On the somatosensory and reinforcement bases of human motor learning

Ananda Sidarta

Department of Psychology

Integrated Program in Neuroscience

McGill University, Montreal

October 2018

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Acknowledgement

I would like to thank Professor David Ostry for his instructions, support, care, and supervision over the last six years. I am also thankful to ALL members of the Motor Control Lab from whom I have learned a great deal and received endless supports. I am honestly grateful to work in a place where there are ample lab resources and equipment. Special thanks are also given to the undergraduates in the Department of Psychology who, from time to time, have joined the lab and helped in conducting the experiments. I also thank my thesis committee, Dr. John Kalaska and Dr. Alexander Thiel, for their feedback and time during our annual review meeting.

Lastly, I am grateful to my wife who has been very supportive throughout this journey. I wish to dedicate this work to my parents back home.

Abstract

Humans have a remarkable capability to adapt and learn new things, including new motor skills. For example, one can learn to play a short sequence of musical notes by navigating one’s fingers on a piano or a guitar. Often, motor skill learning involves a goal-directed pointing or reaching to a clearly-defined visual target, such as making a purchase using a touch screen. In cases such as swimming or skiing, however, visual information is often limited or occluded. In order to successfully learn the task, one has to rely primarily on somatosensory information to form sensory goals or targets, or in effect to discover what it feels like to make a successful movement. In search for actions that are successful, one learns through trial and error and exploration, which leads to either unsuccessful or successful movements. Naturally, an efficient way of learning is to avoid committing the same mistakes, and instead, try to perform repeatedly successful actions. Assuming that moment-to-moment sensory information is important, the best possible candidate for such a mechanism would be a working memory store.

This thesis examines the role of reinforcement and somatosensory information in the initial phases of motor learning using arm-reaching tasks as a model. The tasks involve making reaching movements to a hidden visual target. Binary positive feedback is provided throughout training as a form of reinforcement. Such learning brings about changes in functional networks in specific brain regions. More specifically, increases in functional connectivity related to reduction of movement error were found in sensorimotor areas of the brain, including supplementary motor area and second somatosensory cortex. In contrast, connectivity increases related to the number of successful trials (and thus, reinforcement) were observed in the ventral striatum and ventromedial prefrontal cortex. Thus, learning involves both sensorimotor and reward-related areas of the brain. Trial-by-trial changes in motor behavior (and hence variability) were also examined in this study, where it was found that variability was less following a successful trial than an unsuccessful trial, presumably reflecting a sign of exploration. Changes in connectivity involving putamen and sensorimotor cortices were observed that varied systematically with the magnitude of trial-by-trial variability. Changes in connectivity that were related to motor exploration following unsuccessful trial were also found in second somatosensory cortex and lateral prefrontal cortex, regions known to be involved in somatic working memory and decision-making.

The role of somatic working memory in motor learning was assessed in a second study. In two separate experiments, it was found that individual differences in somatosensory working memory performance were reliably related to the amount of learning. Performance of somatosensory working memory was shown to decay as a function of the number of sequential items in the memory list, where the overall capacity was estimated to be around 2 items. Another experiment measures working memory by using subject’s own movements. During learning period, working memory preferentially stores reinforced movements more than non-reinforced ones. The findings point to the critical roles of somatosensory working memory in reinforcement-based in motor learning.

Résumé

Les humains ont une capacité remarquable à s'adapter et à apprendre de nouvelles choses, y compris de nouvelles compétences motrices. Par exemple, on peut apprendre à jouer une courte séquence de notes de musique en bougeant ses doigts sur un piano ou une guitare. Souvent, l'apprentissage d'habiletés motrices implique un pointage ciblé vers un objectif ou une cible visuelle clairement définie, comme effectuer une sélection sur un écran tactile. Dans des cas tels que la natation ou le ski, les informations visuelles ne sont souvent pas disponibles ou occultées. Pour réussir à apprendre la tâche, il faut utiliser les informations somatosensorielles pour former des objectifs ou des cibles sensorielles, ou pour découvrir ce que cela signifie de réussir un mouvement. À la recherche d'actions réussies, on traverse une série d'essais et erreurs ainsi que d'exploration. Ces processus entraînent des mouvements à la fois infructueux et réussis. Naturellement, un moyen efficace d'apprendre est d'éviter de répéter les mêmes erreurs et, au contraire, de ne mener à bien que des actions réussies. En supposant qu'un stockage sensoriel instantané soit impliqué, le meilleur candidat possible pour un tel mécanisme est la mémoire de travail.

Cette thèse examine le rôle du renforcement et de l'information somatosensorielle dans les phases initiales de l'apprentissage moteur en utilisant des tâches de mouvement d'atteinte avec le bras. Les tâches impliquent de faire des mouvements vers une cible visuelle cachée. Une renforcement positif binaire a été fournie tout au long de l'entraînement. Il a été constaté que le renforcement positif et les informations somatosensorielles étaient importants pour mener l'apprentissage en l'absence d'une cible visuelle claire. Un tel apprentissage entraîne des changements dans les réseaux fonctionnels de régions spécifiques du cerveau. On a constaté que les augmentations de la connectivité fonctionnelle varient en fonction du nombre d'essais réussis. Un schéma similaire a également été observé avec une augmentation de la connectivité impliquant le striatum ventral et le cortex préfrontal ventromédian. Les modifications du comportement moteur effectuées au cours des essais ont également été examinées en fonction des résultats des essais antérieurs, où, comparativement à un essai infructueux, un essai réussi a entraîné une plus faible variabilité des mouvements successifs. Les changements de connectivité variant avec la quantité de variabilité ont été identifiées comme comprenant le putamen et le cortex sensorimoteur. Aussi, les changements de connectivité variant avec la quantité d'exploration après un essai infructueux ont été identifiées comme comprenant le second cortex somatosensoriel et le cortex préfrontal latéral, régions connues pour être impliquées dans la mémoire de travail somatique et les tâches décisionnelles.

Le rôle de la mémoire de travail somatique dans l'apprentissage moteur a été démontré dans une deuxième étude. Dans deux expériences distinctes, il a été constaté que les différences individuelles dans la performance de la mémoire de travail somatosensorielle étaient liées de manière fiable à l’ampleur de l'apprentissage. Pendant l'apprentissage, cette mémoire de travail se souvient préférentiellement des mouvements renforcés par rapport aux mouvements non renforcés. Ces résultats mettent en évidence le rôle critique de la mémoire de travail somatosensorielle dans l'apprentissage par le renforcement du moteur.

Contribution to knowledge

The first study combines both behavioral and neuroimaging work and has been published in a peer-reviewed journal, *Journal of Neuroscience*. The work extends prior studies on reinforcement learning in behavioral psychology and cognitive neuroscience by putting emphasis its role in human motor learning. More importantly, the work shows how reinforcement-based motor learning recruits specific brain regions comprising reward and reward-based decision-making network (ventral striatum and ventromedial prefrontal cortex), and somatic network in the brain (somatosensory cortices and ventrolateral prefrontal cortex). The recruitment of reward-related areas suggests that performing an action successfully produces the same effect as the one observed when getting monetary or other kinds of reward. The work also demonstrates plasticity associated with trial-to-trial change in motor behavior involving putamen and sensorimotor cortices.

The experiments presented in the second study have just been accepted for a publication in the *Journal of Neurophysiology*. The work examines the relationship between somatosensory working memory and reinforcement-based motor learning using two separate experiments. In both experiments, it is found that individual differences in memory performance vary with the amount of motor learning, and that working memory preferentially remembers reinforced movements during learning. The findings especially contribute to the under-represented literature on working memory in the somatosensory domain, especially the significance of sensory working memory to support reinforcement-based motor learning.

Contribution of authors

In all studies, Ananda Sidarta designed the experiments, coded the robot, conducted the experiments, analyzed the data, and wrote the manuscripts. Prof. David Ostry helped in designing the experiment, discussing the findings, and editing the manuscript. Nicolo F. Bernardi helped in designing the behavioral task in Study-1, supported the data analyses, and edited the manuscripts. Shahab Vahdat helped in designing the neuroimaging aspects of Study-1, supported the data analyses, and edited the manuscripts. Floris VanVugt helped in designing the experiments, provided supports in R and Python language, and edited the manuscript for Study-2.

Table of Contents

Chapter 1 : General introduction

1.1. Introduction

During my first winter in Canada, I decided to learn alpine skiing at Mount Sutton. As someone who has never learnt or experienced skiing before, the learning itself proved to be daunting. During training, the ski instructor showed us how the snowplow technique was performed that allows skiers to control speed while moving downhill. Although the instructions sounded trivial, the actual implementation involved arduous trial and error, repeated falls, and bruises at the least. I was finally able to grasp the sense of snowplowing after three consecutive outings.

The situation described above is common to naïve learners when learning for the first time a new motor skill. A striking similarity among these learning instances is the fact that the sensory targets for movement are unclear, meaning, that learners do not initially know the desired sensory state which serves as a reference in order to perform the motor skill. Much of the effort early in learning entails discovering sensory targets. In cases where visual target information is limited such as in swimming or skiing, the challenge is essentially to establish sensory targets in somatosensory terms, that is, to discover the best arm or leg configuration in order to perform the task. Another feature of these initial stages of learning is that information is gathered through trial and error or exploration (Newell, 1991) where learners collect information about the task, possible movements and their sensory outcomes. During this exploratory activity movements are often unsuccessful, so remembering prior movements is important so that the same undesired movements are not repeated. Indeed, what is perhaps more important is remembering the successful ones so that one eventually learns how to do the task correctly. Such temporary storage of prior sensory information in an accessible state for a later use is thought to involve working memory.

This raises a number of questions: what changes to the brain occur in the initial stages of learning while searching for somatosensory targets? What is the role of working memory in motor learning? Is the working memory in this context specific to the somatosensory domain or does it entail a more general cognitive or spatial sensory memory store? Are all movements, both successful and unsuccessful, remembered equally well in working memory at the time of learning? How large or small is the sensory working memory store? That is, do only the most recent movements have an effect on upcoming attempts or is its capacity greater? These questions will be addressed in this thesis. Using arm reaching tasks, I will investigate how humans learn new motor skills by discovering somatosensory targets through trial and error. Although the somatosensory system covers different types of bodily senses, this thesis focuses strictly on proprioception or kinaesthesia (terms which describe limb position sense), and cutaneous information related to reaching movements.

In this chapter, I will review the theoretical background pertaining to the somatosensory system and motor learning through reinforcement, and their role in discovering somatosensory targets early in learning. Neuroimaging is an important non-invasive tool in order to understand brain mechanisms governing this type of motor learning. I will therefore discuss the appeals and challenges in using a technique called resting-state functional MRI. This will be followed by two chapters with experimental work. The first is a study that examines brain plasticity subserving the reinforcement-based motor learning where both behavioral and neuroimaging methods are employed. The subsequent chapter examines the roles of somatosensory working memory in motor learning using two different sets of experiments. The first experiment explores somatosensory working memory capacity and its relation with overall changes in motor behavior following learning. The experiment asks whether individual differences in motor learning outcomes are related to differences in working memory capacity. The second experiment measures working memory of one's own movements on a trial by trial basis directly during learning. This experiment asks whether successful and unsuccessful movements during learning are remembered equally well.

1.2. Different modes of motor learning

Based on differences in the types of information used during learning, there are several ways in which motor learning occurs. Studies of human motor learning in the laboratory are typically based on a well-established adaptation paradigm where error information is used as a learning signal (Shadmehr et al., 2010). The error can be introduced in the form of a visual rotation (in the case of a prism adaptation and visuomotor rotation tasks) or a velocity-dependent field (in the case of force field adaptation) (Shadmehr and Mussa-Ivaldi, 1994). The discrepancy between the predicted reaching movement and the actual movement outcome gives rise to *sensory prediction errors* that are used to drive learning. It is thought that error-based motor learning is mediated by updates of feedforward models of the sensorimotor system that enable the selection of motor commands which progressively correct for systematic errors (Wolpert et al., 1995; Kawato, 1999).

According to theories in error-based learning, adaptation scales in proportion to the size of the error (Kawato et al., 1987; Jordan and Rumelhart, 1992) If the sign of the error changes within the same training period, people can quickly relearn within minutes (Fine and Thoroughman, 2007). Error-based learning can occur regardless whether the error is introduced abruptly or gradually (Malfait and Ostry, 2004; Criscimagna-Hemminger et al., 2010). Consequently, one would expect that learning processes would be compromised if information regarding error magnitude is unavailable. In the early stages of motor learning, desired sensory states relative to which error might be computed are in the process of being formed. In such cases, another form of motor learning that is based on reinforcement may enable progress. Unlike error-based learning, reinforcement learning is restricted to binary feedback related to the outcome of the movement (success or failure) as a learning signal that can simply take the form of positive or negative feedback, but can also serve as reward or punishment. Reinforcement learning can occur in conjunction with sensorimotor adaptation. For instance, using a visuomotor rotation task, Izawa and Shadmehr showed that adaptation can be driven by both sensory prediction error and positive reinforcement (reward). While the adaptation processes result in a sensory recalibration, the latter does not, suggesting that the two types of learning may recruit independent processes (Izawa, 2011). A more detailed discussion of reinforcement learning and how it can be used to model motor learning will be presented in Section 1.4. The neural substrates of error-based adaptation and reinforcement learning will be discussed in Section 1.6 below.

A third form of motor learning is use-dependent learning, which is the motor analogue of the experience-dependent plasticity that occurs when animals experience prolonged sensory manipulation (for example, Jenkins et al., 1990; Gilbert and Wiesel, 1992). In use-dependent learning, motor learning proceeds through an extended repetition of the same movements using the same limb. Such extended training has been shown to promote larger limb representation in the sensorimotor cortices of squirrel monkeys (Nudo et al., 1996) and rats (Kleim et al., 1998), and has been used to study motor learning in humans (Bütefisch et al., 2000; Diedrichsen et al., 2010). In contrast to reinforcement-based learning, use-dependent learning typically occurs without the involvement of positive or negative feedback or error information.

Finally, there is motor sequence learning in which a person learns to perform a series of sequential actions and progress is often measured by the reduction in the serial reaction time (Nissen and Bullemer, 1987; Robertson, 2007). Neuroimaging studies have been useful in elucidating how motor sequence learning develops (Doyon et al., 2002; Lehéricy et al., 2005; Penhune and Doyon, 2005; Steele and Penhune, 2010). Early in learning, brain activity in both cortico-cerebellar and cortico-striatal networks is observed. It is thought that this period involves error correction in order to construct feedforward models of motor sequences. At the same time, learning is also driven by stimulus-response association, linking actions and outcomes (reward). As learning progresses, the involvement of cortico-cerebellar structures is reduced but there is an increase in activity in a specific area of the striatum, the caudoventral portion of the putamen. Overall, the findings are consistent with the idea that the initial phases of sequence learning involve both error-based and reinforcement-based processes which are complementary and can occur in parallel (Doya, 2000; Hikosaka et al., 2002).

1.3. Somatosensory contribution to motor learning

Somatosensation refers to senses associated with one own's body. The word is used interchangeably with bodily or *somatic* senses. Sherrington distinguished three different categories of bodily senses:

(1) *Exteroception* refers to senses that register direct interaction with the external world in the form tactile stimuli (such as gentle touch and vibration), heat (thermoreceptor), and pain (nociceptor).

(2) *Interoception* relates to sensation associated with the internal organs that is important for regulating autonomic functions.

(3) *Proprioception* refers to the sense of limb position and movements in space. Another term that is sometimes used interchangeably with proprioception is *kinaesthesia*.

The somatic senses that are critical in voluntary movements belong to the last category, i.e. proprioception. Finer motor tasks involving finger movements have been shown to also involve tactile exploration, creating another dimension of somatosensation called *haptic* (Srinivasan and LaMotte, 1995; Lederman and Klatzky, 2009). Electrophysiological recording in primates also shows tactile activity in primary somatosensory cortex (S1) during active movement (Cohen et al., 1994). There is also some evidence that inputs related to skin deformation likewise provide proprioceptive information (Edin, 2004). These suggest that somatosensation associated with voluntary movements may well extend beyond just proprioception.

The close interaction between frontal motor area and somatosensory cortex in the parietal lobe is evident given the anatomical connections based on studies in non-human primates (Pons and Kaas, 1986; Darian-Smith et al., 1993; Burton and Fabri, 1995). Neurons in somatosensory cortex (S1) display directional tuning properties associated with active voluntary movements, similar to those in primary motor cortex (M1) (Prud'Homme and Kalaska, 1994). Behaviorally, the contribution of somatosensory system to motor control in humans has been well documented in studies where the sensory afferents were either disrupted by tendon vibration (Cordo et al., 1995) or missing (Ghez et al., 1995; Sainburg et al., 1995; Sarlegna et al., 2006). Together with the vestibular system, somatosensory feedback is essential for postural control and stability (Horak et al., 1990). In situations with limited visual feedback, deafferented patients lacking somatosensory information are unable to properly perform multi-joint movements. Even when visual feedback is present, lacking somatosensory feedback increases the reaction time of reaching to a visually-perturbed target location (Balslev et al., 2007). Sainburg et al. (1995) hypothesized that the afferent inputs are important for updating feedforward models of limb dynamics that is used to control limb movement.

Somatosensory cortex is not only important in providing afferent information, it is also the immediate cause of plasticity in the motor cortex through a long-term potentiation (LTP) (Iriki et al., 1989; Keller et al., 1990). Indeed, somatosensory system have been shown to be important in motor learning by a series of studies. The first type of such studies shows that motor learning is impaired by disrupting the somatosensory system. Lesions to primary somatosensory cortex (S1) in cats (Sakamoto et al., 1989) and non-human primates (Pavlides et al., 1993) are shown to inhibit learning a simple motor task such as picking up food from a rotating platform. Likewise, inhibiting S1 in mice with optogenetics while the animals are moving a lever impairs motor adaptation to force applied to the lever (Mathis et al., 2017). In humans, similar results have been demonstrated using transcranial magnetic stimulation (TMS). Suppressing S1 with 1 Hz repetitive TMS pulses yields a reduction in performance of a tracking task (Vidoni et al., 2010).

Additional evidence comes from studies which show that somatosensory training facilitates motor learning. In a behavioral study with humans, the accuracy of a tracing task with a robotic manipulandum is improved with proprioceptive training that exposes the participants to the desired hand trajectories (Wong et al., 2012). Similarly, if participants are presented with a proprioceptive discrimination task related to their required movements, they are able to benefit and improve their movement accuracy in subsequent force field adaptation tasks (Darainy et al., 2013). Daily passive wrist movements over the period of a month result in greater activity in primary sensorimotor cortex and supplementary motor area (SMA) as measured by functional MRI (Carel et al., 2000). Providing vibratory cues in addition to passive proprioceptive training also improves both wrist position acuity and movement accuracy (Cuppone et al., 2016). The vibratory cues serve as performance feedback to indicate errors and thus facilitate motor learning.

A recent behavioral study in our lab examined the role of somatosensory system in the initial stages of motor learning (Bernardi et al., 2015). The study incorporated a behavioral paradigm that mimicked the natural situation of learning a novel task in which the reference somatosensory states of movements are unclear. The motor task was to move to a target that was occluded from view and binary positive feedback was provided for each successful trial. Under these conditions, learning was based at least in part on trial and error. The authors found that the somatosensory system coupled with positive feedback that serves as reinforcement contributes substantially to motor learning. In the following section, the relevance of reinforcement learning to model the initial stages of motor learning will be discussed.

1.4. Reinforcement-based motor learning

The reinforcement learning that we know today originated with Thorndike (1911) who proposed a set of ideas governing learning by trial and error. Thorndike named the theory “Law of Effect” because it described how one learns a particular action based on the outcome of that action. The theory was later expanded by Skinner (1953) to explain learning behavior in animals using what he called *operant conditioning* principles. According to Skinner, positive (reward) and negative reinforcement (punishment) shape behavior differently (Skinner, 1963). In reinforcement learning, what is learnt is a rule governing performance (or behavior) and the value of its outcome (reward and punishment) (Sutton and Barto, 1998; Doya, 2000; Dayan and Niv, 2008). As Skinner posited, rewards such as food pellets and juice are able to strengthen a behavior by encouraging animals to repeatedly produce the same behavior. Punishment such as electric shock, on the other hand, results in avoidance of unwanted actions. Prior evidence suggested that punishment is generally less effective for motor learning than reward. Compared with practice using punishment as feedback, the use of rewarding feedback improves long-term consolidation of learned motor skills (Abe et al., 2011).

The functional significance of reinforcement learning in finding somatosensory targets during motor learning can be posited as follows. Bernstein (1967) noted that positioning the limbs requires that one solve a *degree-of-freedom* problem where each joint is capable of moving independently, but the total number of degrees of freedom of the limb exceeds that necessary to specify the position of the limb endpoint in space. The need to select among excess degrees of freedom is exacerbated by the fact that any single-joint is controlled by many muscles with varied mechanical actions. Accordingly, there is initially uncertainty in how to select the best joint-configuration to perform a task successfully. Reinforcement learning, or more correctly feedback concerning movement outcome, may contribute to this selection process. Through trial and error, a novice learns to assign credit or value based on the outcome of one's actions, that eventually leads to the "best" possible action-value choices (Daw and Doya, 2006; Gershman et al., 2009). Each time an action is performed the value of one's action is provided by the feedback following the performance of that action. Successful actions are reinforced using positive feedback or reward, and the main goal of learning is to maximize long-term reward by accumulating as many successful actions as possible. These successful actions presumably establish the somatosensory targets needed for future movements. As such, reinforcement learning can be used as an attractive model for the early stages of motor learning.

Reinforcement learning tasks have been used to study human motor learning. Huang and colleagues demonstrated that while adaptation results in error reduction, the association between adapted movements and reward (success) enables faster relearning, a phenomenon reflecting *savings* (Huang et al., 2011). Additional training with binary feedback was also shown to improve motor retention in a visuomotor adaptation task. These findings suggest that even if adaptation and reinforcement learning are independent both processes may contribute simultaneously to learning. Other studies have looked into the role of reinforcement in modulating movement variability. Specifically, the history of past reward was shown to influence movement variability during learning (Pekny et al., 2015), but only if participants were aware of the task dimension that the reward operated on (Manley et al., 2014; Mehler et al., 2017).

In the work that follows, I use arm reaching tasks where the visual target is occluded and movement error information is unavailable. Information on movement outcome which guides learning is provided to the subjects (or *knowledge of results*, KR), that is, whether a particular movement is successful. Although different terminologies exist, such information will be called ‘positive reinforcement’ or ‘positive feedback’, rather than the term ‘reward’ which implies a presumed effect. “Successful action” is used to denote “action that is reinforced or rewarded”.

1.5. The role of working memory in motor learning

As a novice learns a new motor skill through a series of unsuccessful and successful movements, certain online memory storage is presumably involved to plan subsequent movements. The putative role of this kind of memory system will be discussed here.

Atkinson & Shiffrin posited that the human memory system can be divided into three distinct components (Atkinson and Shiffrin, 1968). Although their work related to verbal memory it can be readily extended to motor control and learning:

(1) Sensory memory registers incoming stimuli from the five senses in a transient fashion. This includes bodily or somatic senses. *Iconic* (visual) and *echoic* (auditory) sensory memory have been studied most frequently. According to Sterling, the size of this sensory memory is unlimited but the duration is short-lived in the order of seconds (Sperling, 1960).

