activation and motor control signals. This limits their effectiveness to the control of a distinct range of simple motor tasks (e.g. cursor-based or robotic-arm-based reaching movements). The facilitation of more sophisticated complex tasks (e.g. object manipulation) appears to require enhanced decoding principles with an embedded systemic description of the brain's sensorimotor function.

PLM meets this need in capturing sensorimotor control and learning based on an internal forward model, which describes the brain as a memory based prediction machine that relates actions to consequences and via experience updates its stored representations of the task and the world (Chapters 2 and 3). A meaningful outlook of the present work shall therefore firstly aim at a further f2MOVE-based refinement of our motor learning framework (Chapters 5 and 6) which will introduce all necessary neurophysiological constraints in PLM's computational perspective. Subsequent steps shall proceed to the integration and testing of this approach in BMI setting.

This research direction will be evaluated by the efficiency in which brain-driven artificial devices will produce natural-like movements adapted to the constraints of human motor learning and performance. Successful results can enrich clinical applications through neuroprosthetics and improve rehabilitation for people with disabilities. They can also potentially promote BMIs as a major communication and control technology for general population use in a broad range of daily activities; professional and otherwise.

8.3.2 Clinical implications and personalised medicine

On another front, the findings and methods from our clinically tested paradigm (Chapter 6) address a wide range of diagnostic and patient monitoring applications and suggest a range of aspiring challenges ahead.

Our biomarker toolbox can be flexibly expanded to host further ROI-based processing and network level analysis of fMRI time series (Stam 2014). Such analysis can refine our existing method of extracting neural correlates of distinct clinically relevant behavioural measures (Fig. 7.8). Exploiting the full information of brain functional networks would offer a more systemic insight into the neural mechanisms underlying sensorimotor control and learning and describe our computational framework (PLM, Chapter 3) at the implementation level.

Furthermore, such an expansion will allow a deeper understanding of the neural underpinnings of disease progression when our experimental paradigm is translated to patients with a neurodegenerative condition (e.g. FRDA). In fact, previous studies have shown that by projecting

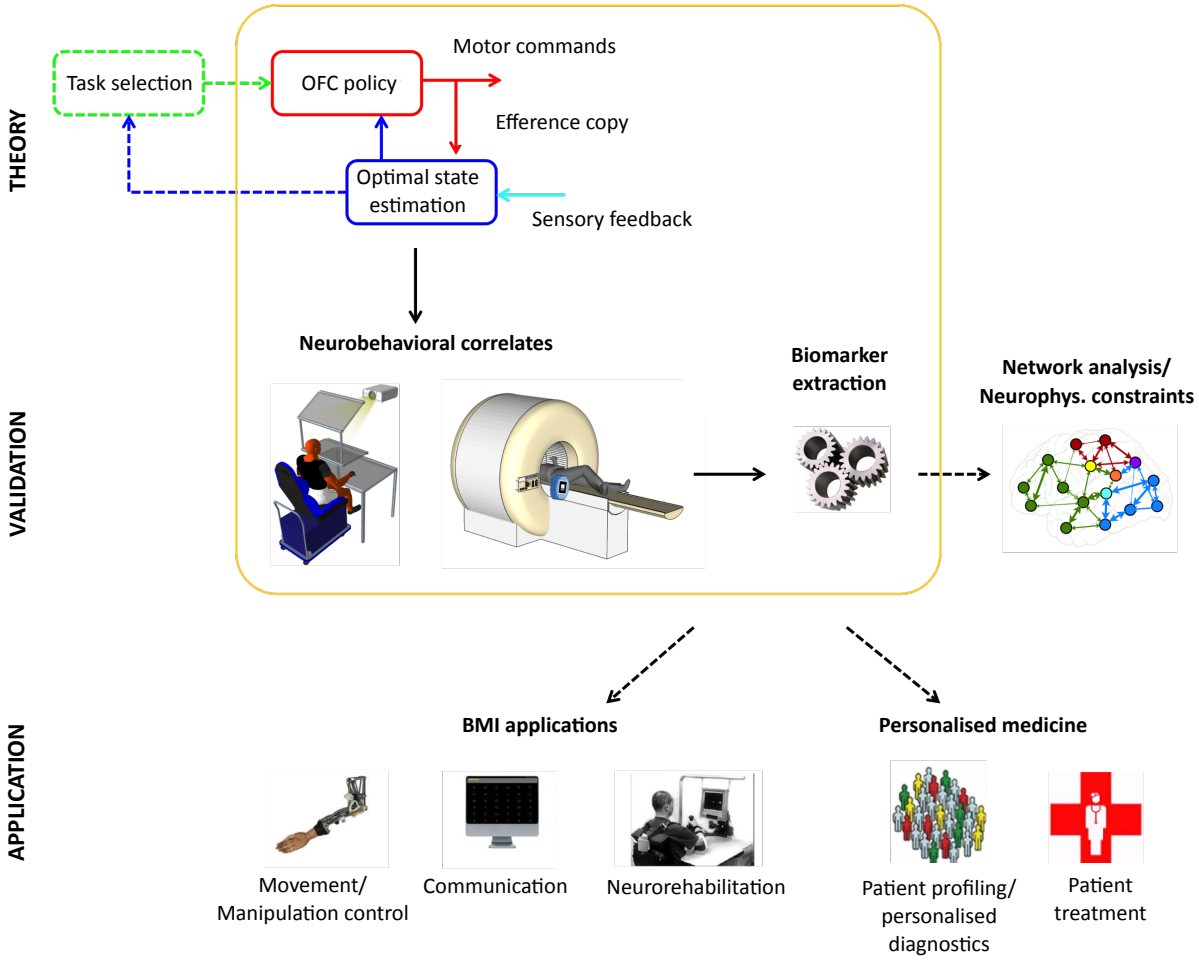

Figure 8.1: Findings, outlook and applications: We developed a framework (PLM) that captures motor learning in complex object manipulation as a process of simultaneous control policy updates and the identification of body and object dynamics. Our assumptions on the learning mechanism and the form of world representations were validated in behavioural paradigms (Chapters 2 and 3). We further developed technology to transfer these closed-loop paradigms into neuroimaging sessions and methodology to set the foundation for eliciting the neural correlates of our computational approach to sensorimotor control and learning (Chapters 5 and 6). In the context of two clinical studies (on healthy subjects and ataxia patients) this methodology supported a first attempt to extract biomarkers which characterize and track the motor performance on a behavioural and implementation level (Chapters 6 and 7). All findings are included in yellow frame. Future research can expand our work in multiple promising directions: e.g. embed into PLM a high-level symbolic action selection mechanism (Abramova et al. 2011) that accounts for cognitive planning, enrich our neurobehavioural metrics with the help of network analysis on the acquired fMRI signal etc. Applications therefore include BMI technology and the fields of personalised patient diagnostics, profiling and treatment.

brain activation signal to a low-dimensional space via network analysis, one can subsequently subject it to standardised regression methods against the behavioural metrics of our platform (Qiu et al. 2015). This process can be employed to predict disease progression based on the estimation of future biomarker states. In conclusion, these steps pave the path towards fuller disease characterisation and patient profiling; goals currently pursued within the context of a broader ongoing study on FRDA (Gavriel et al. 2015). The envisioned personalised medicine framework will advance diagnostics, disease monitoring and patient management with possible implications in drug research through the optimised evaluation of pharmaceutical treatment.