(2) Short-term memory is used to maintain sensory items for a longer period, on the order of seconds or 10s of seconds. Although the influential “magic number 7 plus minus 2” was once regarded as the capacity limit of this short-memory store (Miller, 1956), the exact size is still an on-going debate. Working memory is a more specialized type of short-term memory where the items are readily available for certain purposes (see definition below).

(3) Long-term memory is a much more durable and larger store that consolidates information in short-term memory for long-term storage.

The word “working memory” was coined by Miller et al. (1960) to point to a type of memory for planning and directing behavior such as solving an arithmetic number problem (Miller et al., 1960). Much of the initial work on working memory was done by Baddeley, who focused on verbal memory and defined working memory as a system that temporarily maintains and stores task-relevant information during performance of a cognitive task (Baddeley and Hitch, 1974; Baddeley, 1992). The Baddeley model of working memory describes the system as a combination of an executive controller and two storage subunits, that is, a phonological loop and visuospatial sketchpad. The latter subunit is thought to comprise shared compartments that store visual stimuli, spatial and movement information (Smyth and Pendleton, 1990; Salway and Logie, 1995).

Working memory contributes to cognitive performance in humans in tasks including reasoning, learning, and comprehension. There are numerous pieces of evidence suggesting the importance of verbal working memory capacity in human learning (Baddeley et al., 1998; Gathercole et al., 1999; Gupta and Tisdale, 2009). The digit span task, a cognitive task in which subjects are required to hold a number of individual digits in memory, is thought to tap in verbal working memory and also mental chunking which are often used as predictors for mathematical learning ability (Gobet et al., 2001; Bull et al., 2008). More recently, it was shown that visuospatial working memory capacity correlate with the rate of sequence learning and visuomotor adaptation (Bo and Seidler, 2009; Bo et al., 2011). Using a similar visuomotor task, the role of visuospatial working memory has also been demonstrated when adaptation involved either an explicit strategy or an abruptly introduced rotation, but not implicit strategy (Christou et al., 2016).

Establishing a measure of working memory that is relevant to motor learning requires an appreciation of how this information might be used. To answer this question, I first hypothesize how and why working memory is needed. While making a series of trial and error movements without vision of the limb, subjects have to keep in their memory a number or set of prior action-value pairs. Learning requires that the subject avoid repeating unsuccessful movements while repeating those that receive positive feedback. Eventually, it is even more important to put in memory the successful actions in order to maximize the number of rewarded trials, which is the goal of reinforcement-based learning. On the assumption that short-term memory storage is needed for this purpose, it should be possible to assess this using a somatosensory recognition memory task in which subjects are required to judge if a particular movement direction has been experienced before, and to show that performance on this task is related to measures of motor learning.

Although most emphasis has been given to visual and auditory working memory, some studies in both humans and non-human primates have documented working memory in the somatosensory domain[1]. Such studies have largely involved the use of tactile discrimination tasks, for example, using the shape of objects (Stoeckel et al., 2003; Kaas et al., 2006) or using a set of vibratory stimuli (Romo et al., 1999; Harris et al., 2002; Preuschhof et al., 2006). In Romo's experiments, monkeys had to choose which of the two sequentially presented vibratory stimuli was higher in frequency. Some studies have demonstrated aspects

---

[1] Tactile memory is the least-studied topic in the short-term/working memory literature, let alone movement-related short-term memory. See: Gallace A, Spence C (2009) The cognitive and neural correlates of tactile memory. Psychological bulletin 135:380.

of somatosensory working memory with tasks that involve limb displacement. In a joint-position matching task, the participant's arm is passively displaced by the experimenter to a target location and the task is to reproduce the movement to the same location (Chapman et al., 2001; Goble et al., 2006; Jones and Henriques, 2010).

Regardless of the type of stimuli employed, one similarity of such delayed-recognition tasks is that there is an initial (reference) stimulus presented and, after a certain delay, a second (probe) stimulus is presented. Performance of the task requires that the participant hold information in memory concerning the reference stimulus. To understand better human working memory capacity in the somatosensory domain, one can extend the number of items in the memory list and present the probe item after a brief delay. The technique not only allows us to measure the individual working memory capacity, it allows us to understand how memory performance decays as a function of lags, that is, the sequential position of the item in the memory list.

1.6. Neural underpinnings of reinforcement-based motor learning

The neural substrates through which reinforcement learning results in improvements in human motor performance are less well understood. Functional neuroimaging has been an attractive non-invasive method to track neural activity in the brain by looking at the *blood oxygen level dependent* (BOLD) signal. For the past decade, a neuroimaging technique called *resting-state functional MRI* (rs-fMRI) has gained popularity to study brain activity at rest (Biswal et al., 1995). Originally thought to be noise, spontaneous neural activity has been shown to be valuable in understanding the brain in both healthy and clinical populations[2]. Brain regions that display correlated low-frequency BOLD signal spontaneous fluctuations at rest are said to be *functionally connected*. Several functional networks at rest have been

---

[2] For more information about rs-fMRI, refer to a review paper: Power JD, Schlaggar BL, Petersen SE (2014) Studying brain organization via spontaneous fMRI signal. Neuron 84:681-696.

identified and replicated. These have been found to correspond to the task-based networks involving vision, audition, movement, attention, and learning (Xiong et al., 1999; Cordes et al., 2000; Smith et al., 2009). A non task-related network (the so-called *default mode network*) that consistently shows greater activity during rest than during cognitive tasks has also been identified (Raichle et al., 2001; Greicius et al., 2003).

Performing the scans at rest allows us to separate brain activity associated with task performance and learning-related activities. Consequently, this becomes an attractive method to study brain-related changes specifically as a result of learning since because there is no explicit task, any observed neural activity is not due to task-related signal fluctuations but rather presumably reflects learning related changes. Another appeal of rs-fMRI technique is the ability to track memory formation or consolidation measured hours following learning (Sami et al., 2014; Della-Maggiore et al., 2017). In Sami et al., the authors used explicit and implicit serial reaction time tasks (SRTT) involving button pressing. In the explicit task, subjects were told beforehand that the sequence of stimuli would be repeated. Spontaneous brain activity was measured immediately after training, as well as after 30 min, and 6 h after the task performance. It was observed that after 6 h, the implicit group experienced enhanced functional connectivity in the medial temporal region but reduced connectivity within the sensorimotor network. Conversely, the explicit group had enhanced functional connectivity in the sensorimotor network. The persistence of the resting state signal for a period of 6 hrs following training is consistent with the idea that it provides a measure of motor memory formation.

Some studies have used rs-fMRI as a tool to identify the neural substrates of motor learning. In addition to the study by Sami et al (2017) described above, visuomotor training as short as 11 minutes has been shown to alter functional networks involving fronto-parietal areas and cerebellum (Albert et al., 2009). Another previous study identified separate functional networks associated with force-field adaptation (Vahdat et al., 2011). These authors showed that the secondary somatosensory (SII) cortex, ventral premotor cortex, and supplementary motor area (SMA) are related specifically to perceptual changes that occur in conjunction with force-field adaptation. In contrast, areas such as cerebellum, primary motor and dorsal premotor cortex are related to the motor aspects of adaptation.

It is thought that the cerebellum plays a major role in error-based learning or adaptation. Traditionally, the cerebellum is thought to be the putative neural basis for error-based learning which is guided by an error signal encoded in the climbing fiber inputs (Doya, 2000). The involvement of cerebellum in error-based learning has also been reported in neuroimaging studies (Diedrichsen et al., 2005; Smith and Shadmehr, 2005). In order to compute error, a control system has to have a reference target and sensory information related to the movement outcome. However, early in learning in tasks such as described above without visual targets, the sensory target or goal that is needed to compute error is frequently uncertain. It is unknown whether cerebellar networks also subserve motor learning under these conditions or whether cerebellar activity possibly varies in proportion with reward feedback.

Another aspect of early learning that deserves attention is the role of basal ganglia in shaping somatosensory targets through reinforcement. Early insights in reinforcement learning physiologically were derived from animal work. The neural mechanisms of reward influencing behavior were not known until 1954, when Olds and Milner discovered that electrical stimulation in a region around the ventral striatum in rats caused the animals to press a lever more rigorously to obtain a food reward. The structure was later regarded as the reward center and thought to be the interface between the limbic and motor systems (Mogenson et al., 1980). From subsequent anatomical studies, we now know that the ventral striatum is part of a larger dopaminergic network called the *mesolimbic pathways* (Wise and Rompre, 1989; Schultz et al., 1997; Haber, 2003).

The basal ganglia have been shown to be the putative neural substrate for the reward-based learning in animal studies (Schultz et al., 1993; Schultz et al., 1997; Graybiel, 2005). Schultz and colleagues found that the activity of dopaminergic neurons encode not only the current reward but also expected future reward. The role of the basal ganglia in reinforcement learning, as proposed by Doya, is to evaluate the current action and select an appropriate future action (Doya, 2000). But reward-based action may involve more complex neuronal processes beyond the traditional basal ganglia and sensorimotor loops. For example, it is possible that reward-based decision-making is also involved (Rushworth et al., 2004), such that, during learning, rewards may influence the production of subsequent movements. Prior studies discovered that regions in the prefrontal cortex are involved in this type of activity (Shima and Tanji, 1998; Tremblay and Schultz, 2000). Using fMRI in gambling tasks (Daw et al., 2006), intraparietal sulcus and lateral frontopolar cortex are preferentially active during exploration than exploitation. In contrast, ventral striatum and ventromedial prefrontal cortex are involved in exploitative decision making to accumulate more rewards. The latter is also shown to encode the value of either performed actions or more abstract decision making (Gläscher et al., 2008).

1.7. Processing resting-state data

As mentioned in earlier, rs-fMRI does not involve task performance inside the scanner as in the case of task-based fMRI. Rather rs-fMRI data are analyzed to give a set of temporal correlations that defines a functional network. While we can use a predefined *haemodynamic response function* to model neural activity in response to a stimulus or other events (Friston et al., 1994), the BOLD signal at rest does not possess a known temporal pattern. Therefore, the greatest challenge in analyzing data in rs-fMRI is the effort to separate neural signal (BOLD) from the signals that are noise or artifact. A great deal of effort has gone into understanding and identifying the main sources of artifactual signals in order to clean (denoise) them. A valuable study by Cordes et al suggests that meaningful spontaneous neural activity is limited to low frequency fluctuations. Such fluctuations are found to be lower in frequency than 0.1 Hz (Cordes et al., 2001). This frequency value provides an upper limit of a range of interest within which there is information related to neural activity in resting-state data.

Sudden head motion in the scanner is one source of artifact that has been shown to reduce the correlation value of two brain regions that are supposed to be functionally connected, or to introduce a spurious and fictitious correlation between any brain regions (Power et al., 2012; Van Dijk et al., 2012). Head motion artifacts generally occur along the edges of the brain and to some extent tissue boundaries inside the brain (Jo et al., 2010). In order to correct for head motion related distortion, it is necessary to first correct for volume-wise displacements by assuming that the head is a rigid body and then to model the movements in six degrees-of-freedom (Friston et al., 1996; Jenkinson et al., 2002). Subsequently, volume-by-volume realignment can be performed with respect to a reference volume (e.g. the first volume or the middle volume) once the direction in which movements occurred is known. These displacement parameters are used as unwanted or *nuisance* regressors in a general linear model (GLM) to regress out unwanted distortion from the BOLD signal. The latter is performed because temporal distortion introduced by the head motion has to be modeled and compensated. Another popular technique is called *motion scrubbing* where drastic and sudden changes over time due to head motion are either removed from the analysis or ignored in the GLM (Power et al., 2012; Carp, 2013).

A second source of artifacts is physiologically-related movement due to heartbeat and respiration (Birn et al., 2006; Murphy et al., 2013). Cardiac pulsation is usually seen in brain areas along large arteries and sinuses (Glover et al., 2000) and respiratory effects can introduce slow head and thoracic rhythmic movements (Friston et al., 1996; Raj et al., 2001). Theoretically, physiological fluctuations introduced by respiration (~0.5 Hz) and heartbeat (~1.2 Hz) are well above the range of resting-state BOLD fluctuation (Cordes et al., 2001).

Correcting for physiological noise, however, can be difficult. Most MRI sequences operate with a repetition time (TR) of about 2 seconds, giving rise to a sampling frequency rate of 0.5 Hz. The sampling theorem requires that in order to fully capture the frequency composition of a signal, we sample at least twice the maximum frequency component of the signal. Since the highest frequency component in resting-state data originates from the heartbeat (~1.2 Hz), the ideal sampling rate should be at least 2.4 Hz. This value is equivalent to a TR of 1/(2.4) = 400 msec. Unless we employed an ultra-fast functional imaging sequence (e.g. multiband accelerated sequence), we are unable to fully capture this information.

An alternative strategy is to denoise resting-state data based on an approach called *Independent Component Analysis* (ICA) (McKeown et al., 2003). This approach can be used in both task-based and resting-state fMRI datasets to separate signal-related neural activity from physiological noise and motion artifact. In ICA, the goal is to segregate the fMRI dataset into a set of spatially independent brain maps, each with its own average time courses. Once independent components have been identified, the task is to systematically classify which of the components represent noise. Typically, classification is done manually by a visual inspection (Kelly et al., 2010). Cerebrospinal fluid (CSF) is found in the ventricular system of the brain and is closely related to the blood flow, carrying a frequency spectrum within the range of the heart rate (Bhadelia et al., 1997; Balédent and Idy-peretti, 2001). More recently, efforts have been spent to design a more sophisticated classification method using a *supervised machine learning* algorithm (Salimi-Khorshidi et al., 2014; Pruim et al., 2015)..

In this thesis ICA denoising strategy with manual classification (Kelly et al., 2010) will be used to clean resting state data. Noise components that have been identified will then be regressed out from the time series. A sample of the identified noise components will be given in the supplementary materials of Chapter 2.

1.8. Concluding remarks

Initial stages of learning are often characterized by the lack of known sensory targets or references. In cases where visual or auditory targets of movements are not available, it is necessary to determine sensory targets wholly in somatosensory terms. Discovery of such targets is done through a series of exploratory movements where learning proceeds through the use of positive feedback or reinforcement resulting from one’s own movement outcomes.

In Chapter 2, I will present the results of a study that identifies plasticity associated with reinforcement-based motor learning. Throughout this thesis, I use arm-reaching tasks that mimic the actual scenario of learning where the visual targets are lacking and exploratory behavior is necessary. Chapter 3 reports the results of two experiments that investigate the role of somatosensory working memory in motor learning. The study begins with the first experiment which involves separate working memory and motor learning tasks, conducted on two separate days. This design specifically aims to understand the capacity of somatosensory working memory in humans and how that relates with the change in motor performance following learning. The experiment also aims to examine if the working memory in the context of our learning task is specific to the somatosensory domain. In contrast, a working memory test in the second experiment is interleaved with motor learning trials and directly tests for subject’s memory of their own reaching movements. The immediate role of working memory during learning can then be assessed. A final chapter provides a summary of findings that also includes potential limitations of the current work.

Chapter 2 : Training-induced plasticity in the initial stages of motor learning

2.1. Preface

The study described in the present chapter combines a reinforcement-based motor learning task with resting-state fMRI to investigate the behavioral and neural substrates of the initial stages of motor learning. The behavioral task involves reaching movements to a target that is hidden from view, and the delivery of positive feedback when the movement ends within the target zone. The task goal is to reduce the lateral deviation with respect to a specific reaching direction. An earlier study by Bernardi and colleagues demonstrated the role of both positive feedback, which provides reinforcement, and somatosensory information in this kind of learning (Bernardi et al., 2015). This chapter extends and elaborates on their findings in two important ways. First, I assess how motor behavior is dependent in a trial-to-trial manner on unsuccessful or successful (reinforced or non-reinforced) movements. Second, I examine the neural signatures underlying these processes using resting-state functional MRI (rs-fMRI).

In the present motor learning paradigm, a progressively smaller and hence more difficult target size (the reward zone) is employed as a form of behavioral shaping. The width of the target zone is progressively reduced from one block to the next while maintaining the position of the target center. Behavioral shaping is not foreign to reinforcement learning and has been used in two recent reinforcement-based motor learning studies. In a study by Pekny et al., the goal was to understand whether subjects with Parkinson’s Disease were able to adjust trial-to-trial variability during learning. For this reason, the size of the reward zone in the first half of the experiment was fixed but its center position was rotated 4 degrees after every 100 trials in either direction unbeknownst to the subjects (Pekny et al., 2015). A similar approach was also adopted by Therrien and colleagues in a visuomotor reaching task with a so-called "open-loop" and "closed-loop" reinforcement schedule. In the open loop paradigm, the center of the reward zone was gradually shifted 15 degrees with a degree step size for every 20 trials. In contrast, the shift in the close-loop group was calculated based on the moving average of the reach angle over the past 10 trials (Therrien et al., 2015).

Behavioral analyses described in this chapter include the effect of positive feedback on subsequent movements, which is investigated here using trial-to-trial change in reaching direction as a proxy. It is found that following each non-reinforced trial, the change in reaching direction between the current and immediately subsequent movement is significantly higher than the change in reaching direction following reinforced trials. The findings suggest a form of motor exploration in the absence of reward. Moreover, exploration increased proportionally with an accumulation of unrewarded trials and decreased proportionately with an accumulation of rewarded trials.

In this study, each participant undergoes resting-state scans before and after motor learning. In order to systematically identify learning-specific changes in functional connectivity, I include behavioral measures of learning as covariates in group level analyses. Increases in connectivity occur in primary motor and somatosensory cortices, as well as supplementary motor area. In line with the earlier literature in reinforcement and reward-based learning, people that receive more positive feedback have larger increases in functional connectivity in a set of areas comprising ventromedial prefrontal cortex and ventral striatum. The involvement of ventromedial prefrontal cortex here supports the idea that reward-guided decision-making occurs in conjunction with reinforcement-based motor learning. Plasticity associated with the overall exploratory behavior is identified to include primary sensorimotor cortices, SII, and putamen. Changes in these networks vary systematically with the average trial-to-trial change in movement direction, regardless whether the previous trial is successful or unsuccessful. Another related plasticity is also observed in a fronto-parietal network comprising ventrolateral prefrontal cortex and supramarginal gyrus, where connectivity with second somatosensory cortex varies with the extent of change in movement direction following an incorrect movement. This latter finding is consistent with the involvement of both second somatosensory cortex and ventrolateral prefrontal cortex in somatosensory memory and decision-making tasks.

Sidarta A., Vahdat S., Bernardi NF., Ostry DJ (2016). “Somatic and reinforcement-based plasticity in the initial stages of human motor learning”. *J Neurosci.* 36: 11682-11692.

2.2. Abstract

As one learns to dance or play tennis, the desired somatosensory state is typically unknown. Trial and error is important as motor behavior is shaped by successful and unsuccessful movements. As an experimental model, we designed a task in which participants make reaching movements to a hidden target and receive positive reinforcement when successful. We identified somatic and reinforcement-based sources of plasticity on the basis of changes in functional connectivity using resting-state fMRI before and after learning. The neuroimaging data revealed reinforcement-related changes in both motor and somatosensory brain areas in which a strengthening of connectivity was related to the amount of positive reinforcement during learning. Areas of prefrontal cortex were similarly altered in relation to reinforcement, with connectivity between sensorimotor areas of putamen and the reward-related ventromedial prefrontal cortex strengthened in relation to the amount of successful feedback received. In other analyses, we assessed connectivity related to changes in movement direction between trials, a type of variability that presumably reflects exploratory strategies during learning. We found that connectivity in a network linking motor and somatosensory cortices increased with trial-to-trial changes in direction. Connectivity varied as well with the change in movement direction following incorrect movements. Here the changes were observed in a somatic memory and decision-making network involving ventrolateral prefrontal cortex and second somatosensory cortex. Our results point to the idea that the initial stages of motor learning are not wholly motor, but rather involve plasticity in somatic and prefrontal networks related both to reward and exploration.

Acknowledgments:

This work was supported by the National Institute of Child Health and Human Development R01 HD075740, Les Fonds Québécois de la Recherche sur la Nature et les Technologies, Québec (FQRNT), and the Natural Sciences and Engineering Research Council of Canada (NSERC). We thank Dimitrios Palidis for his assistance in conducting the experiments and Dr. Floris VanVugt for his valuable feedback.

Disclosures:

No conflicts of interest, financial or otherwise, are declared by the authors.

Significance statement

In the initial stages of motor learning, the placement of the limbs is learned primarily through trial and error. In an experimental analog, participants make reaching movements to a hidden target and receive positive feedback when successful. We identified sources of plasticity based on changes in functional connectivity using resting-state fMRI. The main finding is that there is a strengthening of connectivity between reward-related prefrontal areas and sensorimotor areas in the basal ganglia and frontal cortex. There is also a strengthening of connectivity related to movement exploration in sensorimotor circuits involved in somatic memory and decision making. The results indicate that initial stages of motor learning depend on plasticity in somatic and prefrontal networks related to reward and exploration.

2.3. Introduction

One of the challenges a beginner faces in learning a golf swing or a tennis serve is that the desired somatosensory state is initially unknown. Of necessity learning proceeds largely by trial and error and involves a process in which the acquisition of motor commands and the development of somatic targets occur in parallel. The functional brain networks that subserve this stage of learning are largely unknown, and constitute the focus of the present investigation. Much of the current literature on motor learning focuses on adaptation paradigms, in which some form of perturbation impairs movement to well learned sensory targets. A prominent feature in adaptation is the introduction of systematic error followed by a progressive reduction of this error through adjustments to motor commands. Accordingly, neuroimaging studies investigating motor adaptation have highlighted the role of areas such as cerebellum as a key node for error correction (Diedrichsen et al., 2005), and of the posterior parietal cortex, involved in the sensorimotor transformations that are necessary during adaptation to re-plan spatially guided movements (Bernier and Grafton, 2010).

The role that these previously identified networks may play in situations in which sensorimotor targets have to be acquired in the first place is unknown. In such a situation it is possible to hypothesize that a different set of brain areas will show changes in conjunction with motor learning. First, recent behavioral work points to the importance of the somatosensory system for this kind of task. For example, using an experimental manipulation similar to the one employed here, Bernardi et al. (2015) showed that somatosensory experience delivered through passive movements generated learning comparable to that seen in participants trained with active movements. Second, the process of skill acquisition can be memory-dependent in the sense that one must be able to repeat the correct or successful actions and avoid previously incorrect movements. Accordingly, one might expect the recruitment of somatic memory and decision making circuits in this sort of learning (Romo et al., 1999; Romo et al., 2002), and more generally the prefrontal cortex (Miller and Cohen, 2001). Finally, the involvement of reinforcement-related brain networks would be expected, as positive feedback may effectively shape learning and compensate for the lack of detailed error information in the early stages of learning. Previous behavioral studies have shown the contribution of reinforcement to motor learning in tasks such as those involving arm reaching (Izawa and Shadmehr, 2011; Shmuelof et al., 2012; Manley et al., 2014), saccadic eye movement (Takikawa et al., 2002; Madelain et al., 2011), and precision gripping (Dayan et al., 2014). Brain networks that support reinforcement and reward-based learning in general have been studied and comprise, among others, the ventromedial prefrontal cortex and striatum (Schultz et al., 2000; Berns et al., 2001; O'Doherty et al., 2004; Haruno and Kawato, 2006; Bischoff-Grethe et al., 2009).