Bibliography

Abramova, E., Faisal, A., and Kuhn, D. 2011. "Combining Markov Decision Processes with Linear Optimal Controllers." *Proc. Imp. Coll. Comput. Sci. Work.* 3–9. (Cited on pages 60, 78, 96, 154, 158).

Acun, D. E. 2010. "Structure Learning in Human Sequential." *Learning* 6 (12): e1001003. (Cited on page 78).

Acuña, D. E. and Schrater, P. 2010. "Structure learning in human sequential decision-making." *PLoS Comput. Biol.* 6 (12). (Cited on page 53).

Ahmed, A. A., Wolpert, D. M., and Flanagan, J. R. 2008a. "Flexible Representations of Dynamics Are Used in Object Manipulation." *Curr. Biol.* 18 (10): 763–768. (Cited on pages 21, 29, 50).

Ahmed, A. A., Wolpert, D. M., and Flanagan, J. R. 2008b. "Flexible representations of dynamics are used in object manipulation." *Curr. Biol.* 18 (10): 763–8. (Cited on page 54).

Albus, J. S. 1971. "A theory of cerebellar function." *Math. Biosci.* 10 (1): 25–61. (Cited on pages 19, 20, 107, 137).

Atkeson, C. G. and Hollerbach, J. M. 1985. "Kinematic features of unrestrained vertical arm movements." *J. Neurosci.* 5 (9): 2318–30. (Cited on page 30).

Bandettini, P. A., Jesmanowicz, A., Wong, E. C., and Hyde, J. S. 1993. "Processing strategies for time-course data sets in functional MRI of the human brain." *Magn. Reson. Med.* 30 (2): 161–173. (Cited on page 114).

Beckmann, C. F., Jenkinson, M., and Smith, S. M. 2003. "General multilevel linear modeling for group analysis in FMRI." *Neuroimage* 20 (2): 1052–1063. (Cited on page 116).

Bellman, R. 1957. *Dynamic Programming.* 1st ed. Princeton, NJ, USA: Princeton University Press. (Cited on page 53).

Berkes, P., Orbán, G., Lengyel, M., and Fiser, J. 2011. "Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment." *Science* 331 (6013): 83–7. (Cited on page 18).

Berniker, M., Franklin, D. W., Flanagan, J. R., Wolpert, D. M., and Kording, K. P. 2014. "Motor learning of novel dynamics is not represented in a single global coordinate system: evaluation of mixed coordinate representations and local learning." *J. Neurophysiol.* 111 (6): 1165–82. (Cited on pages 21, 29, 50, 51, 54, 153).

Berniker, M. and Kording, K. 2008. "Estimating the sources of motor errors for adaptation and generalization." *Nat. Neurosci.* 11 (12): 1454–1461. (Cited on pages 19, 23, 50, 53, 75, 152, 153).

Berry, D. A. and Fristedt, B. 1985. *Bandit problems.* Dordrecht: Springer Netherlands. (Cited on page 53).

Bertsekas, D. P. 2001. *Dynamic Programming and Optimal Control, Vol. I, 2nd Ed.* Belmont, MA: Athena Scientific. (Cited on page 53).

Bertsekas, D. P. and Tsitsiklis, J. N. 1995. "Neuro-dynamic programming: an overview." In *Decis. Control. 1995., Proc. 34th IEEE Conf.* 1:560–564. IEEE. (Cited on page 53).

Bhat, H. S. and Kumar, N. 2010. "On the derivation of the Bayesian Information Criterion." *Sch. Nat. Sci. Univ. Calif.* (Cited on page 134).

Binkofski, F., Buccino, G., Posse, S., Seitz, R. J., Rizzolatti, G., and Freund, H.-J. 1999. "A fronto-parietal circuit for object manipulation in man: evidence from an fMRI-study." *Eur. J. Neurosci.* 11 (9): 3276–3286. (Cited on page 108).

Bock, O. 1990. "Load compensation in human goal-directed arm movements." *Behav. Brain Res.* 41 (3): 167–177. (Cited on page 30).

Bonilha da Silva, C., Bergo, F. P. G., D'Abreu, A., Cendes, F., Lopes-Cendes, I., and França, M. C. 2014. "Dentate nuclei T2 relaxometry is a reliable neuroimaging marker in Friedreich's ataxia." *Eur. J. Neurol.* 21 (8): 1131–6. (Cited on page 131).

Brammer, M. J. 2001. "Head motion and its correction." *Funct. MRI An Introd. to methods:* 243–250. (Cited on page 114).

Braun, D. A., Aertsen, A., Wolpert, D. M., and Mehring, C. 2009a. "Learning optimal adaptation strategies in unpredictable motor tasks." *J. Neurosci.* 29 (20): 6472–6478. (Cited on pages 53, 77).

Braun, D. A., Aertsen, A., Wolpert, D. M., and Mehring, C. 2009b. "Motor Task Variation Induces Structural Learning." *Curr. Biol.* 19 (4): 352–357. (Cited on pages 53, 78).

Braun, D. A., Mehring, C., and Wolpert, D. M. 2010. *Structure learning in action.* (Cited on pages 53, 54, 78).

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.* (Cited on pages 21, 29, 51, 54, 153).

Brown, G. 1986. "Case, R. (1985). Intellectual Development: Birth to Adulthood." *Br. J. Educ. Psychol.* 56 (2): 220–222. (Cited on pages 30, 108).

Bülthoff, H. 1996a. "A Bayesian Framework for the Integration of Visual Modules." In *Atten. Perform. XVI Inf. Integr. Percept. Commun.* 49–70. (Cited on page 53).

Bülthoff, H. 1996b. "Bayesian decision theory and psychophysics." *Percept. as Bayesian inference:* 123. (Cited on page 53).

Burgess, J. K., Bareither, R., and Patton, J. L. 2007. "Single limb performance following contralateral bimanual limb training." [in English]. *IEEE Trans. Neural Syst. Rehabil. Eng.* 15 (3): 347–55. (Cited on pages 21, 29, 32, 50).

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." *J. Neurosci.* 24 (40): 8662–71. (Cited on page 30).

Chen-Harris, H., Joiner, W. M., Ethier, V., Zee, D. S., and Shadmehr, R. 2008. "Adaptive control of saccades via internal feedback." *J. Neurosci.* 28 (11): 2804–2813. (Cited on pages 53, 120, 129).

Cohen, J. D., McClure, S. M., and Yu, A. J. 2007. "Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration." *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 362 (1481): 933–42. (Cited on pages 81, 154).

Coogler, C. E. 1992. "Falls and imbalance." *Rehab Manag.* 53. (Cited on page 95).

Cooper, R. P. 2010. *Forward and inverse models in motor control and cognitive control.* (Cited on page 80).

Cothros, N., Köhler, S., Dickie, E. W., Mirsattari, S. M., and Gribble, P. L. 2006. "Proactive interference as a result of persisting neural representations of previously learned motor skills in primary motor cortex." *J. Cogn. Neurosci.* 18 (12): 2167–76. (Cited on page 54).

Criscimagna-Hemminger, S. E., Donchin, O., Gazzaniga, M. S., and Shadmehr, R. 2003. "Learned dynamics of reaching movements generalize from dominant to nondominant arm." *J. Neurophysiol.* 89 (1): 168–76. (Cited on pages 21, 29).

Dam, G., Kording, K., and Wei, K. 2013. "Credit assignment during movement reinforcement learning." *PLoS One* 8 (2): e55352. (Cited on page 50).

D'Avella, A. and Bizzi, E. 1998. "Low dimensionality of supraspinally induced force fields." *Proc. Natl. Acad. Sci. U. S. A.* 95 (13): 7711–7714. (Cited on page 76).

Dayan, P. and Daw, N. D. 2008. "Decision theory, reinforcement learning, and the brain." *Cogn. Affect. Behav. Neurosci.* 8 (4): 429–453. (Cited on page 53).

Deiber, M.-P., Wise, S. P., Honda, M., Catalan, M. J., Grafman, J., and Hallett, M. 1997. "Frontal and parietal networks for conditional motor learning: a positron emission tomography study." *J. Neurophysiol.* 78 (2): 977–991. (Cited on page 108).

Diedrichsen, J. 2007. "Optimal Task-Dependent Changes of Bimanual Feedback Control and Adaptation." *Curr. Biol.* 17 (19): 1675–1679. (Cited on pages 21, 22, 28, 29, 40, 49, 51, 54, 153).

Diedrichsen, J., Hashambhoy, Y., Rane, T., and Shadmehr, R. 2005. "Neural correlates of reach errors." *J. Neurosci.* 25 (43): 9919–9931. (Cited on pages 22, 105, 108).

Diedrichsen, J., Shadmehr, R., and Ivry, R. B. 2010. "The coordination of movement: optimal feedback control and beyond." *Trends Cogn. Sci.* 14 (1): 31–9. (Cited on pages 21, 28, 40, 49).

Ding, Q., Stevenson, I. H., Wang, N., Li, W., Sun, Y., Wang, Q., Kording, K., and Wei, K. 2013. "Motion games improve balance control in stroke survivors: A preliminary study based on the principle of constraint-induced movement therapy." *Displays* 34 (2): 125–131. (Cited on page 96).

Donaldson, D. I. and Buckner, R. L. 2001. "Effective paradigm design." In *P. JEZZARD (ED.), Funct. MRI.* Citeseer. (Cited on page 114).

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." *J. Neurosci.* 23 (27): 9032–45. (Cited on pages 22, 29, 32, 50).

Dürr, A., Cossee, M., Agid, Y., Campuzano, V., Mignard, C., Penet, C., Mandel, J. L., Brice, A., and Koenig, M. 1996. "Clinical and genetic abnormalities in patients with Friedreich's ataxia." *N. Engl. J. Med.* 335 (16): 1169–75. (Cited on pages 25, 130).

Fahey, M. C., Corben, L., Collins, V., Churchyard, A. J., and Delatycki, M. B. 2007. "How is disease progress in Friedreich's ataxia best measured? A study of four rating scales." *J. Neurol. Neurosurg. Psychiatry* 78 (4): 411–3. (Cited on pages 131, 150).

Faisal, A. A., Selen, L. P. J., and Wolpert, D. M. 2008. "Noise in the nervous system." *Nat. Rev. Neurosci.* 9 (4): 292–303. (Cited on pages 19, 52, 98).

Faisal, A. A. and Wolpert, D. M. 2009. "Near optimal combination of sensory and motor uncertainty in time during a naturalistic perception-action task." *J. Neurophysiol.* 101 (4): 1901–1912. (Cited on pages 19, 52).

Fillyaw, M. J., Badger, G. J., Bradley, W. G., Tandan, R., Blair, C. J., Fries, T. J., Wilder, D. G., Boerman, J., Young, J., and Witarsa, M. 1989. "Quantitative measures of neurological function in chronic neuromuscular diseases and ataxia." *J. Neurol. Sci.* 92 (1): 17–36. (Cited on page 131).

Flanagan, J. R. and Beltzner, M. A. 2000. "Independence of perceptual and sensorimotor predictions in the size-weight illusion." *Nat. Neurosci.* 3 (7): 737–41. (Cited on page 30).

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." *J. Neurosci.* 17 (4): 1519–1528. (Cited on pages 22, 81).

Flourens, P. 1842. *Recherches exp{é}rimentales sur les propri{é}t{é}s et les fonctions du syst{è}me nerveux dans les animaux vert{é}br{é}s.* Balli{è}re. (Cited on page 107).

Friston, K. J., Frith, C. D., Frackowiak, R. S., and Turner, R. 1995. "Characterizing dynamic brain responses with fMRI: a multivariate approach." *Neuroimage* 2 (2): 166–72. (Cited on page 114).

Friston, K. J., Jezzard, P., and Turner, R. 1994. "Analysis of functional MRI time-series." *Hum. Brain Mapp.* 1 (2): 153–171. (Cited on pages 114, 115).

Friston, K. 2010. "The free-energy principle: a unified brain theory?" *Nat. Rev. Neurosci.* 11 (2): 127–138. (Cited on page 98).

Friston, K. J. 2011. "Functional and effective connectivity: a review." *Brain Connect.* 1 (1): 13–36. (Cited on page 129).

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., and Frackowiak, R. S. J. 1995. "Spatial registration and normalization of images." *Hum. Brain Mapp.* 3 (3): 165–189. (Cited on page 114).

Friston, K. J., Daunizeau, J., and Kiebel, S. J. 2009. "Reinforcement learning or active inference?" *PLoS One* 4 (7): e6421. (Cited on page 98).

Friston, K., Mattout, J., and Kilner, J. 2011. "Action understanding and active inference." *Biol. Cybern.* 104 (1-2): 137–60. (Cited on page 21).

Galea, J. M., Vazquez, A., Pasricha, N., Xivry, J.-J. O. de, and Celnik, P. 2011. "Dissociating the roles of the cerebellum and motor cortex during adaptive learning: the motor cortex retains what the cerebellum learns." *Cereb. Cortex* 21 (8): 1761–70. (Cited on page 19).

Gandolfo, F., Mussa-Ivaldi, F. A., and Bizzi, E. 1996. "Motor learning by field approximation." *Proc. Natl. Acad. Sci. U. S. A.* 93 (9): 3843–6. (Cited on pages 29, 30, 32, 50).

Gavriel, C. and Faisal, A. A. 2013. "Wireless kinematic body sensor network for low-cost neurotechnology applications in-the-wild." In *Neural Eng. (NER), 2013 6th Int. IEEE/EMBS Conf.* 1279–1282. IEEE. (Cited on page 105).