In the present study, we examined changes in functional connectivity in resting-state brain networks that occurred following movements to a small unseen target. When the movement landed within the target zone, positive feedback was provided to indicate success. This task was designed as an analogue to the early stages of learning a novel motor skill, for which reinforcement-based selection of the sensory targets is central, rather than error-based adjustments of the motor commands. We found that training resulted in improvements in movement that were accompanied by changes in functional connectivity in both reinforcement-related networks and those related to memory and decision making. The results point to the idea that reward-related prefrontal regions contribute to the early stages of learning in sensorimotor circuits. Somatic memory and decision-making networks support movement variability and presumably exploration.

2.4. Materials and methods

Experimental Setup

A total of 22 right-handed participants were recruited (14 females, mean age = 22.5 years old, SD = 3.19) and provided written consent. All procedures were approved by the McGill University Institutional Review Board. The participants were healthy adults with no prior physical or neurological conditions. The experimental session for each participant was completed within the same day.

The behavioral paradigm in this study was based on that used by Bernardi et al. (2015). Briefly, we used a two degree-of-freedom robotic manipulandum (Interactive Motion Technologies), with a vertical handle attached to the end-effector. The handle position was provided by a set of 16-bit optical encoders (Gurley Precision Instruments). Participants were seated in front of the robot with their right shoulder abducted to about 70° and the elbow supported by an air sled. A semi-silvered mirror, which served as a display screen, was placed just below eye level and blocked the vision of the arm and the robot handle. A green circle, 20 mm in diameter, was positioned on the display screen along the subject's body midline, and was used as the start position of each movement. To the left, a 1 cm thick target stripe, tilted at 45° with respect to the horizontal, extended the entire width of the display screen (Figure 1A). Within this bar, there was an unseen rectangular target zone, the center of which was located 15 cm from the center of the start circle. A thin yellow line served as a visual cue to indicate the distance of the hand from the target stripe. A small 12 mm diameter yellow circle attached to the yellow line corresponded to hand position. This circle was shown briefly at the beginning of each movement and disappeared as soon as the subject left the start position. No information about the lateral displacement of the hand was provided during movement or at movement end.

Experimental Blocks

Participants were first given a set of instructions about the experiment followed by 15 familiarization trials. They were told to perform outward reaching movements, 45° to the left of the midline, until they reached the target stripe. They were told that the trajectory had to be as straight as possible with no corrective movements throughout. Each trial had to be completed within 800 msec, and participants received feedback about their speed by means of a target color change (red, green and blue corresponding to too fast, correct speed and too slow, respectively). However, there was no penalty if the movement did not end on time. Once the movement ended, the robot would bring the arm back to the start position.

The experiment began with a block of 15 baseline trials in which participants performed reaching movements toward the target stripe. No feedback was provided as to whether the movements were accurate. Following this, they proceeded to the Brain Imaging Centre at the Montreal Neurological Institute (MNI) for a first scanning session. This neuroimaging session comprised two resting-state state scans with eyes closed, followed by a gradient field map acquisition and a T1-weighted scan. A more detailed description of the functional imaging procedures is presented below.

After the initial scanning session, participants returned to the laboratory and completed four training blocks of 50 trials each. They were told that this was the opportunity to learn which movement to the unseen target was successful. A particular movement was defined as successful when the trial ended within the hidden target zone. The success was determined based on the lateral dimension of the movement endpoint, not the movement speed. Following a successful trial, an animated explosion and the words “Nice shot!” appeared on the screen to provide positive feedback. The participant was told to pay attention to the experience of moving to the target correctly and to collect as much positive feedback as possible. The feedback was binary, that is, no information about error magnitude or direction was given for movements that ended outside the target zone.

In order to facilitate learning progressively, the width of the target zone (W) was changed over the course of training, keeping the center position fixed (Figure 1B). We adopted this progressive level of difficulty as a form of behavioral shaping (Skinner, 1965; Darshan et al., 2014). In the first training block, the width of the target zone was calculated as the lateral range within which 50% of the baseline movements ended. In the second block, the width was set to half the distance between the first and the last target width. A final target width of 8 mm was used for the remaining two training blocks and was the same for all participants. A short break was given between successive training blocks.

After the completion of the final training block, the participants were brought to the imaging center for a second series of fMRI scans. The scans consisted of two resting state scans, a gradient field map acquisition, a T1-weighted structural scan, and a task-based movement localizer that will be described below. Following these scans, the participant again returned to the laboratory to perform 15 movements without any feedback. The last block served to evaluate motor performance following learning.

Data Analysis

Motor performance was quantified at movement end based on the unsigned magnitude of the lateral perpendicular deviation, $|PD|$, with respect to a straight line connecting the center of the start position and the center of the target zone (Figure 1C). Movements that ended closer to the center had smaller $|PD|$ scores. For each subject, the average $|PD|$ before (PRE) and after (POST) training was calculated using the 15 trials without feedback, and the difference served as a measure of the participant's improvement in accuracy, $\left(\overline{|PD|}\right)_{\mathrm{PRE}} - \left(\overline{|PD|}\right)_{\mathrm{POST}}$, with larger positive values corresponding to greater learning. We also assessed the training-related performance in terms of the number of successful trials on which feedback was presented. To check the linear dependency between the improvement in accuracy and the overall number of successful trials, we computed Pearson's correlation coefficient between the improvement in movement accuracy from PRE to POST with the total number of successful trials in all training blocks.

Trial and error in search for the correct movement trajectory is presumably important for learning. To see how the feedback or its absence influenced the movement on the following trial, we assessed how trial-to-trial movement direction changed after every successful trial (S = 1) and every unsuccessful trial (S = 0). We quantified this with $\Delta m_n = |PD_{n+1} - PD_n|$, which signifies the difference in PD between trial n and $n + 1$, contingent upon trial n being successful or unsuccessful. For each subject we first computed the mean Δm in these two conditions, and then used the set of means in each condition to estimate the group mean and variability of the sampling distribution. We tested whether the average Δm was different following successful and unsuccessful trials.

MRI Acquisition

MRI data were acquired at the MNI using a 3.0 T MRI scanner (Tim Trio, Siemens). To reduce head motion and scanner noise, foam padding and earplugs were provided to the participants. During resting state scans, each participant was instructed to lie quietly with their eyes closed and avoid any head motion during the scan.

Functional images were obtained using the Simultaneous Multi-Slice BOLD-EPI WIP sequence (Setsompop et al., 2012) as follows: slice acceleration factor = 3x; repetition time (TR) = 1690 msec; echo time (TE) = 25 msec; slices = 63; thickness = 2 mm (no gap); field of view (FOV) = 200 mm × 200 mm; and flip angle (FA) = 90°. Each functional scan lasted for ~7 min and yielded 250 volumes. Two scans were performed before and after training respectively. We acquired two 7 min resting-state runs, rather than a single continuous 14 min scan, for the practical reason that it keeps subjects from falling asleep. Structural images were acquired with a T1-weighted 3D MPRAGE sequence as follows: TR = 2300 msec; TE = 2.98 msec; slices = 192; thickness = 1 mm (no gap); FA = 90°; and FOV = 256 mm × 256 mm, iPAT mode = ON (GRAPPA, acceleration 2x). We used a multiband accelerated imaging sequence in the current studies because we could acquire more data in a relatively short scan time (Moeller et al., 2010). Simultaneous acquisition was achieved using a 32-channel multiarray head coil.

fMRI Data Preprocessing and ICA

Data preprocessing was carried out using FSL v6.0 software packages, www.fmrib.ox.ac.uk, FMRIB, Oxford U.K (Smith et al., 2004). Briefly, image preprocessing consisted of: the removal of the first three volumes in each scan, non-brain removal using BET, motion correction (using a six-parameter affine transformation implemented in FLIRT), spatial smoothing with Gaussian kernel of FWHM 5 mm, and temporal high-pass filtering (Gaussian-weighted least-squares straight line fitting, $\sigma = 100.0$ sec). The boundary-based registration with fieldmap correction aligned the subject's functional image to the subject's structural space (Greve and Fischl, 2009) and the 12 DOF non-linear registration using FNIRT normalized the structural space to the standard MNI152, 2 mm template.

Noise artifacts in the individual datasets were identified using Independent Component Analysis (ICA) in FSL - MELODIC (Beckmann and Smith). There is presently no consensus on the optimal number of components for the noise removal. For our present application, the ICA dimension was determined automatically by the software. On average, the total number of independent components ranged from 45 – 60. From this, components associated with the physiological noise, signal dropout, and sudden head motions, were identified by visual inspection following the guidelines by Kelly et al. (Kelly et al., 2010). The number of components classified as noise and then removed was about ten percent of the total. We found that removing additional components did not yield further changes to the group statistical map.

ROI Identification

Using seed-based analysis, we assessed the temporal correlation of specific brain regions of interest (ROI) with all other voxels in the brain. ROI locations were identified using a task-based localizer fMRI (Vahdat et al., 2011). Briefly, the task involved movement of the right arm with six alternate blocks of movement and rest, each lasting for 30 sec. The movement speed was 1/3 Hz and was paced by visually presented stimuli. During the rest block, the participant remained still.

Subject-level statistical analyses of the localizer task were carried out using the FEAT toolbox in FSL (Beckmann et al., 2003). Here, the block design was convolved with the haemodynamic response function as the main predictor in the linear model. After this analysis was completed for each participant separately, a group-level mixed-effect model analysis (FLAME) was carried out using the same toolbox. The statistical map was subsequently thresholded using $Z = 4.0$ and $p < 0.01$, corrected for multiple comparisons. This map identified regions in the brain that were on average activated across subjects during the task. The map was then used to identify seed locations in the MNI coordinates. Each ROI was represented as a spherical mask of 5 mm radius around the local maximum.

A list of ROIs employed in this study with their corresponding MNI coordinates and the Z-value of the local maximum can be found in Table 1. Briefly, seeds were placed in the primary motor and somatosensory cortices (M1 and S1), the dorsal premotor cortex (PMd), the supplementary motor area (SMA), and the second somatosensory cortex (SII) in the parietal operculum (Vahdat et al., 2014). One seed was placed in the cerebellar lobule V and another seed in the motor region of the left putamen, all of which corresponded to the local maxima as identified by the localizer task.

Seed-based Analysis with Behavioral Factors

Analysis of the resting-state fMRI data was performed using a seed-based approach.

We first obtained the time series of the nuisance components using the ICA process described above. Additionally, to account for further potential artifacts, the average signals within the white matter, the ventricles, and the whole-brain mask were regressed out in the present analysis (Desjardins et al., 2001). To do so, white matter and ventricles were first segmented using FSL – FAST before being mapped into the subject’s native functional space. To increase tissue precision, both images were thresholded using a tissue probability of 90%. We then used the resulting image as a mask to extract the average time series inside the white matter and ventricles.

To extract the temporal correlation between a seed and other brain regions, a multiple regression analysis was carried out using FEAT. Specifically, the ROI time series was the main predictor of interest, while the average time series of white matter, ventricles, global signal, the nuisance components obtained from ICA earlier, and six motion parameters were regressed out from the whole-brain time series. The results were brain regions that were temporally correlated with the seed after regressing out unwanted temporal noise. We repeated this step for all seeds on every run of each subject.

After this stage was completed, a group level repeated-measures *t*-test was carried out for each seed using a mixed-effect model (FLAME) package in FSL. The design matrix consisted of a series of explanatory variables or predictors. The first set explained the subject average or common effect among different runs. The second set comprised a behavioral factor with the aim of finding differences that were associated with our behavioral manipulation (Vahdat et al., 2011). Specifically, we examined changes in functional connectivity (ΔFC) in relation to the number of trials with positive feedback during training. Only successful trials in the last two training blocks were used for this analysis because the width of the target was the same for all subjects. An analysis of connectivity changes related to improvements in movement accuracy was also conducted. The patterns were similar to those reported below for connectivity changes related to successful movements. These are not presented separately because subject feedback during training was restricted to binary feedback on movement success, hence we considered positive feedback as the main factor determining the increase in accuracy in the post-training session thereafter. Moreover, the two behavioral measures of reinforcement and accuracy were significantly correlated (see below).

In a second set of group level analyses, we examined changes in functional connectivity (ΔFC) related to trial-to-trial changes in movement direction. For each subject, we averaged the changes in movement direction (Δm) regardless of the trial outcome and applied this as the behavioral predictor. In a subsequent analysis, we examined ΔFC that were uniquely attributed to the change in movement following either successful or unsuccessful trials. Here, we separately averaged Δm following only successful ($S = 1$) and only unsuccessful trials ($S = 0$). We put these two sets of values as the predictors within one general linear model to determine changes that were uniquely explained by one factor independent of the variability shared with the other factor.

For both group analyses, a correction for multiple comparisons was carried out using Gaussian random field theory using a cluster forming threshold $Z = 2.40$ with $p < 0.05$. Two different contrasts were evaluated, i.e. POST > PRE and POST < PRE, to test for increases or decreases in functional connectivity following training. The thresholded group statistical maps of each seed revealed clusters whose changes in connectivity with the seed region were reliably associated with the corresponding behavioral predictor. To correct for multiple seeds (i.e. Bonferroni correction for choosing seven seeds), clusters obtained from the group-level analyses were considered to be significant if the probability level was lower than $p < 0.05/7$.

The whole-brain global signal in the resting-state data is usually included as one of the unwanted components. However, the removal of the global signal has been controversial as it introduces a negative bias to the resting-state statistical map (Saad et al., 2012). Because we computed the difference between the PRE and POST training scans, this negative bias did not affect the difference maps presented below. In order to quantify the strength of the functional connectivity measure in each scan before and after training trials, we repeated the same analysis but without removing the global signal time series. The results of the group analyses without the global signal removal yielded similar statistical maps.

2.5. Results

Behavioral Performance

Figure 2A depicts movement accuracy as quantified using the absolute lateral deviation at the end of movement during the baseline test (PRE), the four training blocks, and the motor evaluation block (POST). The reduction in the mean $|PD|$ over 15 movements between the PRE and POST training blocks provides a measure of how accuracy improved as a result of training. On average, the reduction was found to be significant ($t_{(21)} = 2.080$, $p < 0.05$) and reliably correlated with the total number of successful trials over the course of training ($r = 0.44$, $p < 0.05$) (Figure 2B). Participants that achieved a greater number of successful trials had a tendency to display a greater improvement in movement accuracy.

We gave the participants the opportunity to improve their movement accuracy with four training blocks during which they received positive feedback if the movement ended within the target zone and no feedback otherwise. Three target zones that gradually decreased in width were incorporated during training trials in order to progressively shape subjects' behavior. During the first and second training blocks, the percentage of success was in the range of 70 – 80%. When the target width was reduced to the smallest, the percentage decreased to 30 – 40 % but nevertheless increased over the course of training (inset of Figure 2A). We checked the relationship between subjects' performance in the first two blocks, in which the width of the target varied and that in the last two blocks, in which target width was fixed (8 mm). We found that subjects that had more successful movements during the first two blocks did so as well in the last two blocks ($r = 0.46$, $p < 0.05$). The fact that the target was smaller in the last two blocks likely contributed to the slowing of learning seen in the third and final block (Zone III) of Figure 2A.

In order to assess the effect of feedback on subsequent movements, we calculated Δm as the absolute change in *PD* between the current and next immediate movement. We used training data from blocks 3 and 4 for this calculation since the target size was uniform across subjects. Figures 2C illustrates the distributions of Δm following successful and unsuccessful trials as a half-normal Gaussian curve. The figure shows that the average Δm after successful trials is significantly less than the average Δm after failed trials ($t_{(21)} = 3.988$, $p < 0.001$). In other words, failing to get positive feedback resulted in a greater trial-to-trial change in movement direction, presumably in search of the correct target zone. The average Δm after successful and unsuccessful trials were linearly related ($r = 0.53$, $p < 0.05$). In addition, subjects who had a greater number of successful movements, and hence received more positive feedback, displayed a smaller change in movement direction following both successful ($r = –0.72$, $p < 0.001$) and unsuccessful movements ($r = –0.58$, $p < 0.005$).

Furthermore, again restricting the analysis to the data from the final two training blocks, we assessed the trial-to-trial change in movement direction (Δm) as a function of the number of consecutive successful trials and of the number of consecutive failed trials. A weighted least squares regression was calculated to predict these relationships (Figure 2D). We found, using a simple linear mixed model, that the average Δm increased with the number of failed trials since the last successful movement [$F(1,17.13) = 6.97$, $p < 0.05$]. On the other hand, there was a reduction in average Δm when preceding movements were successful [$F(1,10.64) = 10.81$, $p < 0.01$]. Thus movement variability, and presumably exploration, progressively increased following unsuccessful movements and decreased following successful movements (Sutton and Barto, 1998).

Selection of Regions of Interest (Seeds)

We assessed changes in functional connectivity (ΔFC) associated with the number of successful trials during learning using a seed-based approach. We identified seven seed locations based on the local maxima in the group level task-based localizer data. The seeds regions were located in the left M1 (primary motor cortex, BA4), left S1 (primary somatosensory cortex, BA2), left dorsal premotor cortex (PMd, BA6), the supplementary motor area (SMA), left second somatosensory cortex SII (parietal operculum, OP1), right cerebellar lobule V (CbV), and the left rostral motor area of putamen (Pu). The seed location of putamen in this study is restricted to the motor region as defined by the Oxford-GSK-Imanova Striatal Atlas (Tziortzi et al., 2014). The MNI coordinates of each seed region along with its corresponding *Z*-score are listed in Table 1.

Changes in Functional Connectivity Related to Training Performance

In order to identify ΔFC associated with the behavioral manipulation, we included as the predictor the number of successful trials in the last two training blocks which had the same target size for all subjects. Figure 3 shows ΔFC that are significantly correlated with the number of successful trials. The seed regions are given in green and, to the right, are those clusters of voxels for which the correlation with the seed region changed in proportion to the number of successes. The scatterplots depict the relationship between the change in functional connectivity and individual differences in the behavioral performance. Additionally, Table 2 provides the list of clusters that show change in connectivity with the individual seed and the coordinate of the maximum *Z*-value in the clusters. The cluster *p*-

value is significant when it is less than 0.05/7 (corrected for multiple seed selection). Functional connectivity measure (strength) before and after training is given by the average Z-score of correlation between the ROI time series and the time series of the corresponding cluster, with a negative value indicates an anti-correlation.

Changes in functional connectivity related to the number of successful trials were observed with seeds in the left M1 and PMd (Figure 3). This measure was associated with increases in functional connectivity in a network comprising M1, PMd, S1 and SMA. The positive correlation indicates that subjects that achieved more successful trials had higher connectivity strength following training. A similar trend was observed in the connectivity between the seed in the left SII and S1. An increase in connectivity strength between SII and S1 was positively correlated with the number of successful trials.

The number of successful trials also predicted both increased and decreased connectivity with the putamen (Pu) seed. An increase in functional connectivity was found with ventral medial prefrontal cortex (vmPFC) that extends to a portion of the ventral striatum. On the other hand, we found a reduction in the functional strength with somatic areas comprising a region in the parietal operculum (SII) and S1 that extends to the anterior intraparietal sulcus. We further tested this observation and found that the increase in connectivity between Pu and vmPFC was strongly correlated with the decrease in connectivity between Pu and S1/SII ($r = –0.58$, $p < 0.01$). This suggests that subjects who were more successful during training, and thus received more positive feedback, had stronger connectivity involving vmPFC, but reduced connectivity with the somatosensory areas of the brain.

Changes in functional connectivity related to feedback-dependent changes in movement direction

Figure 4 depicts the results of a second set of analyses, focusing on functional networks related to trial-to-trial movement direction changes (Δm). We first analyzed ΔFC associated with Δm, regardless of the trial outcome. We then proceeded to segregate networks involved in the repetition of successful movements and those presumably involved in exploration when the preceding movements were unsuccessful. In this case, both factors were included in a single general linear model, enabling us to identify brain areas that were associated with each predictor separately after removing changes in connectivity that were related to the other variable.

In the top panel of Figure 4, the seed regions are shown in green and, to the right, we show the voxels whose correlation with the seed region is dependent on change in the movement direction regardless of whether the preceding movement was successful or unsuccessful. Table 3 summarizes the connectivity measure (strength) between the individual seed locations and the corresponding clusters before and after training. It is seen that connectivity between SII and sensorimotor areas is strengthened as a result of training but the connectivity between the SMA seed and two subcortical clusters is reduced. The clusters were found to be bilateral but with a statistical peak in the left putamen and left thalamus respectively (Table 3, top). Subjects with smaller trial-to-trial changes in direction had greater SMA – putamen connectivity.

Significant changes in connectivity which were dependent on whether the preceding trial was successful or unsuccessful were restricted to movements following unsuccessful trials (Figure 4, lower panel). It was found that change in movement direction after unsuccessful trials predicts the decrease in connectivity between SII and two areas in the right hemisphere. The first area is BA 9/46 in the lateral prefrontal cortex just above the inferior frontal sulcus, and another area is supramarginal gyrus (SMG). The correlation is found to be negative, that is, subjects who explored the space more widely following unsuccessful trials experienced a greater reduction in functional strength. We did not observe any reliable correlation with the left prefrontal region.

The connectivity between M1 and posterior intraparietal sulcus (pIPS) was also found to increase in proportion to the change in direction following unsuccessful trials. The cluster with increased connectivity covers the parieto-occipital border and extends to posterior angular gyrus (Area PGp). A seed placed in SMA shared a similar pattern of change in connectivity with pIPS and angular gyrus. The positive correlation observed here implies that stronger functional interaction between the two regions is associated with a greater change in movement direction following unsuccessful movements. It is noteworthy that there is no direct anatomical connection between pIPS and M1 in macaques. However, the observed functional connectivity between pIPS and M1 might be supported through the dorsal premotor area, which is directly connected with both M1 and pIPS (Tanné-Gariépy et al., 2002).

No reliable changes were observed in connectivity that was uniquely associated with the change in movement direction following successful trials. This might be due to the fact that change in movement direction is substantially less following successful movements and differs little between subjects (refer to Figure 2C). The absence of significant ΔFC under these circumstances thus likely resulted from a lack of variability in the behavioral predictor.

2.6. Discussion

The motivation for this study was to identify changes in functional networks of the brain that are associated with learning sensorimotor targets in the initial stages of human motor learning. In order to focus on somatic target acquisition in the early stages of learning a novel motor skill, we used movements that were already part of the individual’s motor repertoire in combination with target locations that were initially unknown. The task was designed to allow trial and error in search for the correct limb position and to provide positive feedback as reinforcement during training. The results suggest that the initial stages of motor learning are to be understood as not entirely motoric. Evidence of plasticity was obtained in somatic networks that are related to exploration, and also in prefrontal areas, related to reinforcement.

In behavioral terms, we found that on average performance improved compared to baseline. The extent of the improvement varied in proportion to the number of successful training trials, with subjects that were more successful during training having the greatest improvements in movement accuracy.