Gavriel, C., Thomik, A., Lourenco, P., Nageshwaran, S., Athanasopoulos, S., Sylaidi, A., Festenstein, R., and Faisal, A. A. 2015. "Towards neurobehavioral biomarkers for longitudinal monitoring of neurodegeneration with wearable body sensor networks." In *Neural Eng. (NER), 2015 7th Int. IEEE/EMBS Conf.* IEEE. (Cited on pages 27, 132, 151, 159).

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. 2014. *Bayesian data analysis.* Vol. 2. Taylor & Francis. (Cited on page 32).

Genovese, C. R., Lazar, N. A., and Nichols, T. 2002. "Thresholding of statistical maps in functional neuroimaging using the false discovery rate." *Neuroimage* 15 (4): 870–878. (Cited on page 115).

Ghahramani, Z. and Hinton, G. E. 1996. *Parameter Estimation for Linear Dynamical Systems.* Technical report. (Cited on pages 64, 70).

Ghahramani, Z., Wolpert, D. M., and Jordan, M. I. 1996. "Generalization to Local Remappings of the Visuomotor Coordinate Transformation." *J. Neurosci.* 16 (21): 7085–7096. (Cited on pages 21, 28).

Gibilisco, P. and Vogel, A. P. 2013. "Friedreich ataxia." *BMJ* 347:f7062. (Cited on pages 25, 130).

Gillespie, L. D., Gillespie, W. J., Robertson, M. C., Lamb, S. E., Cumming, R. G., and Rowe, B. H. 2003. "Interventions for preventing falls in elderly people." *Cochrane database Syst. Rev.* No. 4: CD000340. (Cited on page 95).

Gittins, J., Glazebrook, K., and Weber, R. 2011. *Multi-armed bandit allocation indices.* John Wiley & Sons. (Cited on page 53).

Glimcher, P. W. 2004. *Decisions, uncertainty, and the brain: The science of neuroeconomics.* MIT press. (Cited on page 53).

Globerson, E. and Nelken, I. 2013. "The neuro-pianist." *Front. Syst. Neurosci.* 7:35. (Cited on page 18).

Goble, D. J., Cone, B. L., and Fling, B. W. 2014. "Using the Wii Fit as a tool for balance assessment and neurorehabilitation: the first half decade of "Wii-search"." *J. Neuroeng. Rehabil.* 11 (1): 12. (Cited on page 96).

Gold, J. I. and Shadlen, M. N. 2007. "The neural basis of decision making." [in en]. *Annu. Rev. Neurosci.* 30:535–74. (Cited on pages 21, 53).

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." *J. Neurophysiol.* 69 (6): 1789–96. (Cited on page 30).

Gotovac, K., Hajnšek, S., Pašić, M. B., Pivac, N., and Borovečki, F. 2014. "Personalized medicine in neurodegenerative diseases: how far away?" *Mol. Diagn. Ther.* 18 (1): 17–24. (Cited on pages 25, 130, 131).

Grafton, S., Woods, R. P., Tyszka, M., et al. 1994. "Functional imaging of procedural motor learning: relating cerebral blood flow with individual subject performance." *Hum. Brain Mapp.* 1 (3): 221–234. (Cited on page 108).

Graydon, F. X., Friston, K. J., Thomas, C. G., Brooks, V. B., and Menon, R. S. 2005. "Learning-related fMRI activation associated with a rotational visuo-motor transformation." *Cogn. brain Res.* 22 (3): 373–383. (Cited on page 108).

Green, D. M. and Swets, J. A. 1966. *Signal Detection Theory and Psychophysics.* New York: Wiley. (Cited on page 53).

Hajnal, J. V., Myers, R., Oatridge, A., Schwieso, J. E., Young, I. R., and Bydder, G. M. 1994. "Artifacts due to stimulus correlated motion in functional imaging of the brain." *Magn. Reson. Med.* 31 (3): 283–91. (Cited on page 113).

Haruno, M., Wolpert, D. M., and Kawato, M. 2001. "Mosaic model for sensorimotor learning and control." *Neural Comput.* 13 (10): 2201–20. (Cited on page 28).

Hatsopoulos, N. G. and Donoghue, J. P. 2009. "The science of neural interface systems." *Annu. Rev. Neurosci.* 32:249–66. (Cited on page 25).

Helmholtz, H. von. 1866. *Handbuch der physiologischen Optik: mit 213 in den Text eingedruckten Holzschnitten und 11 Tafeln.* 874. Voss. (Cited on page 18).

Hogan, N. 1984. "An organizing principle for a class of voluntary movements." *J. Neurosci.* 4 (11): 2745–2754. (Cited on page 75).

Holmes, A. P. and Friston, K. J. 1998. "Generalisability, Random E ects & Population Inference." *Neuroimage* 7:S754. (Cited on pages 116, 136).

Horak, F. B. 2006. "Postural orientation and equilibrium: what do we need to know about neural control of balance to prevent falls?" *Age Ageing* 35 Suppl 2:ii7–ii11. (Cited on page 96).

Houk, J. C., Buckingham, J. T., and Barto, A. G. 1996. "Models of the cerebellum and motor learning." *Behav. Brain Sci.* 19 (3): 368–383. (Cited on pages 19, 107).

Howard, I. S., Ingram, J. N., Körding, K. P., and Wolpert, D. M. 2009. "Statistics of natural movements are reflected in motor errors." *J. Neurophysiol.* 102 (3): 1902–1910. (Cited on page 99).

Howard, I. S., Ingram, J. N., and Wolpert, D. M. 2008. "Composition and decomposition in bimanual dynamic learning." *J. Neurosci.* 28 (42): 10531–40. (Cited on page 30).

Howard, I. S., Ingram, J. N., and Wolpert, D. M. 2010. "Context-dependent partitioning of motor learning in bimanual movements." *J. Neurophysiol.* 104 (4): 2082–91. (Cited on page 30).

Huettel Scott A., S. A. W. M. G. 2004. *Functional magnetic resonance imaging.* Sinauer Associates. (Cited on pages 113, 114).

Ingram, J. N., Howard, I. S., Flanagan, J. R., and Wolpert, D. M. 2010. "Multiple grasp-specific representations of tool dynamics mediate skillful manipulation." *Curr. Biol.* 20 (7): 618–23. (Cited on pages 21, 28, 29, 40, 49, 50).

Ingram, J. N., Körding, K. P., Howard, I. S., and Wolpert, D. M. 2008. "The statistics of natural hand movements." *Exp. Brain Res.* 188 (2): 223–236. (Cited on page 99).

Ingram, J. N. and Wolpert, D. M. 2011. "Naturalistic approaches to sensorimotor control." *Prog. Brain Res.* 191:3–29. (Cited on pages 21, 24, 30, 49, 105, 108, 128).

Izawa, J., Rane, T., Donchin, O., and Shadmehr, R. 2008. "Motor adaptation as a process of reoptimization." *J. Neurosci.* 28 (11): 2883–91. (Cited on pages 19, 52).

Izawa, J. and Shadmehr, R. 2011. "Learning from sensory and reward prediction errors during motor adaptation." *PLoS Comput. Biol.* 7 (3): e1002012. (Cited on page 53).

Jacobi, H. et al. 2013. "Inventory of Non-Ataxia Signs (INAS): validation of a new clinical assessment instrument." *Cerebellum* 12 (3): 418–28. (Cited on pages 131, 150).

Jagannathan, S. and Galan, G. 2004. "Adaptive critic neural network-based object grasping control using a three-finger gripper." [in English]. *IEEE Trans. Neural Netw.* 15 (2): 395–407. (Cited on page 30).

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S., and Passingham, R. E. 1994. "Motor sequence learning: a study with positron emission tomography." *J. Neurosci.* 14 (6): 3775–3790. (Cited on page 108).

Jenkinson, M. and Smith, S. 2001. "A global optimisation method for robust affine registration of brain images." *Med. Image Anal.* 5 (2): 143–56. (Cited on page 116).

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. 2002. "Improved optimization for the robust and accurate linear registration and motion correction of brain images." *Neuroimage* 17 (2): 825–841. (Cited on page 114).

Johansson, R. S. and Westling, G. 1988. "Coordinated isometric muscle commands adequately and erroneously programmed for the weight during lifting task with precision grip." *Exp. brain Res.* 71 (1): 59–71. (Cited on page 30).

Jordan, M. I. and Rumelhart, D. E. 1992. "Forward models: Supervised learning with a distal teacher." *Cogn. Sci.* 16 (3): 307–354. (Cited on pages 20, 107, 137).

Kadiallah, A., Franklin, D. W., and Burdet, E. 2012. "Generalization in adaptation to stable and unstable dynamics." *PLoS One* 7 (10): e45075. (Cited on page 29).

Kaelbling, L. P., Littman, M. L., and Moore, A. W. 1996. "Reinforcement learning: a survey." *J. Artif. Intell. Res.* 4 (1): 237–285. (Cited on page 81).

Kandel, E. R., Schwartz, J. H., Jessell, T. M., et al. 2000. *Principles of neural science.* Vol. 4. McGraw-Hill New York. (Cited on page 19).

Kant, I. 1781. "Kritik der Reinen Vernunft." (Cited on page 19).

Kato, H. and Billinghurst, M. 1999. "Marker tracking and hmd calibration for a video-based augmented reality conferencing system." In *Augment. Reality, 1999.(IWAR'99) Proceedings. 2nd IEEE ACM Int. Work.* 85–94. IEEE. (Cited on page 102).

Kawato, M. and Wolpert, D. 1998. "Internal models for motor control." *Novartis Found. Symp.* 218:291–304, discussion 304–7. (Cited on pages 19, 80, 81).

Kawato, M., Furukawa, K., and Suzuki, R. 1987. "A hierarchical neural-network model for control and learning of voluntary movement." *Biol. Cybern.* 57 (3): 169–185. (Cited on pages 20, 107, 137).

Kawato, M., Kuroda, T., Imamizu, H., Nakano, E., Miyauchi, S., and Yoshioka, T. 2003. "Internal forward models in the cerebellum: fMRI study on grip force and load force coupling." *Prog. Brain Res.* 142:171–188. (Cited on pages 20, 107, 137).

Kennedy, M. W., Schmiedeler, J. P., Crowell, C. R., Villano, M., Striegel, A. D., and Kuitse, J. 2011. "Enhanced feedback in balance rehabilitation using the Nintendo Wii Balance Board" [in English]. In *2011 IEEE 13th Int. Conf. e-Health Networking, Appl. Serv.* 162–168. IEEE. (Cited on page 96).

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." *Front. Hum. Neurosci.* 7 (June): 307. (Cited on page 53).

Knill, D. C. and Pouget, A. 2004. "The Bayesian brain: the role of uncertainty in neural coding and computation." *Trends Neurosci.* 27 (12): 712–9. (Cited on page 18).

Koeppen, A. H. 2011. "Friedreich's ataxia: pathology, pathogenesis, and molecular genetics." *J. Neurol. Sci.* 303 (1-2): 1–12. (Cited on page 130).

Koeppen, A. H. 2013. "Nikolaus Friedreich and degenerative atrophy of the dorsal columns of the spinal cord." *J. Neurochem.* 126 Suppl:4–10. (Cited on page 130).

Kording, K. 2007. "Decision Theory: What "Should" the Nervous System Do?" *Science (80-.).* 318 (5850): 606–610. (Cited on page 53).

Körding, K. P. and Wolpert, D. M. 2004. "Bayesian integration in sensorimotor learning." *Nature* 427 (6971): 244–7. (Cited on pages 18, 50, 154).

Körding, K. P. and Wolpert, D. M. 2006. "Bayesian decision theory in sensorimotor control." *Trends Cogn. Sc.* 10 (7): 319–326. (Cited on page 98).

Körding, K., Tenenbaum, J., and Shadmehr, R. 2006. "Multiple timescales and uncertainty in motor adaptation." In *Neural Inf. Process. Syst.* 19:745. Citeseer. (Cited on page 120).

Krakauer, J. W., Pine, Z. M., Ghilardi, M. F., and Ghez, C. 2000. "Learning of visuomotor transformations for vectorial planning of reaching trajectories." *J. Neurosci.* 20 (23): 8916–8924. (Cited on pages 21, 29, 32, 50).

Krakauer, J. W. and Mazzoni, P. 2011. "Human sensorimotor learning: adaptation, skill, and beyond." *Curr. Opin. Neurobiol.* 21 (4): 636–644. (Cited on pages 18, 53).

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 Biol.* 4 (10): e316. (Cited on pages 21, 28).

Krichmar, J. L. and Wagatsuma, H. 2011. *Neuromorphic and Brain-Based Robots.* 364. Cambridge University Press. (Cited on page 19).

Kuo, A. D. 1995. "An optimal control model for analyzing human postural balance." *IEEE Trans. Biomed. Eng.* 42 (1): 87–101. (Cited on pages 82, 95).

Kwakernaak Huibert/Sivan, R. 1972. "Linear Optimal Control Systems." (Cited on page 62).

Lackner, J. R. and Dizio, P. 1994. "Rapid adaptation to Coriolis force perturbations of arm trajectory." *J. Neurophysiol.* 72 (1): 299–313. (Cited on pages 29, 32, 50).

Lacquaniti, F., Soechting, J. F., and Terzuolo, C. A. 1982. "Some factors pertinent to the organization and control of arm movements." *Brain Res.* 252 (2): 394–7. (Cited on page 30).

Lai, T. L. 1998. *Sequential analysis.* Wiley Online Library. (Cited on page 53).

Li, W. 2006. "Optimal control for biological movement systems." PhD diss., University of California San Diego. (Cited on pages 85, 86).

Libri, V. et al. 2014. "Epigenetic and neurological effects and safety of high-dose nicotinamide in patients with Friedreich's ataxia: an exploratory, open-label, dose-escalation study." [in English]. *Lancet* 384 (9942): 504–13. (Cited on page 133).

MacIntosh, B. J., Mraz, R., McIlroy, W. E., and Graham, S. J. 2007. "Brain activity during a motor learning task: An fMRI and skin conductance study." *Hum. Brain Mapp.* 28 (12): 1359–1367. (Cited on page 108).