We used resting state fMRI to elucidate changes in connectivity in relation to success during learning. We found that learning changed the functional connectivity both in cortical sensory and motor areas of the brain. Participants that had a greater number of successful trials showed larger increases in functional connectivity in a network comprising the left M1, S1, SMA, and PMd. The finding is consistent with previous resting-state imaging work involving both sensorimotor adaptation and somatosensory perceptual learning (Albert et al., 2009; Vahdat et al., 2011; Vahdat et al., 2014). The participation of these same motor regions in reward-related tasks has been observed in prior studies in both humans (Ramnani and Miall, 2003; Kapogiannis et al., 2008) and monkeys (Roesch and Olson, 2003; Sul et al., 2011).

Areas in the prefrontal cortex not typically associated with motor learning were likewise involved and showed a contribution which varied across subjects in a manner related to their behavioral performance. Specifically, changes in connectivity were observed between the putamen and the ventromedial prefrontal cortex (vmPFC) that were related to the number of successful trials. The vmPFC is a region in which activity is associated with stimulus-reward value, selecting actions that are more rewarding (O'Doherty et al., 2003; Rushworth et al., 2004; Daw et al., 2006) and encoding the value of performed decisions (Knutson et al., 2001; Smith et al., 2010).

We observed that across participants, the increase in connectivity between Pu and vmPFC was accompanied by a reduction in connectivity with the primary and secondary somatosensory regions. This suggests that there are individual differences in the participation of putamen in motor learning. In particular, individuals who are more reliant on reward for learning, as indicated by a strengthening of connectivity with prefrontal circuits show a functional dissociation between the putamen and sensorimotor areas. This is consistent with the idea of a competition between the somatic and reward-related neural networks in the basal ganglia during early stages of human motor learning (Mink, 1996; Colder, 2015). More generally, these changes may bear on the relationship between reinforcement-based learning and error-guided behavior that has been the focus of previous research. As the sensorimotor goal takes shape following exploration and reinforcement, motor learning and control processes presumably shift to be more error-based. The finding that following learning individuals who show greater increases in connectivity between putamen and medial frontal cortex, show reduced connectivity between putamen and sensorimotor cortex may reflect the neuroanatomical substrate of this progressive shift.

Prefrontal regions involved in reward-guided decision making such as ventromedial and orbitofrontal cortex have extensive anatomical connections with the ventral striatum (Haber et al., 1995), but not with the putamen. The observed changes in connectivity might be explained by the fact that functional connectivity measures are not only modulated by direct anatomical connections but also by indirect pathways (Koch et al., 2002). A potential indirect pathway underlying the observed result entails the projection of vmPFC to the ventral striatum, and in turn to *substantia nigra pars compacta* and then to sensorimotor striatum (Selemon and Goldman-Rakic, 1985; Haber et al., 2000). As part of the reward system, vmPFC and ventral striatum potentially guide motor learning where one is able to learn the appropriate target position and attempt to repeat successful movements. Such reward-guided action selection is thought to involve putamen (Samejima et al., 2005).

Unlike studies in motor adaptation and sequence learning, the current study did not find a statistically reliable correlation between behavioral predictors and changes in the cortico-cerebellar functional network. Activation of the arm area in the cerebellar cortex was observed in task-based localizer scans so the lack of change in connectivity is not due to an inability to observe activity in cerebellum. Moreover, reliable ΔFC between cerebellar cortex and frontal motor areas have been observed previously in the context of force-field adaptation (Vahdat et al., 2011). If cerebellum plays a role in the correction for error (Diedrichsen et al., 2005; Smith and Shadmehr, 2005), the absence of a reliable relationship in the present data may arise by virtue of the task involved in which the sensory error signal is weak at this stage of learning.

The current study provides an account of spontaneous exploration dynamics during the early stages of learning a novel motor skill. We observed a trial-to-trial change in movement direction that was influenced by the preceding feedback. Change in movement direction was greater following unsuccessful trials. We also identified a relationship between exploration and feedback, such that exploration increased proportionally with an accumulation of unsuccessful trials and decreased proportionately with an accumulation of successful trials. Moreover, subjects that produced more accurate movements had smaller change in their movement direction even when the preceding movement was unsuccessful.

We assessed ΔFC using as predictors the change in movement direction following only either successful or unsuccessful trials. Connectivity between SII and a region in the ventrolateral prefrontal cortex varied systematically with changes in movement direction following unsuccessful trials. This area, in monkeys, which is analogous to BA 9/46v, is somatic and has both inputs and outputs to other somatic regions of the brain, such as ventral premotor cortex, the parietal operculum (SII) and the inferior parietal lobule (Petrides and Pandya, 1984). This area of lateral prefrontal cortex is engaged during somatic memory and discrimination tasks in both non-human primates and human studies (Romo et al., 1999; Stoeckel et al., 2003; Kostopoulos et al., 2007). Other neuroimaging studies indicate that the right prefrontal cortex is involved in tasks involving bimanual motor sequences (Sun et al., 2007) and spatial working memory (d'Esposito et al., 1998; Owen et al., 2005), specifically, in relation to visuomotor adaptation (Anguera et al., 2010). The pattern of connectivity changes observed here suggests that working memory may be one of the elements through which reinforcement results in learning especially during movement exploration.

It is worthwhile noting that the changes in functional connectivity observed here were obtained from scans which occurred one hour following the end of the behavioral training. The persistence of learning observed following brief periods of training with a motor task is consistent with a considerable body of behavioral and neuroimaging data. This has been shown behaviorally in the context of reinforcement-based motor learning (Bernardi et al., 2015), as well as in force-field adaptation and visuomotor rotation (Shadmehr and Brashers-Krug, 1997; Krakauer et al., 2005). Persistence of learning in these studies has been observed at intervals up to one week. Similarly, neuroimaging studies have observed that changes in resting-state networks persist for at least 6 hours following brief periods of motor learning (Sami et al., 2014). The persistence of these changes is likely supported by cellular mechanisms such as long-term potentiation and depression (LTP and LTD). These mechanisms affect neuronal metabolism and oxygen consumption, which in turn are reflected in the resting-state signal following learning (Logothetis, 2002).

Overall, it is observed that the acquisition of sensorimotor targets in the early stages of motor learning is dependent on both exploration and positive reinforcement. It is found that reinforcement is associated with an increase in functional connectivity in traditional sensorimotor circuits (M1, S1, PMd, SMA). Areas of prefrontal cortex are also important, subserving both reward-guided behavior (medial prefrontal cortex) and exploratory movement (ventrolateral prefrontal cortex). In future studies, it would be desirable to directly test the role of ventrolateral prefrontal cortex in providing somatic working memory during exploratory behavior. It would also be meaningful to test the idea that as learning progresses there is progressive shift from reinforcement-based learning during the formation of sensorimotor targets to error-based control as learning progresses.

Figures and tables

Figure legends:

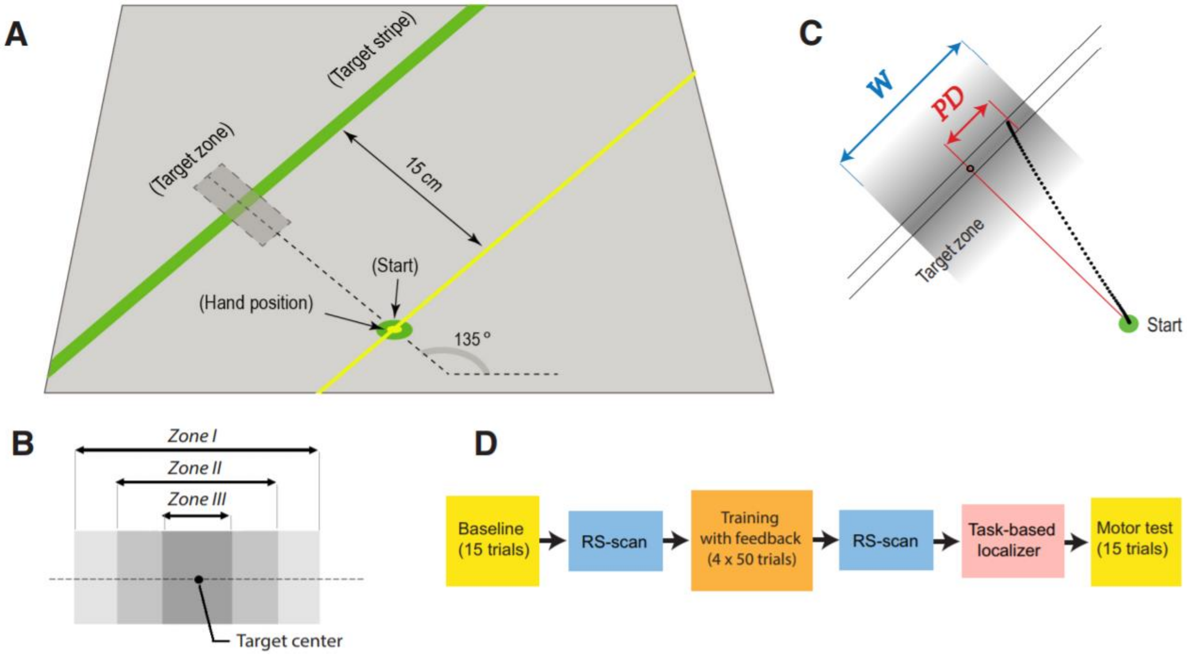


Figure 1. **A**: Schematic diagram showing what the subject sees during the experiment. The robotic manipulandum and the subject's own arm are situated underneath the display mirror.

Participants are supposed to reach out 45° to the left towards the target stripe. Movement trajectories are not visible at any time during the experiment. **B:** Three different target zone widths (*W*) are used during the training blocks. If the movement ends inside the target zone, positive feedback is given. **C**: Schematic illustrating the lateral perpendicular deviation (*PD*) at the movement endpoint. Motor improvement of each subject was quantified as the reduction in the average magnitude of *PD* between the PRE and POST training blocks. **D**: Block diagram showing the overall experimental sequence.

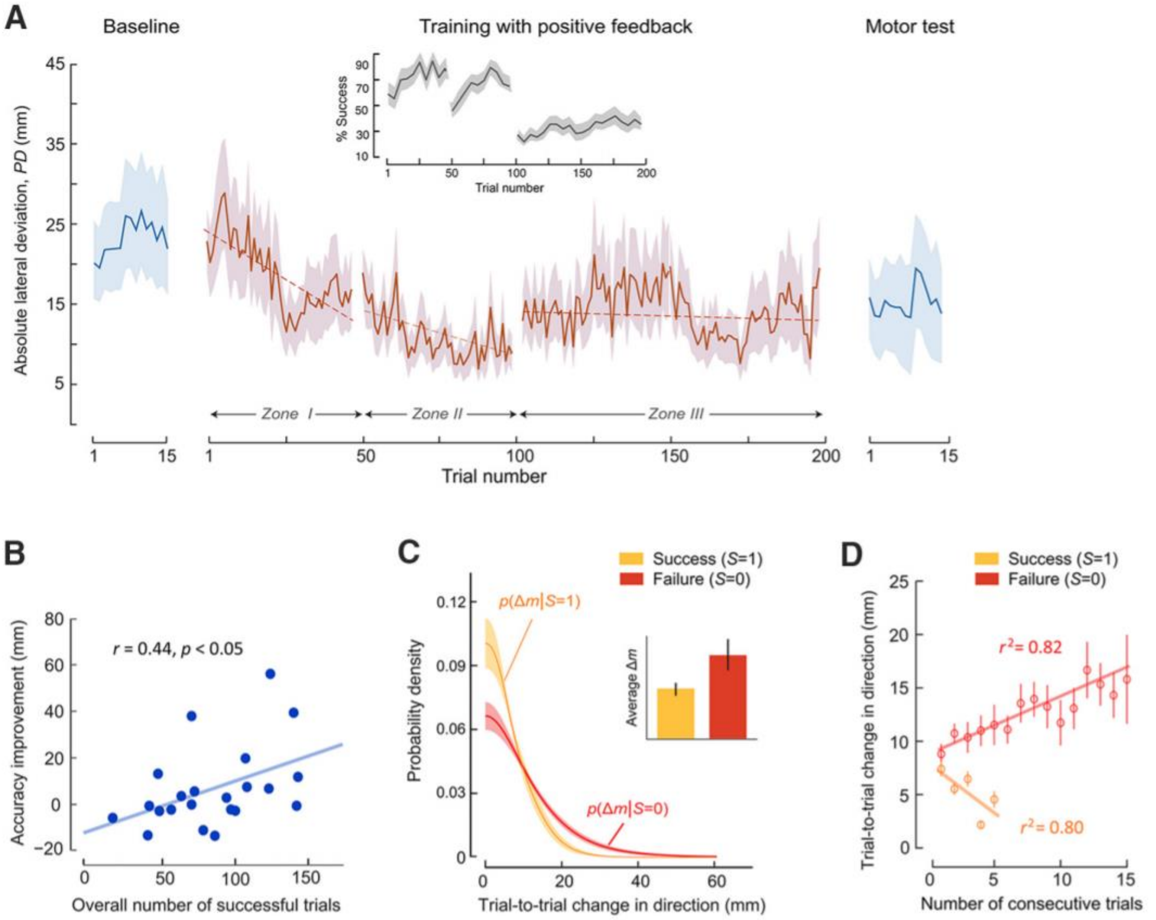


Figure 2 **A:** Behavioral performance (*n* = 22 participants) during baseline movements prior to training (PRE), training trials with feedback, and motor assessment following learning (POST). Zone I to Zone III refers to different sizes of the target zone. Positive feedback was

provided following a successful trial when the reaching movement ended in the corresponding zone. The Y-axis denotes the average $|PD|$ in millimeters. In the inset, the percentage of successful trials over the course of training is given. Shaded colors refers to the standard error. **B:** Linear dependency between the overall number of successful trials and the improvement in accuracy following training. **C:** The overall distribution of Δm, the absolute difference in PD between trial n and $n + 1$ when the current trial n is unsuccessful ($S = 0$, red) and successful ($S = 1$, orange). The bar plot shows the average Δm across subjects. **D:** The average change in direction as a function of number of consecutive successful (orange) and unsuccessful movements (red), fitted with a weighted least-squares regression.

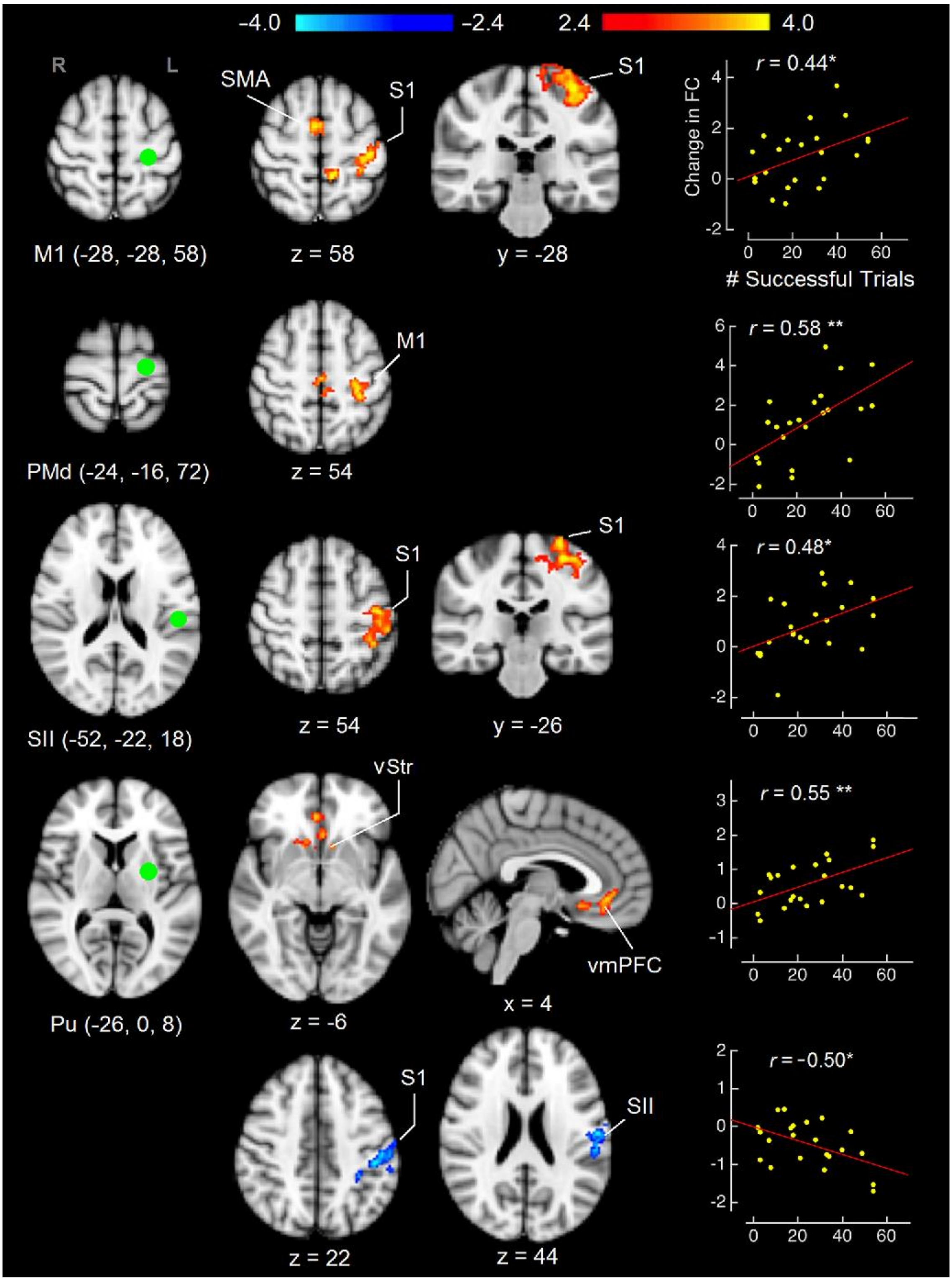


Figure 3: The left column shows seed regions (ROIs) within the sensorimotor cortices. The middle column shows cluster maps indicating statistically significant change in connectivity strength (ΔFC) with each ROI, which are reliably correlated with the number of successful trials as the behavioral predictor. The graphs on the right column are scatterplots illustrating

the linear relationship between ΔFC and the behavioral predictor. An increase in connectivity is represented by the red-to-yellow colorbar, while a decrease is shown with a light-to-dark blue colorbar; thresholded at Z = 2.40, corrected $p < 0.05$. The r represents Pearson's correlation coefficient.

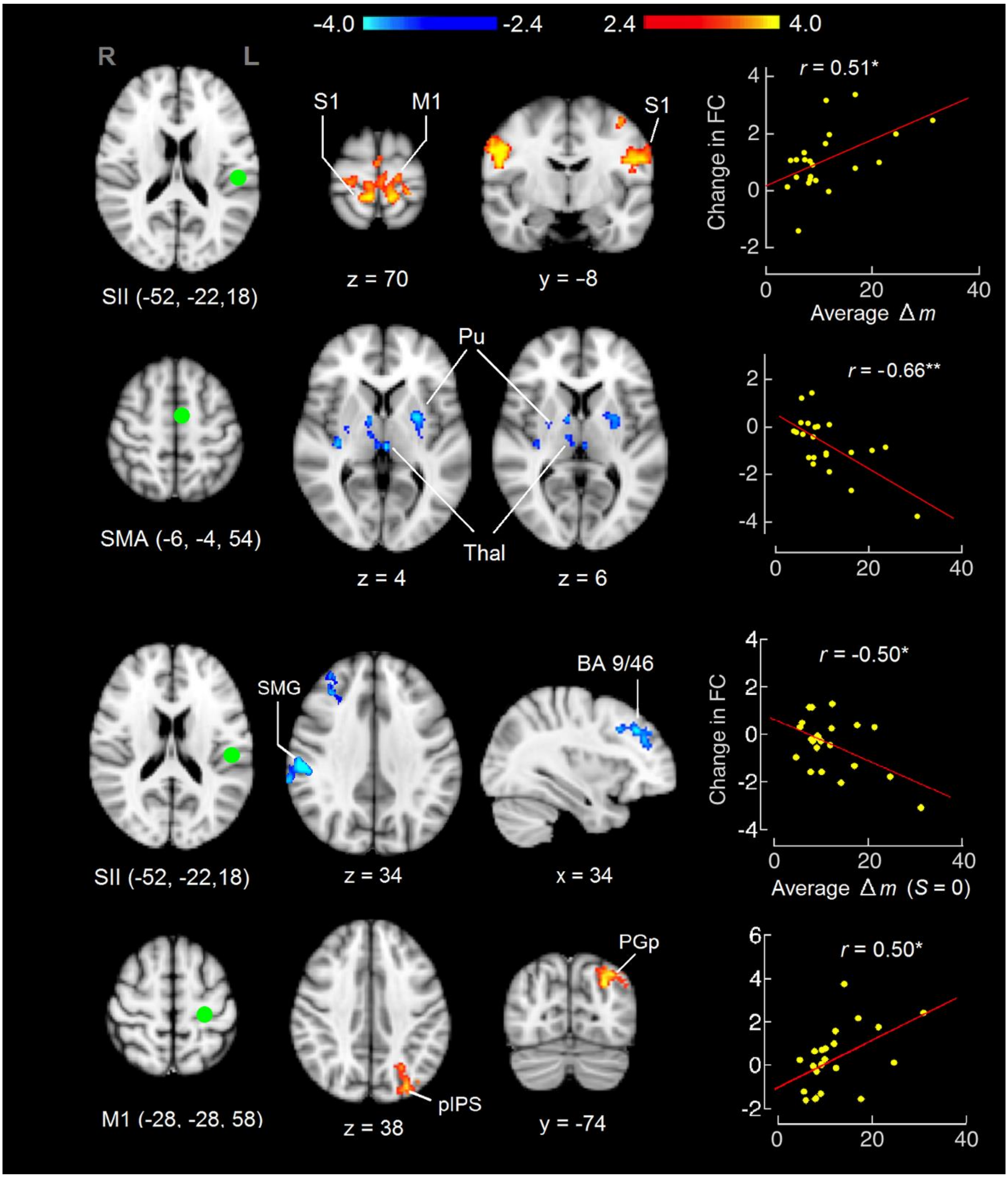


Figure 4: Changes in connectivity related to changes in movement direction (Δm) regardless of whether the previous movement was successful or not (upper panel) and to changes in movement direction following unsuccessful trial (bottom panel). As in Figure 3, the left panel gives seed regions, the middle panel shows clusters whose connectivity with the seed regions varies with change in direction, the right panel shows the relation between change in connectivity and movement direction change.

Table legends

Table 1: List of ROIs used in the resting-state analyses.

The coordinates (in millimeters) are according to the MNI Standard Template. The Z-values correspond to the local maxima at the ROI center obtained from the group task-based localizer.

Seed	Anatomical Label	X	Y	Z	Z-value
S1	Primary somatosensory cortex, BA2, left hemisphere	-28	-42	58	7.25
M1	Primary motor cortex, BA4, left hemisphere	-28	-28	58	7.44
CbV	Cerebellar lobule V, right hemisphere	22	-42	-24	6.33
PMd	Dorsal premotor cortex, BA6, left hemisphere	-24	-16	72	7.12
Pu	Putamen, rostral motor region, left hemisphere	-26	0	8	4.80
SII	Second somatosensory cortex, OP1, left hemisphere	-52	-22	18	5.68
SMA	Supplementary Motor Area	-6	-4	54	5.54

<u>Table 2: Δ*FC* associated with the number of successful trials.</u>

Cluster labels are according to standard atlases. The *p*-cluster is the corrected cluster-level *p*-value related to the predictor of interest. The peak *Z*-value and the coordinates (in millimeters) inside the corresponding cluster are also given according to the MNI Standard Template respectively. The two rightmost columns provide the functional connectivity measure (FC) before (PRE) and after (POST) training, with a negative value indicates an anti-correlation.

Cluster label	***p*-cluster**	**Total voxels**	**Peak *Z***	**MNI-X**	**MNI-Y**	**MNI-Z**	**FC (Pre)**	**FC (Post)**
Postcentral gyrus (S1) – Left	< 0.0001	5309	4.70	-40	-24	54	3.21	4.10
Precentral gyrus (M1) – Left	< 0.0001	1683	3.75	-30	-28	56	4.38	5.51
Precentral gyrus (M1) – Left	0.0089	222	3.88	-28	-16	68	3.65	4.75
Subcallosal gyrus, ventromedial PFC	< 0.0001	654	3.90	2	42	-2	-0.19	0.39
Parietal operculum (SII) – Left	< 0.0001	788	4.25	-56	-14	22	0.76	0.30

Table 3: Δ*FC* associated with the average change in direction.

Details are similar to Table 2. The top shows results predicted by Δm as a whole, regardless of the trial outcomes. The bottom shows results predicted by Δm specifically after unsuccessful trials only ($S = 0$).

Cluster label	***p*-cluster**	**Total voxels**	**Peak *Z***	**MNI-X**	**MNI-Y**	**MNI-Z**	**FC (Pre)**	**FC (Post)**
After both successful and unsuccessful trials								
Precentral, postcentral gyrus (M1/S1)	< 0.0001	3700	4.25	-10	-40	74	1.10	2.18
Thalamus – Left	< 0.0001	654	4.97	-2	-24	4	1.20	0.48
Putamen – Left	< 0.0001	600	3.91	-26	0	4	1.40	0.66
After unsuccessful trials only								
Middle frontal gyrus (BA 9/46) – Right	0.0031	308	4.54	30	38	28	0.75	0.63
Angular gyrus (PGp) – Left	0.0008	358	3.75	-28	-66	30	0.85	1.36
Angular gyrus (PGp) – Left	0.0130	326	4.32	-32	-68	34	0.00	0.52

Supplementary Materials

In this section, analyses which were not parts of the published article but potentially informative will be described. The first part gives additional behavioral results, while the second discusses the rationale behind the number of independent components (ICs) that were removed to clean the resting-state data in the study.

1. Additional kinematic measures

Overall, endpoint movement accuracy improved after training as measured by the absolute lateral deviation, $|PD|$ (see main text and Figure S1-A) [$t(21) = 2.080$, $p < 0.05$]. Changes in endpoint variability were assessed by quantifying the standard deviation of the signed lateral deviation (PD) before and after training blocks. Although the reduction in endpoint variability following training was not significant [$t(21) = 1.443$, $p > 0.10$], participants that received more reinforcement had a greater reduction in variability ($r = 0.68$, $p < 0.001$). Additionally, participants that had more accurate movements (less absolute deviation) had more consistent (less variable) movements ($r = 0.46$, $p < 0.05$).

We also examined trial-to-trial change in direction (Δm) in early and late phases of each training block. Figure S1-B shows variation of Δm following successful (orange color) and unsuccessful (red color) trials. Early and late phases were defined as the first and the last 25 training trials in each block, respectively. Zone-I through Zone-III denote the widest and smallest target zones respectively. As mentioned in the main text, the magnitude of change in direction (Δm) following failed trial was reliably different from the one following successful trial [$F(1,21) = 26.98$, $p < 0.001$]. However, there were no significant differences in Δm across blocks ($p > 0.1$). The fact that the level of trial-to-trial exploration did not change may suggest that subjects continued to sample the various reaching directions over the course of learning. This is presumably due to the fact that the width of the target zone, and thus, task difficulty changed from one block to the next, causing a need to continuously explore other reaching directions in order to be successful.

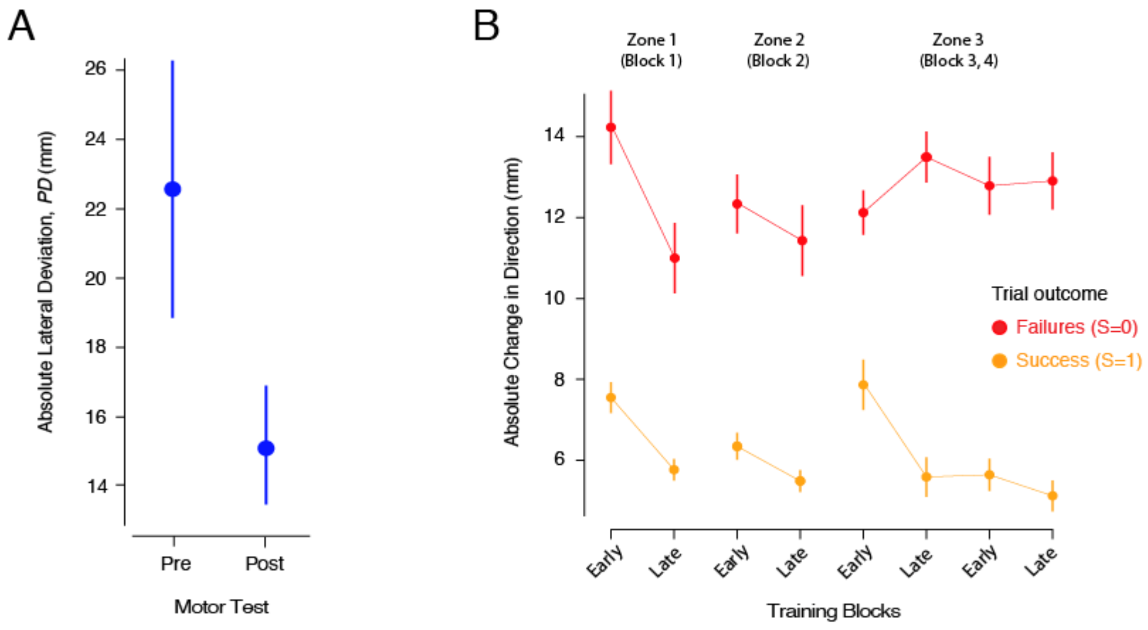


Figure S1: (A) Improvement in accuracy is shown as a reduction in absolute lateral deviation (|*PD*|), measured before and after training, during which the binary feedback was removed. (B) Evolution of trial-to-trial change in movement direction following successful (*S*=1) and unsuccessful (*S*=0) trial. Each zone is associated with a different target width (see main text). Trial-to-trial change in direction was averaged across the first and last 25 trials of each training block, denoted as “Early” and “Late” phase.

2. Resting-state data cleanup

In the neuroimaging analyses, ICA denoising was performed in order to remove artifact from the resting-state data. As mentioned in Section 2.3 of the main text (Methods section: *fMRI Data Preprocessing and ICA*), noise components were selected manually and comprised 10% of the total independent components (ICs) identified using FSL- MELODIC. These nuisance components were included in the general linear model (GLM) in the subject-level analyses (see main text). To validate our choice, we evaluated the effects on removing different percentages of the total number of ICs. Denoising was done by systematically removing none, 5%, 20%, 40% and 50% of the total ICs for each run in each subject. It should be noted that in our pre-processing steps, we let FSL-MELODIC automatically decide the total number of ICs for each run, of each subject. As a result, different subjects could have different total number of ICs,

As mentioned in the main text, the time series of each noise component was included as a predictor in the general linear model (GLM) together with the ROI time series of the predictor of interest and the voxelwise time series of the whole brain (uncleaned resting-state data) as the dependent variable. This analysis yielded a set of brain regions whose BOLD time-series was temporally correlated with that of the ROI, after regressing out the noise components. In order to evaluate how well the model explained the temporal features of the resting-state data, we then first computed the R-squared value within the grey matter region of the brain. Subsequently, the adjusted R-squared (R^2_{adj}) was calculated as a measure of the overall model fit. The use of R^2_{adj} rather than R^2 is preferred, as the latter does not take into account the number of predictors in the model, making it vulnerable to overfitting. In other words, it is theoretically possible to keep adding predictors that would eventually explain the entire pattern of the dependent variable, leading to an artificially high R^2. In contrast, R^2_{adj} increases only when the new predictors improve the model by more than would be expected by chance alone. This validation was systematically done on two selected ROIs, i.e. one in the cerebral cortex (SMA) and the other subcortically (putamen). Subsequently, we computed the average R^2_{adj} across all subjects for different percentages of nuisance ICs.

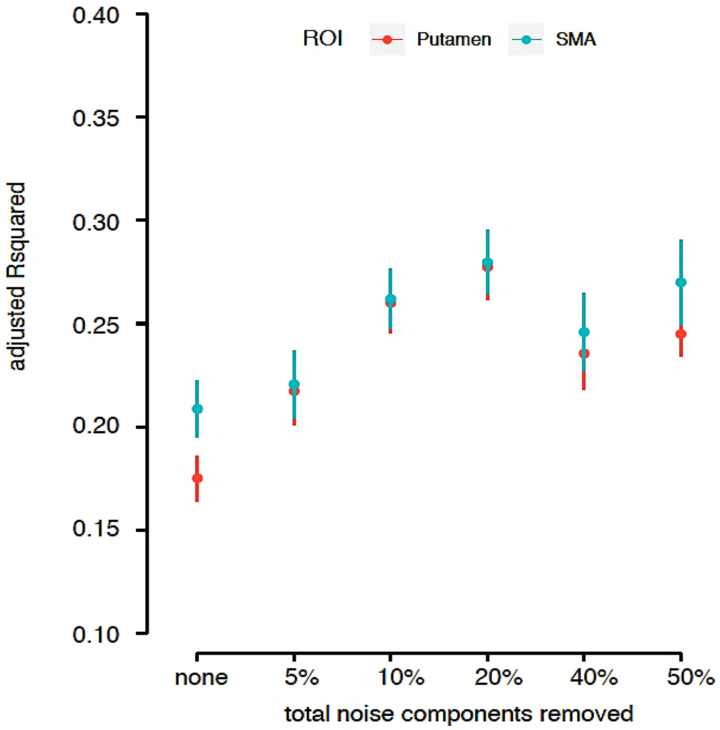


Figure S2: Adjusted R-squared of the general linear model (GLM), averaged across all subjects and 2 seeds, for different number of nuisance components. The abscissa represents percentage of nuisance independent components removed. For example, if there were 45 ICs identified by FSL-MELODIC algorithm, then removing 10% here means removing 5 components (after rounding up to the nearest integer).

Overall, differences in R^2_{adj} values among different number of nuisance ICs removed were observed [2-way repeated measure ANOVA, $F_{(5,105)} = 33.29$, $p < 0.0001$, with Greenhouse-Geiser correction] (Figure S2). Post-hoc pairwise *t*-test with Tukey correction showed that there was a significant improvement in R^2_{adj} between “none” and 5% ($p < 0.005$) category, and between 5% and 10% category ($p < 0.005$). In contrast, no significant difference in the R^2_{adj} between 10% and 20% category was found ($p > 0.5$). Likewise, the difference R^2_{adj} between 20% and 40% ($p > 0.05$), or between 40% and 50% category was not significant ($p > 0.5$). This suggests that removing more components (> 10%) does not provide additional benefits to the R^2_{adj}. Thus, in our work, the number of components classified as noise and then removed was 10% of the total identified components. Refer also to Figure S3 and S4.

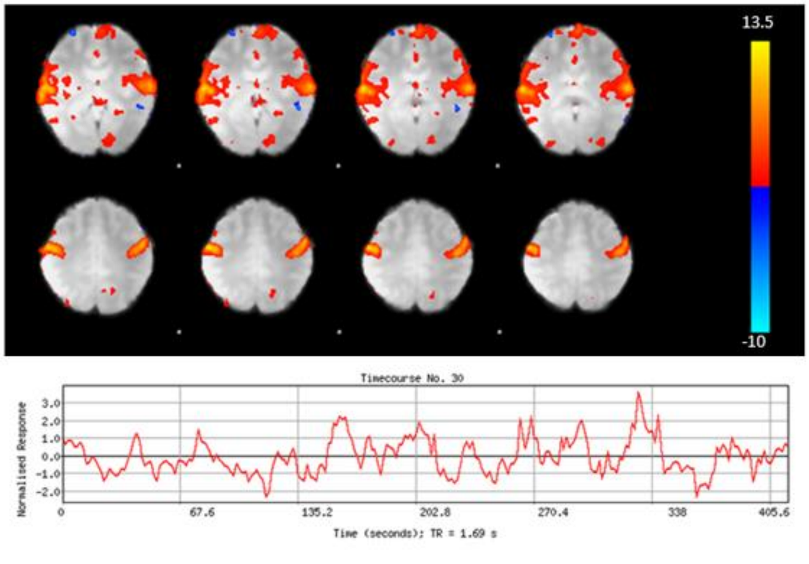


Figure S3: An example of a resting network (sensorimotor cortices) as produced by FSL – MELODIC. The colour bar represents statistical values of the connectivity strength. The BOLD time series of the maximum area is also given below it.

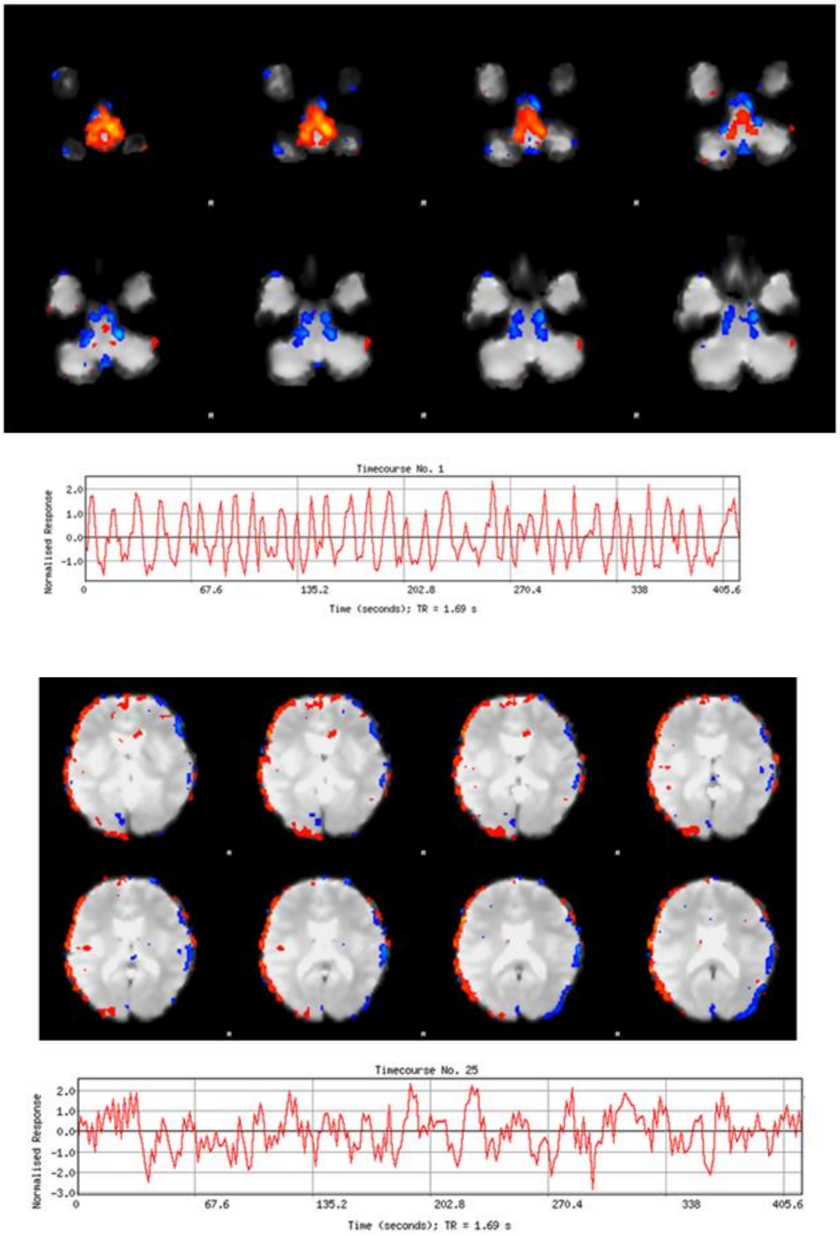


Figure S4: Examples of unwanted independent components (ICs), presumably those related to heart beat (above) and head motion (below). See (Kelly et al., 2010).

Chapter 3 : Somatosensory working memory and its relation to motor learning

3.1. Preface

In the following chapter, I investigate the role of somatosensory working memory in reinforcement-based motor learning in humans. One characteristic of the initial stages of learning is trial and error and exploration, marked by a relatively high trial-to-trial change in movement direction (Sidarta et al., 2016). Feedback related to trial outcome can act as reinforcement such that learners understand which movements are more desirable than others. Inevitably, it is necessary to keep in memory which prior movements are successful and which are not. For this reason, working memory is hypothesized to be important for motor learning. While there has been prior work investigating other aspects of learning, specifically, motor memory consolidation, retention of motor performance, and savings, the role of sensory working memory in the learning process has been explored in only a small number of papers and is not well understood. This chapter studies the relationship between working memory and reinforcement-based motor learning in two different experiments.

In the first experiment, an *offline experiment*, working memory and motor learning are tested on separate days. This procedure is designed to assess working memory capacity in the somatosensory domain and to test if there is any relationship between individual differences in memory capacity and changes in motor performance following training. Subsequently, I examine whether the relationship between working memory and learning is context or domain-specific to somatosensory modality using three control tasks (visuospatial memory task, and the standard forward and backward digit span tasks). The somatosensory working memory task involves a recognition-based memory test where subjects have to indicate whether a test item is one of a previously presented set of memory items. During this task, the subjects remain passive as the robot displaces the arm in different directions. The minimum angular difference between each direction in the memory set is 10°, threefold higher than the minimum detectable change in angle found using joint-position matching task (Goble et al., 2005). In the motor learning, the task is similar to the one used in Chapter 2 and also in Bernardi et al. (2015). One difference is that the target zone has a constant width of 10 mm rather than a variable width as in the previous studies. The advantage of using a constant target width is the consistency of the training environment and task demand across blocks, and accordingly, the ability to interpret differences in variables such as movement accuracy or the number of reinforced trials with respect to a common reference.

Overall, it was found that subjects show improvement in motor performance over the course of training as measured by the absolute deviation, signed deviation (bias), and movement variability. In this offline experiment, the capacity of somatosensory working memory was found to be at least 2 items. This is consistent with the role of short-term memory of prior movements is limited to movement planning in the immediate future. It was also found that improvement in reaching accuracy to the rewarded direction is uniquely related to somatosensory working memory. Changes in motor performance over the course of learning were did not correlate with digit span memory which is known to tap in the verbal component of working memory. A reduction in movement variability was found to be related to both somatosensory and visuospatial working memory, suggesting that this change in performance is not domain-specific.

A second experiment is conducted that involves an *online procedure* in which working memory tests are interleaved with motor learning trials. The merits of combining memory and learning tests are as follow. Firstly, the offline procedure does not test working memory of subject's self-generated movement. To overcome this limitation, the memory test is presented in between learning trials by taking the subject's own movement as the reference direction. Secondly, because the movements during training blocks can be both successful and unsuccessful, we are able to assess the influence of positive feedback on working memory performance. In this procedure learning is quantified based on measures obtained from within the training blocks, that is, the total number of reinforced movements.

In the motor learning task, an improvement in the design is used. One limitation found in previous work (Bernardi et al, 2015; Sidarta et al., 2016) that we sought to correct in the current design was that subjects were explicitly told the required reaching direction (45 deg). It is possible that subjects that had a better sense of the geometry of the workspace, might perform better under these instructions. As a design improvement, the target bar was replaced by an arc in the form of a quarter circle. The task goal is the same, that is, to reduce the lateral deviation with respect to a specific reaching direction. The arc was shown during familiarization trials to give a sense of distance but later removed during the actual experiment. Due to the nature of the task, it is more natural and reasonable to measure movement accuracy by the angular deviation with respect to the rewarded reaching direction instead of lateral deviation at movement end point.

As with the offline experiment, there is a correlation between somatosensory working memory and learning. It is seen that people with better memory learn more. It is also found that overall subjects remembered reinforced movements better than non-reinforced movements, consistent with the idea that working memory has a differential preference to retain movements that are rewarded. The online working memory capacity seems limited to 2 prior movements, suggesting once again that only a limited number of most recent information is incorporated in planning upcoming movements.

Sidarta A., VanVugt FT, Ostry DJ. "Somatosensory working memory in reinforcement-based motor learning". *J Neurophysiology*.

3.2. Abstract

Recent studies using visuomotor adaptation and sequence learning tasks have assessed the involvement of working memory in the visuospatial domain. The capacity to maintain previously performed movements in working memory is perhaps even more important in reinforcement-based learning in order to repeat accurate movements and avoid mistakes. Using this kind of task in the current work we tested the relationship between somatosensory working memory and motor learning. The first experiment involved separate memory and motor learning tasks. In the memory task the participant's arm was displaced in different directions by a robotic arm and the participant was asked to judge whether a subsequent test direction was one of the previously presented directions. In the motor learning task, participants made reaching movements to a hidden visual target and were provided with positive feedback as reinforcement when the movement ended in the target zone. It was found that participants that had better somatosensory working memory showed greater motor learning. In a second experiment, we designed a new task in which learning and working memory trials were interleaved, allowing us to study participants' memory for movements they performed as part of learning. As in the first experiment we found that participants with better somatosensory working memory also learned more. Moreover, memory performance for successful movements was better than for movements that failed to reach the target. These results suggest that somatosensory working memory is involved in reinforcement motor learning and that this memory preferentially keeps track of reinforced movements.

Keywords:

Sensory working memory, somatic, exploration, reinforcement learning

New & Noteworthy

The current work examined somatosensory working memory in reinforcement-based motor learning. Working memory performance was reliably correlated with the extent of learning. Using a paradigm, in which learning and memory trials were interleaved, memory was assessed for movements performed during learning. Movements that received positive feedback were better remembered than movements that did not. Thus, working memory does not track all movements equally but is biased to retain movements that were rewarded.

3.3. Introduction

When learning motor skills such as swimming or dance, it is necessary to discover the limb configuration that enables successful movement. In motor tasks such as these, there is limited visual information and the only performance measure available is success or failure. Learning under these conditions proceeds at least in part through exploration and trial and error. In this model of motor learning, sensory working memory, which enables maintenance and decision-making related to prior sensory information, is presumably involved in movement selection by allowing repetition of successful movements and the avoidance of errors. However, to date little is known about the relation between sensory working memory and this kind of motor learning.