Malfait, N., Gribble, P. L., and Ostry, D. J. 2005. "Generalization of motor learning based on multiple field exposures and local adaptation." *J. Neurophysiol.* 93 (6): 3327–3338. (Cited on pages 21, 29, 53).

Malfait, N., Shiller, D. M., and Ostry, D. J. 2002. *Transfer of motor learning across arm configurations.* Technical report 22. McGill University, Montreal, Quebec, Canada H3A 1B1. (Cited on pages 21, 29, 30, 53).

Mamassian, P., Landy, M. S., and Maloney, L. T. 2002. "Bayesian modelling of visual perception." *Probabilistic Model. brain Percept. neural Funct.* 13–36. (Cited on page 18).

Mangel, M. and Clark, C. W. 1988. *Dynamic modeling in behavioral ecology.* Princeton University Press. (Cited on page 53).

Mantovan, M. C., Martinuzzi, A., Squarzanti, F., Bolla, A., Silvestri, I., Liessi, G., Macchi, C., Ruzza, G., Trevisan, C. P., and Angelini, C. 2006. "Exploring mental status in Friedreich's ataxia: a combined neuropsychological, behavioral and neuroimaging study." *Eur. J. Neurol.* 13 (8): 827–35. (Cited on page 131).

Marr, D. 1969. "A theory of cerebellar cortex." *J. Physiol.* 202 (2): 437–70. (Cited on pages 19, 20, 107, 137).

Marr, D. 1982. "Vision: A computational investigation." (Cited on page 98).

Mattar, A. A. G. and Ostry, D. J. 2007. "Modifiability of generalization in dynamics learning." *J. Neurophysiol.* 98 (6): 3321–9. (Cited on pages 29, 32, 50).

Mattar, A. A. G. and Ostry, D. J. 2010. "Generalization of dynamics learning across changes in movement amplitude." *J. Neurophysiol.* 104 (1): 426–38. (Cited on pages 29, 32, 50).

Mazzoni, P. and Krakauer, J. W. 2006. "An implicit plan overrides an explicit strategy during visuomotor adaptation." *J. Neurosci.* 26 (14): 3642–5. (Cited on pages 19, 61).

McNamara, J. and Houston, A. 1980. "The application of statistical decision theory to animal behaviour." *J. Theor. Biol.* 85 (4): 673–690. (Cited on page 53).

Mehta, B. and Schaal, S. 2002. "Forward models in visuomotor control." *J. Neurophysiol.* 88 (2): 942–53. (Cited on page 80).

Menon, S., Yu, M., Kay, K., and Khatib, O. 2014. "Haptic fMRI: Accurately estimating neural responses in motor, pre-motor, and somatosensory cortex during complex motor tasks." In *Eng. Med. Biol. Soc. (EMBC), 2014 36th Annu. Int. Conf. IEEE,* 2040–2045. IEEE. (Cited on pages 105, 108).

Miall, C. 2002. "Modular motor learning" [in English]. *Trends Cogn. Sci.* 6 (1): 1–3. (Cited on page 28).

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." *J. Physiol.* 567 (Pt 2): 651–64. (Cited on page 81).

Mnih, V. et al. 2015. "Human-level control through deep reinforcement learning" [in en]. *Nature* 518 (7540): 529–533. (Cited on page 156).

Mombarg, R., Jelsma, D., and Hartman, E. 2013. "Effect of Wii-intervention on balance of children with poor motor performance." *Res. Dev. Disabil.* 34 (9): 2996–3003. (Cited on page 96).

Montague, P. R., Dayan, P., and Sejnowski, T. J. 1996. "A framework for mesencephalic dopamine systems based on predictive Hebbian learning." *J. Neurosci.* 16 (5): 1936–47. (Cited on pages 53, 81).

Montcel, S. T. du et al. 2008. "Composite cerebellar functional severity score: validation of a quantitative score of cerebellar impairment." *Brain* 131 (Pt 5): 1352–61. (Cited on pages 131, 150).

Monti, M. M. 2011. "Statistical Analysis of fMRI Time-Series: A Critical Review of the GLM Approach." [in English]. *Front. Hum. Neurosci.* 5:28. (Cited on page 113).

Morasso, P., Baratto, L., Capra, R., and Spada, G. 1999. "Internal models in the control of posture." *Neural Networks* 12 (7-8): 1173–1180. (Cited on pages 80, 82, 95).

Mordatch, I., Lowrey, K., and Todorov, E. n.d. "Ensemble-CIO: Full-Body Dynamic Motion Planning that Transfers to Physical Humanoids." (Cited on pages 23, 152).

Mumford, J. and Nichols, T. 2006. "Modeling and inference of multisubject fMRI data" [in English]. *IEEE Eng. Med. Biol. Mag.* 25 (2): 42–51. (Cited on page 116).

Mumford, J. A. and Poldrack, R. A. 2007. "Modeling group fMRI data." *Soc. Cogn. Affect. Neurosci.* 2 (3): 251–7. (Cited on page 116).

Nagengast, A. J., Braun, D. A., and Wolpert, D. M. 2009. "Optimal control predicts human performance on objects with internal degrees of freedom." *PLoS Comput. Biol.* 5 (6). (Cited on pages 23, 54, 64, 72, 77, 152).

Nashner, L. M. and McCollum, G. 1985. "The organization of human postural movements: A formal basis and experimental synthesis" [in English]. *Behav. Brain Sci.* 8 (01): 135–150. (Cited on page 82).

Nichols, T. E. 2012. "Multiple testing corrections, nonparametric methods, and random field theory." *Neuroimage* 62 (2): 811–815. (Cited on page 115).

Niv, Y., Daw, N. D., and Dayan, P. 2006. "Choice values." *Nat. Neurosci.* 9 (8): 987–8. (Cited on page 53).

Nowak, D. A., Koupan, C., and Hermsdörfer, J. 2007. "Formation and decay of sensorimotor and associative memory in object lifting." *Eur. J. Appl. Physiol.* 100 (6): 719–26. (Cited on page 30).

O., B. J. 1985. *Statistical decision theory and Bayesian analysis.* Vol. Springer s. New York: Springer-Verlag. (Cited on page 53).

Orbán, G. and Wolpert, D. M. 2011. "Representations of uncertainty in sensorimotor control." *Curr. Opin. Neurobiol.* 21 (4): 629–35. (Cited on pages 19, 52).

Parker, S. T. and Gibson, K. R. 1977. "Object manipulation, tool use and sensorimotor intelligence as feeding adaptations in cebus monkeys and great apes." *J. Hum. Evol.* 6 (7): 623–641. (Cited on pages 30, 108).

Pélisson, D., Alahyane, N., Panouillères, M., and Tilikete, C. 2010. "Sensorimotor adaptation of saccadic eye movements." *Neurosci. Biobehav. Rev.* 34 (8): 1103–20. (Cited on page 22).

Piaget, J. 1954. *The construction of reality in the child.* New York: Basic Books. (Cited on pages 30, 108).

Poldrack, R. A. 2007. "Region of interest analysis for fMRI." *Soc. Cogn. Affect. Neurosci.* 2 (1): 67–70. (Cited on page 128).