Short-term memory has been previously shown to store feedforward control of reaching transiently before being consolidated in more stable and long-term memory (Brashers-Krug et al., 1996; Krakauer et al., 1999; Tong et al., 2002). Individual differences in working memory capacity have been assessed in relation to the amount of motor learning. It was shown that estimates of visuospatial working memory capacity correlate with the rate of sequence learning and visuomotor adaptation (Bo and Seidler, 2009; Bo et al., 2011). In a related study that included neuroimaging, spatial working memory was involved early in visuomotor adaptation and was associated with task-related neural activity in right dorsolateral prefrontal cortex and bilateral inferior parietal cortex (Anguera et al., 2010). The role of visuospatial working memory in visuomotor adaptation has also been demonstrated when adaptation involved either an explicit strategy or adaptation to an abruptly introduced perturbation (Christou et al., 2016). In contrast, using a gradual perturbation which minimizes explicit strategies, working memory capacity was no longer a reliable predictor of learning.

In situations where there is only success or failure information about movement outcome (reinforcement learning), learning is partly driven by positive feedback and reward which serve as reinforcement. Prior studies have reported the influence of reward on motor learning when other types of information are available as well. Specifically, positive feedback during training increases memory for reaching direction in a visuomotor adaptation task (Shmuelof et al., 2012; Galea et al., 2015; Therrien et al., 2015) and memory of pinch force (Abe et al., 2011). While it has been established that reward during training plays a role in motor learning, its effects on working memory remain unclear.

Recent behavioural studies have investigated somatosensory processes involved in reinforcement-based motor learning in a task in which participants received binary feedback on their movement outcome (Bernardi et al., 2015, Therrien, 2018 #97). In the present paper, we assessed the relationship between somatosensory working memory and human motor learning in a similar task in which participants made movements to hidden targets and received positive feedback when the movement finished within a target zone. We hypothesized that given the paucity of visual information, the task would be heavily reliant on somatosensory information (Bernardi et al., 2015) and accordingly that participants with better somatosensory working memory would show better learning. This hypothesis was addressed in two separate experiments. The first was an *offline* experiment in which working memory and learning tasks were completed on separate days and working memory capacity was tested as a predictor for motor learning performance. The second experiment aimed to understand what movements participants held in memory during the experiment. One hypothesis is that participants held both successful and unsuccessful movements in memory, because these are the movements to adopt or avoid, respectively, in the future. An alternative hypothesis is that predominantly rewarded movements are remembered, since in principle repeating these movements accurately is sufficient for performing the task. These contrasting hypotheses were tested in an *online* experiment, in which motor learning and memory trials were presented in an interleaved fashion to assess participants' memory for their own movements. The online technique enabled us to examine in a trial-by-trial manner whether successful or unsuccessful movements were remembered more or less well. Overall, it was observed that participants who had better somatosensory memory learned more in the motor task. The online experiment also revealed that that successful trials (trials with positive feedback as reinforcement) were better remembered.

3.4. Materials and Methods

We conducted two separate studies which measured *offline* and *online working memor*y respectively. A total of 30 right-handed participants were recruited (6 males, *M* age = 22.11 years old, *SD* = 2.85) for the offline working memory experiment which consisted of two experimental sessions, one testing motor learning and the other testing working memory. Each session was completed on a separate day with the order counterbalanced across participants. For the online working memory experiment, we recruited another 30 right-handed participants (4 males, *M* age = 20.9 years old, *SD* = 2.45) for a single-day study. All procedures were approved by the McGill University Faculty of Medicine Institutional Review Board. The participants were healthy adults with no prior physical or neurological conditions.

Both experiments used a two degree-of-freedom robotic manipulandum

(Interactive Motion Technologies) with a vertical handle attached to the end-effector. Participants were seated in front of the robot with their right shoulder abducted to about 70 degree and the elbow supported by an air sled. A semi-silvered mirror, which served as a display screen, was placed just below eye level and blocked the vision of the arm and the robot handle. A white start circle, 20 mm in diameter, was positioned on the display screen at about 30 cm in front of the participant's body midline.

Offline Working Memory Experiment

Sensorimotor Learning Paradigm

The task in this study was similar to that used in previous work (Bernardi et al., 2015; Sidarta et al., 2016). Briefly, in the left part of the workspace the participant was shown a 1 cm thick white target stripe or bar, within which there was a hidden rectangular target zone that also had a width of 1 cm (Figure 1B). The center of the zone was located 15 cm from the center of the start circle. Parallel to this target stripe was a thin yellow line that indicated the distance of the hand from the stripe. A small 12 mm diameter yellow circle attached to the yellow line corresponded to hand position. This circle was shown briefly at the beginning of each movement and disappeared as soon as the robot handle left the start position. No information about the lateral deviation of the hand was provided during movement so the participants could not use the error information associated with lateral distance from the target as a learning signal.

The participant was first given 15 familiarization trials with instructions. In both the familiarization trials and in the actual experiment, participants were told that after a "Go" cue appeared, they had to perform straight outward reaching movements to the target stripe without making corrections. Each movement had to stop within the stripe and be completed within 500 - 700 msec. The participant was given feedback about the movement speed verbally if they were consistently too slow or too fast. However, there was no penalty if the movement did not end on time or ended outside of the stripe. Once the movement ended, the robot brought the arm back to the start position.

The experiment began with a block of 25 baseline trials without any feedback regarding movement accuracy. Participants were instructed to reach at an angle of 45 degrees to the left. Then, they performed 4 training blocks of 50 trials each and were told to learn which movement to the unseen target was successful, that is, ended within the target zone. The goal was to reduce the deviation with respect to the hidden target. Success was determined solely by the lateral deviation at the movement endpoint, not the movement distance or speed. Following a successful trial, an animated explosion and the words “Nice shot!” along with a pleasant tone, and a running score appeared on the screen to provide positive feedback as reinforcement. The width (1 cm) and center position of the target zone (45 degrees to the left workspace) was fixed. In this offline experiment, the participant was told to pay attention to the arm configuration when successful and to make as many successful movements as possible (no such instruction was given in the online experiment). The session ended with a final set of 25 movements without any feedback which evaluated motor accuracy following learning.

Offline Somatosensory Working Memory Task

An offline somatosensory working memory task tested recognition memory. A set of memory items was presented one at a time followed by a test item (probe). The participant had to indicate whether the test item was in the memory set or not. In the present working memory task, the to be remembered items were passive limb displacements produced by the robot in directions in the left part of the workspace as in the sensorimotor reaching task described above. During the experiment, the view of the arm was occluded and the screen was completely blank.

Each trial of the memory test began with the words “New Round” presented on the screen as a visual cue. The participant was instructed to remain passive as the robot displaced the right arm outward in four different directions (Figure 1C), each had an amplitude of 15 cm and took 900 msec to complete. After a brief hold time at the destination, the manipulandum moved the arm back to the start position. There was a delay of 500 msec between consecutive movements. Once the participant had experienced the four memory items, a tone was played which was a cue indicating the following displacement would be the test direction. The participant responded verbally after having experienced the test direction, that is, “Yes” if they felt that the test item was one of the four directions presented in the memory set and “No” otherwise.

In a given block of memory test trials, the test item was one of the four directions in the memory set on half of the total trials. In such trials, the test items were presented with varying lags separating the test item and the to-be remembered item in the memory set. For example, a lag 2 memory trial means that the test direction was the same as the memory set direction presented two items ago (Figure 1C, dotted arrow). In the remaining half, the test item was a lure, that is, it was a totally new direction. The order in which the test direction was a lure or was one of the previous memory set items was randomized across trials.

On a given memory test trial, the set of four memory items and the test item (probe) was obtained as follows. In all cases we first started with six directions equally spaced with 10 degrees separation. These six directions d_i were found using the following formula: $d_i = 10i + 100° + j$, where $i \in \{1, 2, 3, .., 6\}$ and $j \in \{-11°, -10°, .., 10°, 11°\}$. Two out of these six directions were then chosen pseudorandomly to be discarded, resulting in the memory set of four directions. The directions that were discarded were not at the two extremities and not adjacent to one other. Lastly, the test direction was selected as follows. If this was a lure trial, one of the two removed directions became the test direction, otherwise one of the four memory directions was chosen at random to be the test direction.

Before the start of the actual task, each participant went through six familiarization trials with feedback (correct/incorrect) to ensure that they understood the task. The actual task consisted of 6 blocks of 24 trials that lasted for about 8 – 10 minutes each with a short break after each block. Overall, there were 18 trials at each lag.

Visuospatial Working Memory and Digit Span Tasks

To determine whether the relationship between working memory and learning was specific to somatosensory memory, we invited the participants who did the above tasks to participate in a set of control conditions. The first task was a visuospatial working memory task which assessed the ability to remember the locations of items presented visually in space. The task resembled the somatosensory working memory task with the exception that there was no displacement of the arm and the stimuli were locations of the end of movement that were shown as white circles on the screen (Figure

1D). Each participant underwent a series of familiarization trials before the actual test. They were told to pay attention to a 20 x 20 cm bounded area on the left side of the screen with the white start position in its lower right corner. Then, four white circles would be presented one after the other, followed by a tone and a test circle. The participant then had to verbally indicate whether the test circle was in the memory set or not. As with the somatosensory working memory task, the actual test consisted of 6 blocks of 24 trials each and the set of memory and test items were generated using similar procedures.

Two other cognitive tests were also employed as control tasks. Following the visuospatial working memory task, the participants were presented with forward and backward digit span tasks to tap into the verbal short-term memory (Wechsler, 1999). In this task, participants were presented a sequence of digits on the screen and then had to report the sequence in forward or backwards order (as specified by an instruction) using the keypad. At the beginning of the test, a message would appear on the monitor screen to tell the participants if the task was a forward or backward task. During the experiment, a series of numbers at a pace of 1 second per digit was presented with 1 second pause in between sequential digits. Both tasks began with a set of 3-digit numbers and continued up to 9-digit numbers. Within a set, there was no single number that was repeated, and the digit sequence was random. Prior to the experiment, we provided the participants with 3 familiarization trials with instructions using 2-digit numbers. Subsequently, they began the actual task which consisted of the forward and backward digit span task (with the order counterbalanced). Task performance was quantified as the proportion of correct trials. Out of 30 original participants, 25 participated in the control conditions.

Data Analysis

Motor performance in each trial was quantified as the perpendicular deviation (*PD*) at movement end point from a straight line originating at the start position and passing through the center of the target bar, which is exactly 45 degrees to the left of a straight-ahead movement. If the movement ended beyond the target bar, the perpendicular lateral deviation was computed with respect to this movement end point (Figure 1B). Movements that ended closer to the center had smaller *PD* scores. For each participant, the average absolute deviation ($|PD|$) before (Pre) and after (Post) training was calculated using the 25 trials without any positive feedback, and the difference served as a measure of improvement in accuracy, with larger positive values corresponding to greater learning. Using the same set of trials, we also assessed accuracy in terms of movement bias (or the average value of signed *PD*), and end point variability (or standard deviation of signed *PD*) which evaluated movement consistency. During training blocks, to assess whether positive feedback or its absence influenced the movement on the immediately following trial, we calculated the absolute change in movement direction following each successful and unsuccessful trial as $\Delta m_n = |PD_{n+1} - PD_n|$, which gives the difference in *PD* between the trial n and $n + 1$, contingent upon trial n being successful or not.

In the somatosensory working memory task, we quantified both the hit rate (proportion of “yes” responses when the test item was part of the memory set) and the false alarm rate (proportion of incorrect “yes” responses) for each lag, and the difference between hit and false alarm rates was obtained. Using ANOVA we assessed differences in hit – false alarm rates across lags. The same analyses were conducted for the visuospatial working memory test. Tests for normality and assumption of sphericity of the dataset were conducted using Shapiro-Wilk test and Mauchly's test respectively. Relevant post-hoc analyses were done with Bonferroni-Holm correction.

A composite somatosensory working memory score of each participant was computed as the average of hit – false alarm rates over all 4 lags. A similar approach was used to obtain individual's visuospatial working memory score. Performance on the forward and backward digit span tasks was measured by proportion of correct trials. Subsequently, we computed the correlation between each of the memory scores and the measures of learning together with the 95% bootstrapped confidence interval (CI).

Online Working Memory Experiment

Sensorimotor Learning Paradigm

Whereas the first experiment was designed to test the relationship between the somatosensory working memory performance and the amount of motor learning measured separately (offline), a second experiment measured participants' memory for their own movements on a trial-by-trial basis during the learning processes. As such, the working memory test was interleaved with the motor learning trials themselves. This experiment used the same basic setup as the offline experiment but was divided into two parts, which involved sets of movements to a hidden target at the right and the left of the workspace respectively. Memory testing was restricted to 2 lags, one tested at the right and the other at the left. The assignment of movement direction and memory test lag was random across participants (see Figure 1E). There was a 10-minute break halfway through the experiment at which time participants switched movement directions and lags.

The study began with familiarization trials where a quarter arc with a 1 cm thickness was shown on the screen (Figure 1F). As before, vision of the participant's arm was blocked. Participants were instructed to move to any point on the arc after the “Go” cue appeared and to make straight movements without corrections. The yellow hand cursor position was removed once the arm moved outside of the white start circle. The required movement duration was 500 – 700 msec but there was no penalty if the movement did not end on time or outside the target arc. Once the movement ended, the robot brought the arm back to the start position. Directional error was measured in terms of angular deviation (*AD*) from the true target direction at the maximum movement speed. The width of the target zone was 5 degrees and positive feedback was provided if the angular deviation was within ± 2.5 degrees (Figure 1F).

Following the familiarization trials, the arc was removed. However, the participant was instructed to move in the direction of the arc and was told there was a target located in the now hidden arc. The task was to search for the correct direction to the target and then to continuously move in the same direction. When the direction was correct, the trial was considered successful and the participant was given the same positive feedback as in the offline experiment (an animated explosion, a pleasant tone, and a score). This positive reinforcement was independent of the movement length although we told the participants during the familiarization trials whether the movement was too far beyond the arc or too short. For each participant, we chose a participant-specific target direction as follows. The participant first made 15 baseline movements (without feedback). The target direction was then set to the direction of the first movement after the 15th trial that fell within the range of 20 to 70 degrees relative to the horizontal at the right of the workspace, or 110 to 160 degrees at the left. This provided at least 15 movements in which participants randomly explored the workspace before the first reinforced (successful) trial. It also eliminated the use of explicitly defined directions to the target. Throughout training, the width and position of the reinforced direction did not change. After the random exploration phase, participants completed 4 blocks of 60 training trials with positive feedback when successful. This was followed by 25 further movement trials with no feedback.

Online Somatosensory Working Memory Test

The online working memory test was designed to assess participants' memory for their own movements during motor learning in a trial-by-trial manner (Figure 1G). Individual reaching movements were recorded from movement start to movement endpoint. On a memory test trial, the robot would replay a rotated version of the previous movement (in the case of a lag 1 memory test) or the movement two trials before (in the case of lag 2). The rotated movement was 5 degrees to the left or right of the participant's original movement, selected at random. The movement that was used for the working memory test will be referred to as the *seed movement*. The task in the online working memory test was to indicate the direction of the rotation relative to their seed movement direction. Participants responded "Left" or "Right" for this purpose.

The online working memory tests were presented once every 5 – 8 trials according to a probability distribution shown on Figure 1H. A visual cue on the display screen appeared for 1500 msec indicating that the upcoming movement was a memory test. After responding, participants continued the training by again making reaching movements to the occluded target. Participants were explicitly informed whether lag 1 or lag 2 memory judgements were required for a given workspace direction.

Data Analysis

In the online experiment, movement accuracy was quantified using absolute angular deviation $|AD|$ measured at the maximum movement speed. We used an arc along with the angular deviation so that the target location could be made different for each participant while still ensuring that all participants made movements of equal distance. The total number of trials with positive feedback over the course of the 4 training blocks was used to quantify a reinforcement index of learning. As before, we quantified the effect of positive feedback on the current trial on movement direction on the following trial with $\Delta m_n = |AD_{n+1} - AD_n|$, contingent upon trial n being successful or unsuccessful. Since the working memory test was interleaved in between two training trials (trial n and $n + 1$), we also examined whether the presence of the online working memory test had any influence on the change in movement direction (Δm) immediately after the memory test.

Working memory performance at each lag was quantified using the proportion of correct responses. To assess the effects of positive feedback on memory for movements, we examined whether memory was different following successful versus unsuccessful movements. For each participant and each lag, working memory performance contingent upon successful and unsuccessful seed movements was calculated separately. A 2-way repeated-measures ANOVA was used to evaluate whether the success of a seed movement affected working memory performance at different lags. Tests for normality and assumption for sphericity were conducted using Shapiro-Wilk test and Mauchly's test respectively. The Greenhouse-Geiser corrected *p*-value was used if the sphericity assumption was violated. The correlation between motor learning and the overall memory performance was computed together with the 95% bootstrapped confidence interval (CI).

3.5. Results

Offline Working Memory Experiment

In this study, movement accuracy before and after learning was quantified as the absolute perpendicular deviation $|PD|$ at the movement endpoint, based on the 25 movements in the baseline (Pre) and motor evaluation blocks (Post). Movement bias was measured as the average value of signed *PD*. Overall, participants showed learning as indicated by a reliable decrease in the mean absolute lateral deviation, $|PD|$ [$t(29) = 4.82$, $p < 0.001$] (Figure 2A), and in the magnitude of bias [Pre: $M = 1.57$ cm, $SD = 0.73$ cm, Post: $M = 1.01$ cm, $SD = 0.57$; $t(29) = 5.19$, $p < 0.001$]. Using the standard deviation of signed *PD* as a measure of movement endpoint variability, it was found that there was decrease in variability from before to after learning [Pre: $M = 1.29$ cm, $SD = 0.47$ cm, Post: $M = 0.96$ cm, $SD = 0.32$ cm; $t(29) = 3.16$, $p < 0.005$]. Moreover, there was a correlation between the reduction in error magnitude and the reduction in variability ($r = 0.45$, $p = 0.013$, CI [0.29, 0.67]), indicating that participants that showed greater improvement in accuracy also had a greater reduction in variability.

To quantify the reinforcement rate over time, a linear function was fit to reinforcement rate (with a bin size of 5 trials) for each participant to provide an estimate of the slope. The average slope across participants was shown to be significantly different than zero [one sample *t*-test: $t(29) = 3.18$, $p < 0.005$] suggesting that the amount of positive reinforcement increased over training trials (Figure 2B). Absolute change in signed *PD* between the current and next immediate movement (Δm) was computed to assess the effect of the positive reinforcement on subsequent movements. Non-reinforced movements resulted in a greater trial-to-trial change in movement direction than reinforced movements, which presumably reflects exploration to find the correct direction when movements fail to end in the target zone [$t(29) = -6.33$, $p < 0.001$] (Figure 2C).

In terms of somatosensory working memory, it was found that response accuracy decreased as a function of lag [$F(3,87) = 54.29$, $p < 0.001$], indicating that more recently experienced movements were remembered more accurately (Figure 2D). Performance at the first two lags was significantly different from zero (Bonferroni corrected, $p < 0.01$), suggesting that for this task, people could reliably maintain 2 previous movement directions in working memory. Analyses of working memory were also conducted for the visuospatial memory task ($N = 25$) which likewise yielded differences in performance across lags [$F(3,72) = 17.26$, $p < 0.001$]. In general, visuospatial working memory performance was better than that for somatosensory working memory [$F(1,24) = 106.43$, $p < 0.001$] in a manner that varied across lags [2-way interaction: $F(3,72) = 3.53$, $p = 0.018$]. In somatosensory working memory, reliable differences were observed between lag 1 and lag 2, and between lag 2 and lag 3 ($p < 0.005$), but not between lag 3 and lag 4 ($p = 1.0$). In contrast, visuospatial memory scores between lag-1 and lag-2 were found to be different ($p = 0.011$), but there was no difference in scores in the subsequent lags ($p > 0.52$).

In the forward version of the digit span test, the overall proportion of correct responses was 68.6% (*SD* = 3.1 %), whereas for the more difficult backward digit span test, the proportion correct was 60.4% (*SD* = 4.5 %). We estimated the degree of association between somatosensory working memory and the 3 other memory tasks, visuospatial working memory, forward, and backward digit span (Table 1). We found that somatosensory and visuospatial working memory scores showed a positive correlation (r = 0.43, p = 0.038, CI [0.19, 0.79]). In contrast, there was no reliable correlation between somatosensory working memory and either the forward digit span (r = 0.18, p = 0.39, CI [-0.21, 0.53]) or the backward digit span test (r = 0.33, p = 0.09, CI [0.08, 0.70]).

Somatosensory working memory performance was positively correlated with the accuracy improvement such that individuals with better memory showed greater reduction in $|PD|$ (r = 0.49, p = 0.006, CI [0.26, 0.81]) (Figure 2E). Better somatosensory working memory performance was also related to lower movement variability following learning (r = 0.49, p = 0.005, CI [0.27, 082]) (Figure 2F). Visuospatial working memory had no reliable relationship with the reduction in $|PD|$ (r = 0.12, p =0.55, CI [-0.28, 0.49]), but was positively correlated with the reduction in variability, such that individuals with higher visuospatial working memory performance had less variable movements (r = 0.65, $p < 0.005$, CI [0.30, 0.82]). Performance on the digit span tasks was not related to any of the learning measures ($r < 0.20$, $p > 0.10$).

To assess whether the relationship between the reduction in absolute error and the memory score was specific to the somatosensory modality, we conducted multiple linear regression with the reduction in error as the dependent variable and the four memory scores (somatosensory, visuospatial, and two digit-span tasks) as predictors. It was found that somatosensory working memory was able to explain the reduction in error (*p* = 0.027), but not the other predictors (*p* = 0.58 for visuospatial, *p* = 0.25 and *p* = 0.18 for forward and reverse respectively). In a second model, we used the reduction in variability as the dependent variable and found that visuospatial working memory score was a reliable predictor (*p* = 0.016), but not somatosensory working memory (*p* = 0.19), or the remaining two predictors (*p* = 0.89, *p* = 0.67 for forward and reverse digit span respectively).

It has been demonstrated previously that task-relevant baseline variability in reinforcement-based learning is able to predict the amount of learning (Wu et al., 2014). To address the concern that the correlation between working memory and motor learning was driven by differences in baseline variability, we conducted the following analysis. Baseline variability was quantified using the standard deviation of the signed *PD* during Pre (trials without feedback). After controlling for the baseline variability using partial correlation, the relationships remained significant between somatosensory working memory and reduction in absolute error (*r* = 0.43, *p* = 0.018, CI [0.20, 0.73]) and in reduction end point variability (*r* = 0.48, *p* = 0.009, CI [0.29, 0.72]).

Online Working Memory Experiment

In a second experiment, working memory trials were interleaved with motor learning trials, allowing us to test memory for movements that the participants actually performed during learning.

We first obtained behavioral measures of learning in the sensorimotor task. Because this experiment involved blocks of testing in which movements were made either to the right or the left of the workspace (with order balanced) we tested for the possibility of order effects on motor learning. The order in which participants experienced the two movement directions was not found to significantly affect the overall amount of learning in either direction as assessed using the total number of reinforced movements [$t(29) = -0.42$, $p = 0.67$], the average accuracy ($|AD|$) measured during the last block of training trials [$t(29) = -1.72$, $p = 0.1$] and during Post test [$t(29) = 0.68$, $p = 0.52$]. We examined differences in learning performance between movements in the left and right workspace. There was no reliable difference in terms of accuracy during the last block of learning [$t(29) = 1.02$, $p = 0.31$] or during the Post test [$t(29) = -1.19$, $p = 0.08$] between the two directions. We found a reliable difference in terms of total reinforced trials [$t(29) = -2.87$, $p = 0.005$], indicating that movements on the left were overall less successful ($M = 82.4$, $SD = 24.5$) than on the right ($M = 103.3$, $SD = 32.7$)., In subsequent analyses, the behavioral measures of learning for the individual participants were averaged across the two reaching directions. The mean movement distance travelled towards the hidden arc was 20.06 cm ($SD = 5.4$ cm). Because the target arc was invisible throughout training, it is possible that differences in the extent of reaching and movement speed might have an effect on the overall accuracy. Taken together the data from all participants, it was found that neither movement distance ($r = 0.02$, $p = 0.92$) nor speed ($r = 0.02$, $p = 0.07$) influenced the accuracy ($|AD|$).

Figure 3A and 3B show the movement accuracy as defined by the absolute angular deviation $|AD|$, and reinforcement rate. The angular deviation at maximum speed was significantly correlated with the angular deviation at the movement endpoint ($r = 0.82$, $p < 0.001$), as well as with the perpendicular deviation (PD) measured at movement endpoint ($r = 0.49$, $p < 0.001$). To assess whether there was learning, a linear function was fit to the absolute angular deviation over all training trials for each participant to provide an estimate of the learning slope. We took this approach rather than measuring differences between baseline and post-test movements, because in the present experiment there was no actual target defined until its direction was set on or about trial 16, based on each individual participant's movement direction. We found that the average slope across participants was significantly different than zero, indicating that the error magnitude decreased over training trials [$t(29) = -3.17$, $p < 0.01$].

Figure 3B shows that average reinforcement rate (with a bin size of 5 trials) across participants increased over the course of learning [one sample *t*-test, $t(29) = 2.17$, $p = 0.022$]. It was further found that participants that received more total reinforcement typically had less variable ($r = -0.46$, $p = 0.011$, CI [-0.70, -0.12]) and more accurate movements (smaller $|AD|$) ($r = -0.50$, $p = 0.005$, CI [-0.73, -0.21]) during Post test trials without positive feedback. In addition, participants that made more accurate movements during Post test trials also produced less variable movements ($r = 0.78$, $p < 0.0001$, CI [0.59, 0.89]). As with the offline sensorimotor task, more positive reinforcement was associated with a smaller magnitude of change in movement direction (Δm) in the next immediate trial ($r = -0.89$, $p < 0.0001$, CI [-0.91, -0.75]), consistent with the idea that positive reinforcement reduces trial-to-trial variability.

Average proportion of correct responses for lag 1 and lag 2 test were $M = 76.1\%$ ($SD = 1.9$ %) and $M = 71.8\%$ ($SD = 2.2\%$) respectively, where 50% denotes chance level. No significant difference was observed in overall working memory performance between lag 1 and lag 2 [$t(29) = 1.52$, $p = 0.16$], or between the movement direction tested first and the one that was tested second [$t(29) = 1.62, p = 0.15$]. Average memory performance for movements in the right and left workspace was $M = 77.6\%$ ($SD = 9.9\%$) and $M = 70.5\%$ ($SD = 11.7\%$) and the difference was reliable [$t(29) = 2.73$, $p = 0.02$]. We investigated whether longer/shorter movements were better remembered as follows. For each subject, all memory trials were grouped according to the extent of the seed movement using median split. A similar analysis was performed to examine whether movement speed influenced memory performance. It was found that there was no reliable difference in memory performance between seed movements that were long and short [$t(865) = 0.11$, $p = 0.43$], nor between seed movements that were fast and slow [$t(865) = 0.41$, $p = 0.96$]. This suggests the average memory score is insensitive to both movement extent and speed.

An overall measure of working memory performance was computed for each participant as the mean proportion of correct answers combining both lags and workspaces. This approach was adopted because there was no significant difference in online working memory performance between lag 1 and lag 2, or between the movement direction tested first and the one that was tested second ($p > 0.10$ respectively). Subsequently we assessed the relationship between performance during training and online working memory. The working memory score was found to be reliably associated with the total number of reinforced movements ($r = 0.47$, $p = 0.009$, CI [0.19, 082]) (Figure 3C). Participants with higher working memory scores also achieved better asymptotic performance as indicated by smaller angular deviation ($|AD|$) in the last block of training ($r = -0.41$, $p = 0.039$, CI [-0.64, -0.072]).

By interleaving the working memory task with training movements, we were able to evaluate possible differential effects on memory of making movements that successfully ended in the target zone (reinforced movements) and those that missed the target and did not receive reinforcement. Figure 3D shows the working memory score for each lag according to whether the corresponding seed movement was successful or not. It can be seen that for both lags memory was better when tests involved successful seed movements than that when seed movements were not reinforced [$F(1,29) = 6.08$, $p = 0.019$], and this was not different across lags [$F(1,29) = 0.153$, $p = 0.68$]

The presence of a working memory test interleaved between two consecutive training trials may affect the movement trial immediately following it. For example, it is possible that the movement direction deviates more from the target zone following a working memory test but in a certain angular direction. In order to investigate this possibility, we quantified the change in direction (Δm) when working memory trial intervened between trial n and $n + 1$, in cases when trial n was reinforced and not reinforced. Figure 3E shows that reaching movement deviated more following a working memory test but only when the current trial n was reinforced [2-way interaction: $F(1,29) = 33.11$, $p < 0.0001$]. Post-hoc analyses showed that the effect of the memory task on Δm was greater following a successful trial ($p < 0.001$) than an unsuccessful trial ($p = 0.351$). We also found that this additional amount of change in movement direction after a working memory test, $\Delta m_{\mathrm{WM}} - \Delta m_{\mathrm{noWM}}$, was negatively correlated with the working memory score ($r = -0.48$, $p = 0.008$, CI [-0.72, -0.16]), that is, participants who had higher working memory test scores were affected less (Figure 3F).

In the offline experiment, the capacity of the somatosensory working memory was found to be roughly 2 items. In the online experiment where participants were tested using their own movements, memory performance of each lag was reliably greater than chance level, suggesting they were able to remember two movements as well [one sample t-test, $t(29) = 11.91$, $p < 0.001$ and $t(29) = 10.51$, $p < 0.001$)]. Because the working memory test in the present study made use of the participant's own movements, it was not possible to control for the angular differences between two consecutive movements, one of which may serve as the seed movement for the working memory test. Accordingly, we asked whether participants were using the actual seed movement information, in particular, when doing the lag 2 working memory tests. To assess whether this was the case, we examined participant’s responses on the subset of lag 2 memory trials for which the correct answer would be different if participants in fact were using the lag 2 versus the lag 1 movement as the reference for their judgements. For every participant we computed the proportion of answers that matched the lag 2 reference answer and the proportion of answers that matched the lag 1 reference answer. Similar analysis was also done for the working memory lag 1 test where the wrong reference in this context was the movement performed two trials before. It was found that participant’s answers matched more often the answers that would be expected if they were basing their response on the actual seed movement than if they based it on the wrong movement (Figure 3G) [$F(1,29) = 16.87$, $p < 0.005$] and the pattern was the same for both lags [$F(1,29) = 0.99$, $p = 0.34$]. This result suggests that participants were capable of basing their answers on the correct reference (seed) movements as instructed, and not simply substituting as a basis for judgement with the wrong movements.

Finally, we assessed whether the change in direction following a working memory test was influenced by either the rotation direction used in the working memory trial (left or right) or by the direction indicated in the participants response. Figure 3H shows the pattern of the signed directional change immediately after a working memory test. The top bar shows the change in direction relative to the direction of the memory test where a positive value means that the movement direction is biased towards the direction of the rotation in the memory test. The bottom bar shows the change in direction relative to participant’s judgement. The negative value means the movement is biased in a direction opposite to participant’s verbal response. It is seen that the movement direction following the working memory trial was opposite to the participant’s judgement, regardless whether the response was correct or not [one sample *t*-test, $t(29) = 9.22$, $p < 0.005$]. This suggests that the participant’s perceptual judgment introduced a bias in planning the direction of the subsequent movement.

3.7. Discussion

The current studies demonstrated a relationship between sensory working memory and reinforcement-based motor learning. The sensorimotor learning task was based on a reinforcement learning paradigm in which participants made arm movements to unseen targets and when the movement ended within the target zone, participants received positive feedback as reinforcement to enable learning. In each experiment we observed an improvement in movement accuracy over the course of training, which was also reflected in an increase in reinforcement rate. Somatosensory working memory was assessed using participants judgements of the direction of passive displacements of the arm. In one experiment, memory tests and learning were performed separately in time. In the other, memory tests and learning trials were interleaved such that the memory tests probed the participants' memory for the movements they performed themselves in the context of learning. In both studies we found that people with better somatosensory working memory learned more. The experiment involving interleaved memory and learning trials enabled us to examine the contribution of positive feedback to working memory performance. It was found that successful trials, that is, trials that received positive reinforcement, were better remembered.

Somatosensory working memory predicts human motor learning

The term somatosensory working memory is used in the present study to refer to recognition memory and decision-making for arm configurations associated with reaching movements. Prior work in both humans and non-human primates has documented instances of working memory in the somatosensory domain. Such studies have often involved the use of tactile discrimination tasks in which, for example, one has to make judgments about the shape of an object (Stoeckel et al., 2003; Kaas et al., 2006) or to compare two sets of vibratory stimuli (Romo et al., 1999; Preuschhof et al., 2006) or to recognize patterns by tracing lines in the absence of vision (Fiehler et al., 2007).

Other studies have documented aspects of somatosensory working memory with tasks that involve limb displacement. For example, the participant's arm was passively displaced by the experimenter to a target location and the task was to reproduce the movement to the same location (Chapman et al., 2001; Goble et al., 2006; Jones and Henriques, 2010). When a delay was introduced between the passive presentation and participant's reproduction, reaching was less accurate than immediate reaching, suggesting short-term sensory memory decays over time. The present study also found that somatosensory working memory accuracy decreased for movements that were presented longer ago (at longer lags). In both experiments in the present dataset, participants could reliably retain at least two prior movements in memory.

Individual differences in somatosensory working memory performance were found to correlate with the amount of reinforcement motor learning. This is consistent with previous work demonstrating a link between sensory working memory and visuomotor adaptation (Anguera et al., 2010; Christou et al., 2016). In another demonstration of this same relationship, when subjects perform a secondary task that depletes spatial working memory capacity subsequent visuomotor adaptation is also impaired (Seidler et al., 2012). Likewise, in reinforcement learning it has been shown that the use of a secondary task impairs learning (Codol et al., 2018; Holland et al., 2018). Taken together, those findings are consistent with the idea that working memory is involved in motor learning.

Is the memory involved in motor learning specific to the somatosensory domain or is it a general memory capacity? To answer this question, we assessed whether other types of working memory might account for the individual differences in learning that were observed. To do this, a series of control tasks were used that involved visuospatial or verbal working memory. The forward and backward digit span tasks tested for the possibility that memory performance and possibly motor learning were related to verbal memory capacity. The visuospatial working memory task tested for the possibility that even though there was no explicit visual target, learning performance involved visuospatial information. Our results showed that motor learning was not related to digit span memory. In contrast, visuospatial working memory was reliably correlated with a reduction in movement variability but not to measures of improved movement accuracy. This suggests that reinforcement-based motor learning may contain several components, such as reduction of error and reduction of variance, which proceed in parallel but may be driven by different processes and thus differentially dependent on working memory. The contribution of working memory to the observed reduction in absolute movement error was specific to the somatosensory domain. The reduction in movement variability is likely to entail more general memory processes as it is reliably associated with both visuospatial and somatosensory working memory. Such domain-general memory capacities have been found in other studies, for example, in tasks that tap into both verbal and visuospatial working memory (Kane et al., 2004; Bo et al., 2011).

Working memory and positive reinforcement

In the current work, both experiments showed that movement accuracy increased over the course of learning. Positive reinforcement was shown to promote learning in terms of improvement in movement accuracy (less absolute error) and reduction in movement variability. In addition, trial-to-trial movement variability was influenced by reinforcement, with unsuccessful trials resulting in larger changes in movement direction, as was observed previously (Pekny et al., 2015; Sidarta et al., 2016).

If somatosensory working memory contributes to motor learning, are all movements equally well remembered? This issue was addressed in the present studies using a motor learning task in which at pseudorandom intervals during training, a working memory test was delivered that tested how well participants remembered their own past movements. Specifically, we presented a participants' own movement with either a rightward or leftward deflection. This online task presumably draws upon the natural learning situation where one keeps track of prior somatosensory states. By deflecting the movement, we also probed participants' ability to make a perceptual judgment associated with their own actions by comparing test displacements with the information held in the somatosensory working memory.

Somatosensory working memory scores during motor learning were found to be higher for movements that received positive feedback (Figure 3D). This finding in conjunction with the observation that working memory for movements is limited to roughly two items, suggests that the nervous system deals with this limitation by prioritizing the retention of successful movements.

Memory bias towards rewarded movements may be due to factors such as attention, saliency or the arousing effects of reinforcement that the participant received when movements were successful. This result forms part of an increasing body of evidence documenting that memory is enhanced for items or events associated with reward. Electrophysiological studies in non-human primates have found that reward influences neuronal discharge in areas of prefrontal cortex that are known to be implicated in working memory. Activity in a subset of dorsolateral prefrontal neurons was found to be modulated by reward of previously performed memory-guided saccades (Leon and Shadlen, 1999; Tsujimoto and Sawaguchi, 2004). Reward was also observed to modulate performance in a spatial memory task such that the discharge pattern of neurons in ventrolateral prefrontal cortex was associated with both spatial cues and reward (Kennerley and Wallis, 2009). The influence of reward on memory in these cases may be driven by projections of midbrain dopaminergic neurons to prefrontal cortex as shown by prior anatomical studies in non-human primates (Gaspar et al., 1992; Williams and Goldman-Rakic, 1998). Similar influences of reward on memory are found in visual and auditory working memory in humans. Performance of visual working memory is modulated by reward and activity in prefrontal cortex is correlated with reward value (Krawczyk et al., 2007; Gong and Li, 2014; Klink et al., 2017). Based on functional connectivity analyses involving auditory cortex, prefrontal cortex, and ventral striatum, pleasurable music is thought to be encoded more strongly in auditory working memory (Zatorre and Salimpoor, 2013).

Previously, Pekny et al. (2015) and Holland et al. (2018) investigated trial-to-trial changes in movement direction in a reinforcement learning task as a function of the history of prior rewards. In each paper the authors computed the difference in movement direction between a particular trial and the immediately preceding trial, as a function of the sequence of rewards going back as far as three preceding trials. It was found that the memory for a sequence of rewards influences the change in movement direction on the current trial. It is possible that the change in direction is due to memory decay with increasing distance from the last successful trial, which would be consistent with the lag effect on memory seen in the present study (Figure 2D). An alternative possibility is that the increase in variability following a string of unsuccessful movements reflects an exploration strategy, or a combination of memory decay and exploration.

It was observed that the online working memory tests presented during learning increased the variability of the next reaching movement, but people with better working memory performance were less affected. To better understand the nature of this effect, we examined whether the subsequent change in movement direction followed the direction of the rotation introduced by the robot or the direction indicated in the participant’s verbal response. We found that the reaching direction did not shift toward the rotation direction presented in the memory test. Instead, it was biased in a direction opposite to participants’ judgements, which may indicate an attempt to correct for the presumed direction of rotation (Figure 3H). Because the online task required participants to make a perceptual judgment, this finding is consistent with the idea that the perceptual judgment appeared to introduce a bias in planning the next immediate movement. Such finding may be due to a top-down influence on the motor system by the prefrontal neurons (Cisek, 2007).

Recent work in reinforcement-based motor learning has focused on the involvement of awareness, exploration, and explicit processes (Manley et al., 2014; Pekny et al., 2015; Therrien et al., 2015; McDougle et al., 2016; Cashaback et al., 2017; Codol et al., 2018; Holland et al., 2018), both in clinical and healthy population, as well as in learning-related brain plasticity (Sidarta et al., 2016). Several of these studies incorporate a reward zone that shifts either gradually (Pekny et al., 2015; Holland et al., 2018) or dynamically based on the performance in previous trials (Therrien et al., 2015). In contrast, the size and position of the reinforced direction in the current study was fixed. The advantage of using a constant target width is the consistency of the environment (and task demand) across blocks, and accordingly, the ability to interpret differences in variables such as movement accuracy or the number of reinforced trials with respect to a common reference.

It is possible that the role of working memory is different for a fixed versus rotated reward zone. Specifically, previous work showed spatial working memory capacity predicts explicit visuomotor adaptation (Christou et al. 2016), and that learning to aim to a shifting reward zone is dominated by explicit processes (Holland et al 2018, Codol et al 2018). Therefore, one might expect that a shifting reward zone involves visuospatial working memory components. However, in the present paradigm there was an average initial bias prior to learning (average magnitude = 1.57 cm), and the magnitude of the bias significantly reduced following learning. In order to correct for the bias participants may rely on spatial rotation, which implies this fixed target paradigm may share some features with a shifting reward zone.

Potential neural bases of somatosensory working memory in human motor learning

While previous studies have suggested that the dorsolateral prefrontal cortex is implicated in visuospatial working memory in human motor learning (Seidler et al., 2012), the neural substrates of somatosensory working memory for limb configuration in relation to motor learning are less certain. Using a sensorimotor learning task similar to that in the present study, it was found that learning-related changes in resting-state function connectivity involved second somatosensory cortex (SII), the right supramarginal gyrus, and right BA 9/46v (Sidarta et al., 2016). These areas are known to be inter-connected anatomically. Specifically, studies in non-human primates have identified bi-directional projections linking areas PF and PFg in the inferior parietal lobe (supramarginal gyrus in humans), the parietal operculum, and ventral area 46 below the principal sulcus (Preuss and Goldman-Rakic, 1989; Petrides and Pandya, 2002; Gerbella et al., 2012). Indeed area 9/46v is the only region in prefrontal cortex to project to hand area structures in cortical motor areas, specifically to ventral premotor cortex and pre-SMA (Luppino et al., 1993; Lu et al., 1994). In other primate studies, second somatosensory cortex, ventral premotor cortex, supplementary motor area, and the lateral prefrontal cortex have been shown to be involved somatosensory memory and decision-making tasks involving vibrotactile stimuli (Romo et al., 1999; Romo et al., 2002; Romo et al., 2004). These findings suggest that lateral prefrontal cortex may be involved in online guidance of reaching movements (Goldman-Rakic, 1996) by providing motor areas with sensory information stored in working memory.

One potential limitation of the current work is that baseline assessment of participants' somatosensory acuity was not performed. It is possible that the perceptual acuity may have an effect on our estimate of memory performance. Nevertheless, regardless of whether memory items were spaced far apart (10 degree spacing, as in the offline task) or close together (5 degrees, as in the online task), memory performance was able to predict learning. This suggests that the memory tasks capture a type of memory that is largely invariant to spatial scale. Presumably, this in turn makes it less likely that somatosensory acuity influences working memory estimates.

In conclusion, two experiments are presented here that provide evidence for the idea that somatosensory working memory supports reinforcement-based motor learning in humans. In the future, it would be desirable to directly modulate neural activity in areas in frontal and prefrontal cortex that contribute to working memory to assess their individual contributions to human motor learning.

Table 1: Correlation coefficient (*r*) between somatosensory working memory and other measures of working memory.

* means $p < 0.05$, *** $p < 0.001$

Task	1. SWM	2. VSWM	3. DSf	4. DSb
1. Somatosensory WM (SWM)	-	0.43*	0.18	0.33
2. Visuospatial WM (VSWM)		-	0.32	0.23
3. Digit span (forward) (DSf)			-	0.71***
4. Digit span (backward) (DSb)				-

Acknowledgments:

This work was supported by the National Institute of Child Health and Human Development R01 HD075740, Les Fonds Québécois de la Recherche sur la Nature et les Technologies, Québec (FQRNT), and the Natural Sciences and Engineering Research Council of Canada (NSERC). We thank David St-Amand and Katrine Bergeron for their assistance.

Disclosures:

No conflicts of interest, financial or otherwise, are declared by the authors.

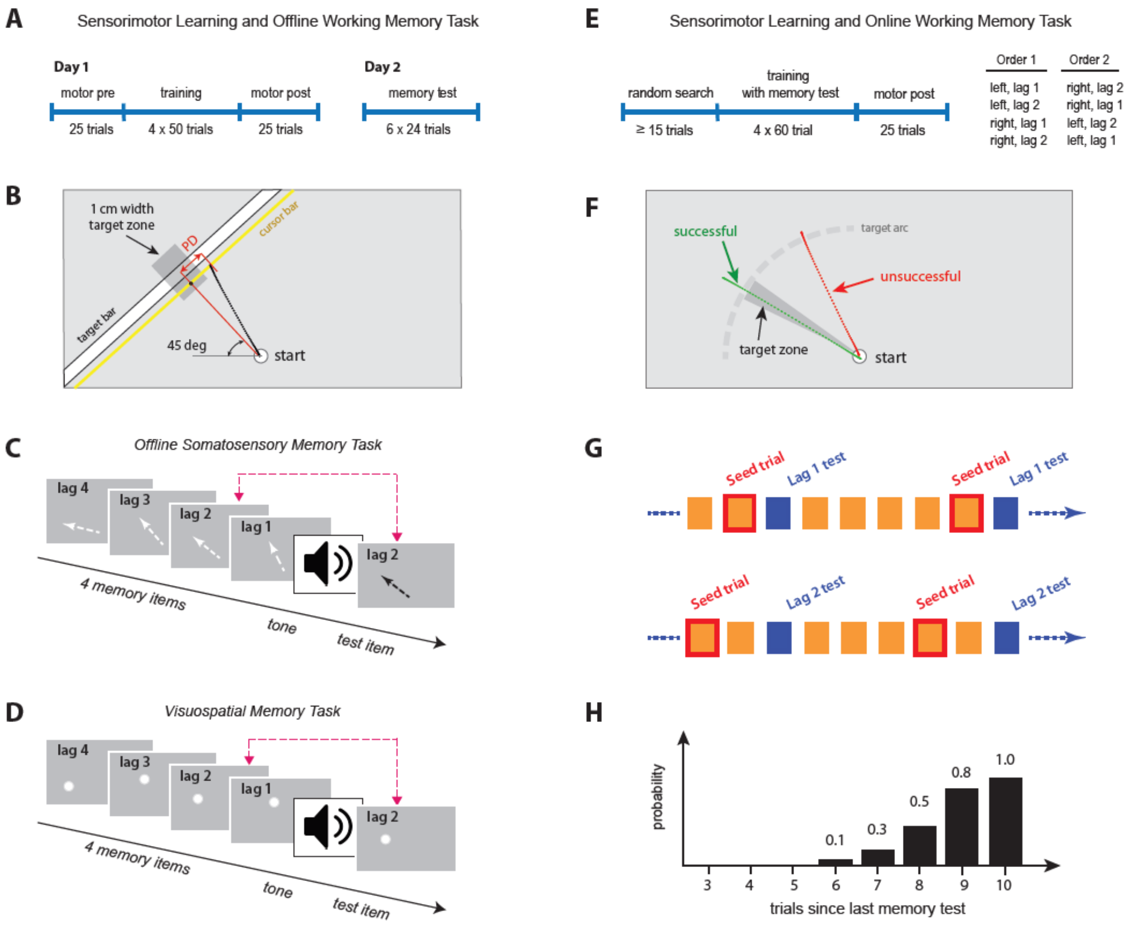
Figure captions:

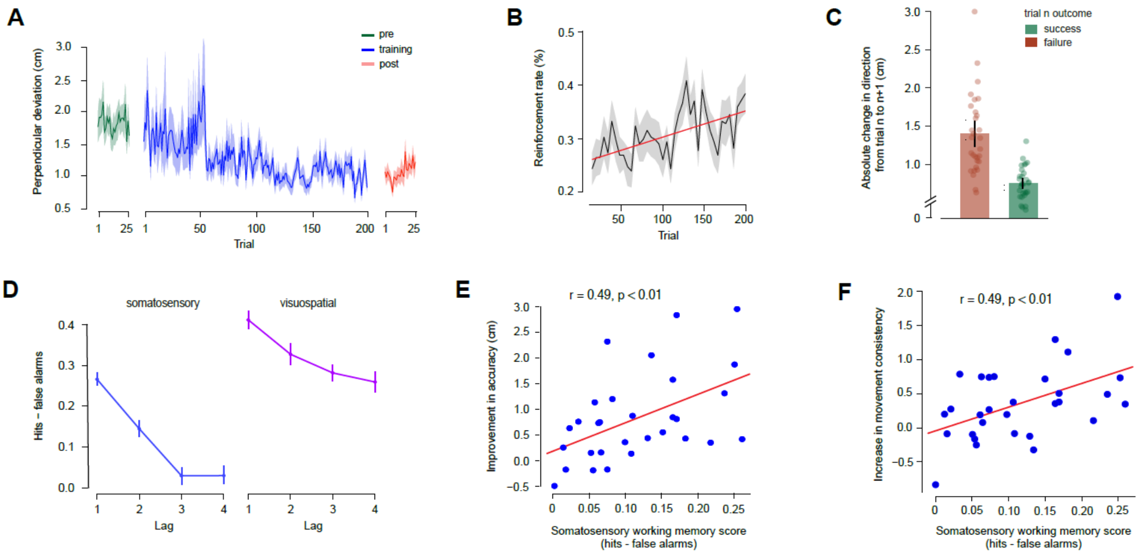
Figure 1: Experimental setup. (A) Offline working memory testing sequence. (B) Participants learned to make movements to an unseen visual target (grey). Movement accuracy was quantified as the lateral perpendicular deviation (*PD*) at the movement end point. (C) Somatosensory working memory test design. (D) Visuospatial working memory test design. (E) Online working memory testing sequence. (F) Online sensorimotor learning with examples of successful (green) and unsuccessful (red) movements. A movement was successful when it fell within 2.5 degrees of the target direction (shown in grey). (G) Online working memory test trials (blue) were interleaved with the sensorimotor training trials (orange). Each working memory trial was based on a previous movement (seed trial, red). (H) Working memory test trials were placed according to a predefined probability distribution, which was a function of the number of trials since the last working memory test.