Poline, J.-B. and Brett, M. 2012. "The general linear model and fMRI: does love last forever?" *Neuroimage* 62 (2): 871–80. (Cited on page 114).

Pouget, A., Beck, J. M., Ma, W. J., and Latham, P. E. 2013. "Probabilistic brains: knowns and unknowns." *Nat. Neurosci.* 16 (9): 1170–8. (Cited on page 18).

Puterman, M. L. 2014. *Markov decision processes: discrete stochastic dynamic programming.* John Wiley & Sons. (Cited on page 53).

Qiu, A., Lee, A., Tan, M., and Chung, M. K. 2015. "Manifold learning on brain functional networks in aging." *Med. Image Anal.* 20 (1): 52–60. (Cited on pages 151, 159).

Rabe, K., Livne, O., Gizewski, E. R., Aurich, V., Beck, A., Timmann, D., and Donchin, O. 2009. "Adaptation to visuomotor rotation and force field perturbation is correlated to different brain areas in patients with cerebellar degeneration." *J. Neurophysiol.* 101 (4): 1961–71. (Cited on page 19).

Risacher, S. L. and Saykin, A. J. 2013. "Neuroimaging biomarkers of neurodegenerative diseases and dementia." *Semin. Neurol.* 33 (4): 386–416. (Cited on page 131).

Rodriguez, M., Sylaidi, A., and Faisal, A. A. 2014. "Developing a Novel fMRI-Compatible Motion Tracking System for haptic motor control experiments." In *Neurotechnix.* (Cited on pages 27, 99, 105).

Roth, F. 2007. "Explicit design and adaptation in self-construction." PhD diss., Swiss Federal Institute of Technology, ETH Zurich. (Cited on page 18).

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Sasaki, Y., and Pütz, B. 1998. "Transition of brain activation from frontal to parietal areas in visuomotor sequence learning." *J. Neurosci.* 18 (5): 1827–1840. (Cited on page 108).

Santello, M. and Soechting, J. F. 2000. "Force synergies for multifingered grasping." *Exp. brain Res.* 133 (4): 457–67. (Cited on page 76).

Schmitz-Hübsch, T. et al. 2006. "Scale for the assessment and rating of ataxia: development of a new clinical scale." *Neurology* 66 (11): 1717–20. (Cited on pages 131, 135, 150).

Schmitz-Hübsch, T. et al. 2008. "SCA Functional Index: a useful compound performance measure for spinocerebellar ataxia." *Neurology* 71 (7): 486–92. (Cited on pages 131, 150).

Schultz, W. 1997. "A Neural Substrate of Prediction and Reward." *Science (80-.).* 275 (5306): 1593–1599. (Cited on page 156).

Scott, G., Hellyer, P. J., Hampshire, A., and Leech, R. 2015. "Exploring spatiotemporal network transitions in task functional MRI." *Human brain mapping* 36 (4): 1348–1364. (Cited on page 129).

Scott, S. H. 2004. "Optimal feedback control and the neural basis of volitional motor control." *Nat. Rev. Neurosci.* 5 (7): 532–546. (Cited on pages 19, 20, 22, 54, 98).

Scott, S. H. 2006. "Neuroscience: converting thoughts into action." *Nature* 442 (7099): 141–2. (Cited on page 25).

Shadmehr, R. and Brashers-Krug, T. 1997. "Functional stages in the formation of human long-term motor memory." *J. Neurosci.* 17 (1): 409–19. (Cited on page 30).

Shadmehr, R. and Moussavi, Z. M. 2000. *Spatial generalization from learning dynamics of reaching movements.* Technical report. (Cited on pages 21, 29, 53, 98).

Shadmehr, R. and Mussa-Ivaldi, F. A. 1994. "Adaptive representation of dynamics during learning of a motor task." *J. Neurosci.* 14:3208–3224. (Cited on pages 21, 22, 28–30, 40, 49, 53, 98).

Shadmehr, R. and Krakauer, J. W. 2008. "A computational neuroanatomy for motor control." *Exp. Brain Res.* 185 (3): 359–381. (Cited on pages 19, 52).

Shadmehr, R., Smith, M. A., and Krakauer, J. W. 2010. "Error correction, sensory prediction, and adaptation in motor control." [in en]. *Annu. Rev. Neurosci.* 33:89–108. (Cited on page 53).

Shih, J. J., Krusienski, D. J., and Wolpaw, J. R. 2012. "Brain-computer interfaces in medicine." *Mayo Clin. Proc.* 87 (3): 268–79. (Cited on page 25).

Simmonds, A. J., Leech, R., Iverson, P., and Wise, R. J. S. 2014. "The response of the anterior striatum during adult human vocal learning." *J. Neurophysiol.* 112 (4): 792–801. (Cited on page 108).

Simpkins, A. and Todorov, E. 2011. "Complex object manipulation with hierarchical optimal control" [in English]. In *2011 IEEE Symp. Adapt. Dyn. Program. Reinf. Learn.* 338–345. IEEE. (Cited on page 30).

Skovronsky, D. M., Lee, V. M.-Y., and Trojanowski, J. Q. 2006. "Neurodegenerative diseases: new concepts of pathogenesis and their therapeutic implications." *Annu. Rev. Pathol.* 1:151–70. (Cited on pages 25, 130).

Smith, A. P. R., Dolan, R. J., and Rugg, M. D. 2004. "Event-related potential correlates of the retrieval of emotional and nonemotional context." *J. Cogn. Neurosci.* 16 (5): 760–75. (Cited on pages 113, 135).

Smith, M. A., Ghazizadeh, A., and Shadmehr, R. 2006. "Interacting adaptive processes with different timescales underlie short-term motor learning." *PLoS Biol.* 4 (6): 1035–1043. (Cited on page 129).

Smith, S. M. 2012. "The future of FMRI connectivity." *Neuroimage* 62 (2): 1257–66. (Cited on page 129).

Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., Ramsey, J. D., and Woolrich, M. W. 2011. "Network modelling methods for FMRI." *Neuroimage* 54 (2): 875–91. (Cited on page 129).

Smith, S. M. and Nichols, T. E. 2009. "Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference." *Neuroimage* 44 (1): 83–98. (Cited on page 115).

Smith, S. M. et al. 2009. "Correspondence of the brain's functional architecture during activation and rest." *Proc. Natl. Acad. Sci. U. S. A.* 106 (31): 13040–5. (Cited on page 125).

Stam, C. J. 2014. "Modern network science of neurological disorders." *Nat. Rev. Neurosci.* 15 (10): 683–695. (Cited on pages 148, 149, 151, 157).

Stefanescu, M. R., Dohnalek, M., Maderwald, S., Thürling, M., Minnerop, M., Beck, A., Schlamann, M., Diedrichsen, J., Ladd, M. E., and Timmann, D. 2015. "Structural and functional MRI abnormalities of cerebellar cortex and nuclei in SCA3, SCA6 and Friedreich's ataxia." *Brain* 138 (Pt 5): 1182–97. (Cited on page 131).

Subramony, S. H., May, W., Lynch, D., Gomez, C., Fischbeck, K., Hallett, M., Taylor, P., Wilson, R., and Ashizawa, T. 2005. "Measuring Friedreich ataxia: Interrater reliability of a neurologic rating scale." *Neurology* 64 (7): 1261–2. (Cited on page 131).

Sutton, R. S. and Barto, A. G. 1998. *Introduction to reinforcement learning.* MIT Press. (Cited on pages 23, 53, 81, 95).

Sylaidi, A. and Faisal, A. 2012. *What is the hierarchical representation of tasks involving objects with complex internal dynamics?* (Cited on pages 27, 54, 80, 98).

Sylaidi, A., Lourenco, P., Nageshwaran, S., Lin, C.-H., Rodriguez, M., Festenstein, R., and Faisal, A. A. 2015. "fMRI-compatible haptic object manipulation system for closed-loop motor control studies." In *Neural Eng. (NER), 2015 7th Int. IEEE/EMBS Conf.* IEEE. (Cited on page 27).

Szymański, P., Markowicz, M., Janik, A., Ciesielski, M., and Mikiciuk-Olasik, E. 2010. "Neuroimaging diagnosis in neurodegenerative diseases." *Nucl. Med. Rev. Cent. East. Eur.* 13 (1): 23–31. (Cited on page 131).

Taylor, J. A., Klemfuss, N. M., and Ivry, R. B. 2010. "An explicit strategy prevails when the cerebellum fails to compute movement errors." *Cerebellum* 9 (4): 580–6. (Cited on page 19).

Tcheang, L., Bays, P. M., Ingram, J. N., and Wolpert, D. M. 2007. "Simultaneous bimanual dynamics are learned without interference." *Exp. brain Res.* 183 (1): 17–25. (Cited on page 30).

Tee, K. P., Franklin, D. W., Kawato, M., Milner, T. E., and Burdet, E. 2010. "Concurrent adaptation of force and impedance in the redundant muscle system." *Biol. Cybern.* 102 (1): 31–44. (Cited on page 53).

Thomik, A. A. C., Haber, D., and Faisal, A. A. 2013. "Real-time movement prediction for improved control of neuroprosthetic devices." In *Neural Eng. (NER), 2013 6th Int. IEEE/EMBS Conf.* 625–628. IEEE. (Cited on page 99).

Thorndike, E. L. 1911. *Animal intelligence; experimental studies, by Edward L. Thorndike.* 1–324. New York, The Macmillan company, (cited on page 156).

Thoroughman, K. A. and Taylor, J. A. 2005. "Rapid reshaping of human motor generalization." *J. Neurosci.* 25 (39): 8948–53. (Cited on page 29).

Todorov, E. 2004. "Optimality principles in sensorimotor control." *Nat. Neurosci.* 7 (9): 907–915. (Cited on pages 22, 54, 75, 85, 86, 98).

Todorov, E. and Jordan, M. I. 2002. "Optimal feedback control as a theory of motor coordination." *Nat. Neurosci.* 5:1226–1235. (Cited on pages 22, 54, 76).

Vaadia, E. and Birbaumer, N. 2009. "Grand challenges of brain computer interfaces in the years to come." *Front. Neurosci.* 3 (2): 151–4. (Cited on pages 25, 157).

Vaziri, S., Diedrichsen, J., and Shadmehr, R. 2006. "Why does the brain predict sensory consequences of oculomotor commands? Optimal integration of the predicted and the actual sensory feedback." *J. Neurosci.* 26 (16): 4188–4197. (Cited on pages 19, 21, 98).

Verduzco-Flores, S. O. and O'Reilly, R. C. 2015. "How the credit assignment problems in motor control could be solved after the cerebellum predicts increases in error." *Front. Comput. Neurosci.* 9:39. (Cited on page 50).

Wagner, D. and Schmalstieg, D. 2007. *Artoolkitplus for pose tracking on mobile devices.* na. (Cited on page 102).

Wagner, M. J. and Smith, M. A. 2008. "Shared internal models for feedforward and feedback control." *J. Neurosci.* 28 (42): 10663–73. (Cited on page 80).

Wei, K. and Körding, K. 2009. "Relevance of error: what drives motor adaptation?" *J. Neurophysiol.* 101 (2): 655–64. (Cited on page 50).

White, O. and Diedrichsen, J. 2010. "Responsibility assignment in redundant systems." *Curr. Biol.* 20 (14): 1290–5. (Cited on page 50).

Wiestler, T. and Diedrichsen, J. 2013. "Skill learning strengthens cortical representations of motor sequences." *Elife* 2:e00801. (Cited on page 108).

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., and Nichols, T. E. 2014. "Permutation inference for the general linear model." *Neuroimage* 92:381–97. (Cited on page 116).

Witney, A. G., Goodbody, S. J., and Wolpert, D. M. 2000. "Learning and decay of prediction in object manipulation." *J. Neurophysiol.* 84 (1): 334–43. (Cited on page 30).

Witney, A. G. and Wolpert, D. M. 2003. "Spatial representation of predictive motor learning." *J. Neurophysiol.* 89 (4): 1837–43. (Cited on page 30).

Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. 1995. "An internal model for sensorimotor integration." *Science* 269 (5232): 1880–2. (Cited on pages 19, 80).

Wolpert, D. M., Diedrichsen, J., and Flanagan, J. R. 2011. "Principles of sensorimotor learning." *Nat. Rev. Neurosci.* 12 (12): 739–51. (Cited on pages 18–20, 52, 98, 107, 155).

Wolpert, D. M. and Flanagan, J. 2001. "Motor prediction" [in English]. *Curr. Biol.* 11 (18): R729–R732. (Cited on page 28).

Wolpert, D. M. and Landy, M. S. 2012. "Motor control is decision-making." *Curr. Opin. Neurobiol.* 22 (6): 996–1003. (Cited on page 50).

Wolpert, D. M., Miall, R. C., and Kawato, M. 1998. "Internal models in the cerebellum." *Trends Cogn. Sci.* 2 (9): 338–347. (Cited on pages 20, 107, 137).

Wolpert, D. and Kawato, M. 1998. "Multiple paired forward and inverse models for motor control." *Neural Networks* 11 (7-8): 1317–1329. (Cited on page 28).

Woolrich, M. W., Ripley, B. D., Brady, M., and Smith, S. M. 2001. "Temporal autocorrelation in univariate linear modeling of FMRI data." *Neuroimage* 14 (6): 1370–86. (Cited on page 115).

Wu, H. G., Miyamoto, Y. R., Gonzalez Castro, L. N., Ölveczky, B. P., and Smith, M. A. 2014. "Temporal structure of motor variability is dynamically regulated and predicts motor learning ability." *Nat. Neurosci.* 17 (2): 312–21. (Cited on pages 68, 75).

Yang, C., Ganesh, G., Haddadin, S., Parusel, S., Albu-Schaeffer, A., and Burdet, E. 2011. "Human-Like Adaptation of Force and Impedance in Stable and Unstable Interactions" [in English]. *IEEE Trans. Robot.* 27 (5): 918–930. (Cited on page 53).