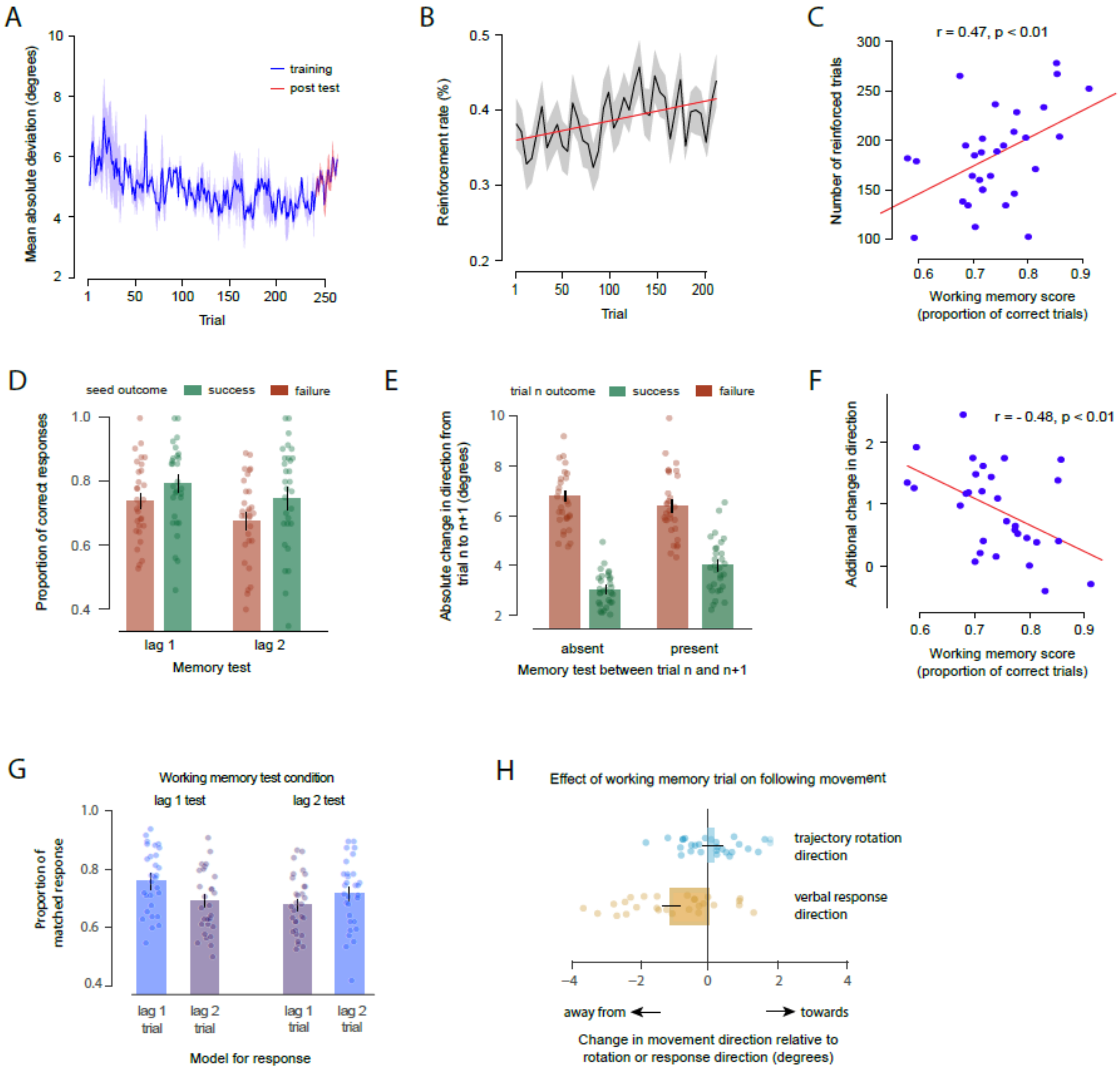
Figure 2: Results of the offline working memory experiment. (A) Movement accuracy as measured by the perpendicular deviation at movement end point $|PD|$ in three different task phases (pre, training, and post). (B) Reinforcement rate averaged across participants (bins of 5 trials) shows an increase in the number of successful trials over the course of training. (C) There was a larger change in movement direction following unsuccessful trials. (D) Somatosensory and visuospatial working memory decreases for movements that were presented longer ago (increasing lag). (E) Somatosensory working memory

performance was positively correlated with improvement in accuracy in the motor learning task. Participants who had better working memory learned more. (F) Somatosensory working memory performance was positively correlated with reduction in end point variability.

Figure 3: Online working memory experiment. (A) Participants showed learning as indicated by a reduction in angular deviation, |*AD*|. (B). Reinforcement rate increased over trials. (C) Participants with better working memory learned more as shown by a greater number of reinforced trials. (D) Successful movements were remembered better. (E) The presence of a working memory test resulted in an increase in directional change in the following trial but only when the preceding trial was successful. (F) Participants with better working memory show smaller changes in movement direction following memory tests, indicating that the working memory trials disrupted the learning less. (G) Participants correctly based their memory test responses on the appropriate seed (reference) movement. For each participant, model responses were computed based on lag 1 or lag 2 movements and these were then matched to participant’s actual responses (indicated by the bar height). (H) Following a working memory test, movements were biased away from the participants' judgments (verbal responses) of the memory test direction but were not affected by the actual trajectory rotations.


A
Sensorimotor Learning and Offline Working Memory Task
Day 1
motor pre
training
motor post
25 trials
4 x 50 trials
25 trials
Day 2
memory test
6 x 24 trials
E
Sensorimotor Learning and Online Working Memory Task
random search
training with memory test
motor post
≥ 15 trials
4 x 60 trial
25 trials
Order 1
left, lag 1
left, lag 2
right, lag 1
right, lag 2
Order 2
right, lag 2
right, lag 1
left, lag 2
left, lag 1
B
1 cm width target zone
PD
cursor bar
target bar
45 deg
start
F
successful
target arc
unsuccessful
target zone
start
C
Offline Somatosensory Memory Task
lag 4
lag 3
lag 2
lag 1
lag 2
4 memory items
tone
test item
G
Seed trial
Lag 1 test
Seed trial
Lag 1 test
Seed trial
Lag 2 test
Seed trial
Lag 2 test
D
Visuospatial Memory Task
lag 4
lag 3
lag 2
lag 1
lag 2
4 memory items
tone
test item
H
probability
0.1
0.3
0.5
0.8
1.0
3
4
5
6
7
8
9
10
trials since last memory test


A
pre
training
post
Perpendicular deviation (cm)
3.0
2.5
2.0
1.5
1.0
0.5
1
25
1
50
100
150
200
1
25
Trial
B
Reinforcement rate (%)
0.5
0.4
0.3
0.2
50
100
150
200
Trial
C
trial n outcome
success
failure
Absolute change in direction from trial n to n+1 (cm)
3.0
2.5
2.0
1.5
1.0
0
D
somatosensory
visuospatial
Hits - false alarms
0.4
0.3
0.2
0.1
0.0
1
2
3
4
Lag
1
2
3
4
Lag
E
r = 0.49, p < 0.01
Improvement in accuracy (cm)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
0
0.05
0.10
0.15
0.20
0.25
Somatosensory working memory score (hits - false alarms)
F
r = 0.49, p < 0.01
Increase in movement consistency
2.0
1.5
1.0
0.5
0.0
−0.5
0
0.05
0.10
0.15
0.20
0.25
Somatosensory working memory score (hits - false alarms)


A
Mean absolute deviation (degrees)
training
post test
Trial
B
Reinforcement rate (%)
Trial
C
r = 0.47, p < 0.01
Number of reinforced trials
Working memory score
(proportion of correct trials)
D
seed outcome
success
failure
Proportion of correct responses
lag 1
lag 2
Memory test
E
trial n outcome
success
failure
Absolute change in direction from trial n to n+1 (degrees)
absent
present
Memory test between trial n and n+1
F
r = - 0.48, p < 0.01
Additional change in direction
Working memory score
(proportion of correct trials)
G
Working memory test condition
lag 1 test
lag 2 test
Proportion of matched response
lag 1 trial
lag 2 trial
lag 1 trial
lag 2 trial
Model for response
H
Effect of working memory trial on following movement
trajectory rotation direction
verbal response direction
away from
towards
Change in movement direction relative to rotation or response direction (degrees)

Supplementary Materials

The following supplementary analyses which were not included in the original manuscript focus on the evolution of trial-to-trial change in movement direction (Δm) over the course of learning in both the offline and online procedures. The final set of analyses are for the online working memory procedure and examine whether there was any change in terms of working memory performance across learning.

1. Trial-to-trial change in direction

In the previous chapter, average trial-to-trial movement variability did not change significantly over the course of learning. If the task demand and the environment remain the same, one theoretically would no longer need to continue exploring the space as one would get more reinforced movements near the end of learning. To test this hypothesis, the change in direction (Δm) following successful and unsuccessful trial was averaged in each of the early and late phases of training. As in the previous chapter, the early and late phases were defined as the first and the last 25 training trials in each block of training. Figure S1-A and S1-B show change in direction following successful (blue) and unsuccessful trials (red) during motor learning in the offline and online procedures, respectively. It can be seen that the amount of directional change progressively declines over the course of learning. This reduction was found to be reliable in the offline procedure [$F(7,203) = 3.45$, $p < 0.05$ with Greenhouse-Geiser correction], as well as in the online procedure [$F(7,203) = 4.81$, $p < 0.01$ with Greenhouse-Geiser correction].

The result suggests that exploratory behavior becomes less necessary to achieve reward as movements get more accurate nearer to the end of training. In contrast with the previous Chapter 2, the constant reward zone used in the current experiment means that the environment and task demands remain the same. This points to the idea that subjects are able to voluntarily reduce trial-to-trial change in directional if the task demand remains fixed.

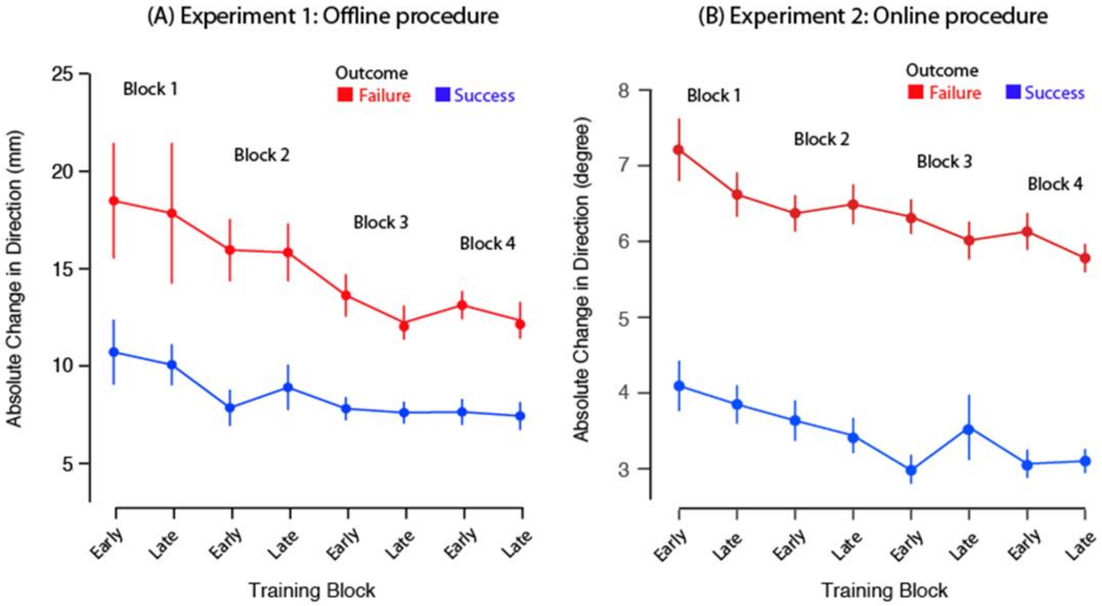


Figure S1: (A) Trial-to-trial change in direction in the early and late phase of each training block in the offline procedure. (B) Same as (A) but in the online procedure. In both panels, red and blue curves denote change in direction following unsuccessful and successful trial respectively.

2. Online memory performance

The role of working memory in motor learning may not be constant. One possibility is that trial-to-trial memory is utilized more in early stages of learning, whereas later on, participants can instead simply repeat prior actions. We tested for this possibility by assessing in the online procedure whether there were changes in working memory performance over the course learning. To do that, the memory score (proportion of correct responses) was computed for each block, depending on the outcome of the reference or seed movement. It was found that there was no significant change in the performance across blocks [$F(3,87) = 0.01$, $p > 0.5$] (Figure S2-A), but the there was a reliable difference in memory score in terms of whether the seed movement was successful, and thus rewarded. This is line with the idea that somatosensory working memory in our task is equally important throughout learning.

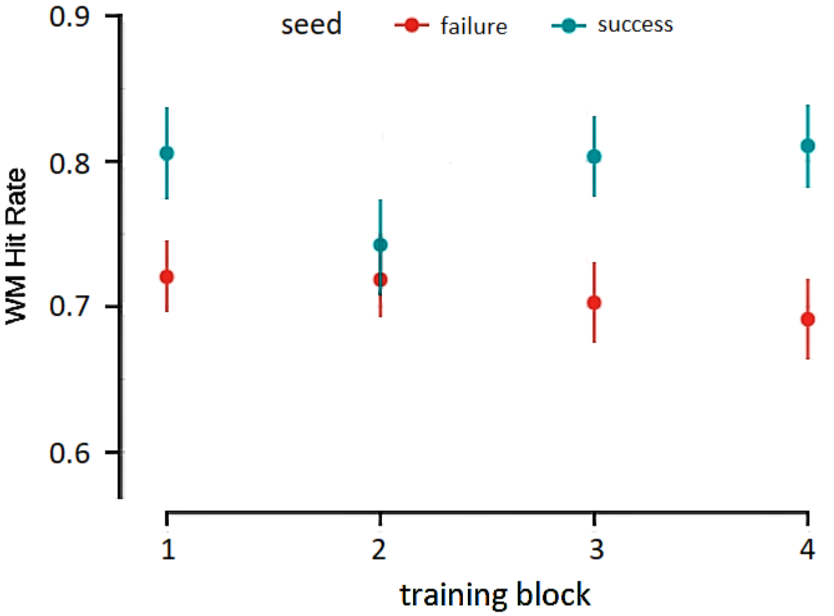


Figure S2: Proportion of correct responses (hit rate) as a function of training block in the online working memory experiment. The seed movement is the reference movement produced by the subject during training where at times it can be either successful or unsuccessful.

Chapter 4 : General discussion

4.1. Introduction

The main focus of this thesis are on the initial stages of motor learning in tasks where visual information is limited. Study-1 involved a combination of behavioral and neuroimaging procedures with the aim of understanding brain plasticity associated from this type of motor learning. This study also extends the previous findings in Bernardi et al. (2015) in understanding the behavioral mechanisms behind reinforcement-based motor learning. In Study-2, the contribution of somatosensory working memory to motor learning was examined using two different experiments (offline and online procedures). The offline procedure assessed whether individual differences in motor learning outcomes were related to differences in working memory capacity. The online procedure measured working memory of one's own movements on a trial-by-trial basis directly during learning. This experiment assessed whether successful and unsuccessful movements were equally remembered. Overall, both positive feedback and somatosensory working memory have been shown to be important for motor learning and changes in functional networks in specific brain regions associated with this type of learning have been identified. This chapter will address the findings of the present studies in relation to general concepts in motor learning, and will identify some potential limitations that would be addressed in future studies.

4.2. Motor learning through reinforcement

Humans have a remarkable capability to adapt and learn new motor skills and there are many varieties of motor learning in daily lives. While visual targets have been widely studied as the main sensory source of motor control and learning, there are instances where visual information is occluded. In order to appreciate this kind of situation, we can draw upon situations such as when a novice learns to swim or dance. While it can be argued that the novice sees the external environment to which he has to go (visual information), the information arising from such sensory input contributes only partially as a learning tool. Instead, one has to learn the arm (or leg) configuration that is necessary to successfully perform the task. Likewise, discovering the most appropriate arm configuration is also required while learning to play golf or performing a javelin throw. In each case, the novice must discover the appropriate somatosensory state, or more technically the targets or goals, during motor learning. Exposing subjects to somatosensory information alone, however, is not sufficient. As Bernardi et al. (2015) showed, passive movement (yoked to the movements of an active participant) but without reinforcement provided limited benefit to improving movement accuracy following training. The combination of positive feedback (as reinforcement) and somatosensory information obtained from passive movement was sufficient to produce learning comparable in magnitude to that obtained when subjects trained actively.

The mechanisms behind motor learning in this thesis can be explained in terms of reinforcement learning. According to reinforcement learning concepts, a novice (or agent) learns a certain task (in an environment) by exploring different actions. Each action is associated with two types of information: relevant states (somatosensory feedback) and an outcome (positive feedback, reward). The novice then learns to assign credit or value based on the outcome of one's actions, that will eventually lead to the "best" possible action-value choices (Sutton and Barto, 1998; Daw and Doya, 2006; Gershman et al., 2009). In the tasks used here, positive feedback is given when the movement ends within a specific target zone. Somatosensory states associated with this rewarded actions define the subsequent somatosensory targets or goals.

In Chapter 1, it was stated that in the absence of visual target information, error-based learning is not well suited to model the earliest stages of motor learning. This is because sensory targets relative to which error might be computed are still in the process of being discovered. This, however, does not mean that the mechanism behind motor learning is exclusively reinforcement-based. Even with imprecise targets, motor adaptation (error-based learned) can precede or occur in conjunction with reinforcement learning. In swimming, for example, one has to adapt their limbs to the pressure under water. In the javelin throw, while finding the best arm configuration is largely driven by reinforcement, one arguably has to adapt to the weight of the spear and is guided by a distant visual goal.

4.3. Behavioral signatures of reinforcement learning

It has been proposed that the goal of reinforcement learning is to find actions that maximize future reward through trial and error and exploration (Sutton and Barto, 1998). Successful actions are expressed more often as a result of reward or reinforcement, which eventually lead to more accurate movements. In support of this idea, I have shown that the rate of reinforcement increases over the course of learning, and that positive feedback yields a reliable reduction in movement error in all motor learning tasks described in this thesis. Moreover, subjects who had more accurate movements following learning were also less variable. The reduction in movement variability varies in proportion to the total number of reinforced trials.

This thesis also examined exploratory behavior in search of more rewarding actions by looking at differences (or variability) in movements from one trial to the next, e.g. using trial-to-trial change in reaching direction (Δm). Consistent with the idea that exploration follows unsuccessful outcomes, one sees a progressive increase in trial-to-trial change in direction (see Chapter 2, Figure 2D). In Study-1, the task difficulty varied between blocks as a result of changes in the width of the reward zone. The magnitude of Δm remained relatively high throughout training. In contrast, in Study-2 a constant reward zone width was used for both the offline and online procedures. When trained under this type of environment, there was a reliable reduction in Δm over training blocks. When the task difficulty remained the same across blocks, the chance of getting positive feedback, or more technically reward probability in the environment did not change. In such situation, the need to explore the environment subsequently diminished, resulting in a gradual reduction of the Δm. Interestingly, the gradual reduction in Δm occurred following both successful and unsuccessful movement. This was most likely due to the reinforcing effect of positive feedback. As training continued, participants became more skilful in moving in the desired reaching direction and received more positive feedback. Following unsuccessful trial, they could more easily re-discover the successful direction without having to sample a wider area. Similarly, following successful trial, trial-to-trial variability reduced as a result of more precision resulting from an increasing amount of positive feedback. These results are in line with the findings observed in other motor learning studies. For example, Pekny and colleagues observed similar trial-to-trial movement variability when studying the sensitivity of one's behavior to reward. The authors showed that healthy adults were able to modulate the amount of exploration in the face of changing reward conditions (Pekny et al., 2015). On the other hand, the fact that reward influences trial-to-trial change in motor behavior is in contrast with error-based learning where it is the magnitude of error of one trial that is associated with the amount of correction in another trial (Kawato et al., 1987; Jordan and Rumelhart, 1992; Fine and Thoroughman, 2007).

Recently, the importance of movement variability in motor learning has been emphasized by findings of Wu and colleagues (Wu et al., 2014) who found that people with more variable movements in baseline are able to learn more, both in error-based as well as reinforcement-based tasks. Similar behavior has been observed in animal studies involving songbirds (Ölveczky et al., 2005; Tumer and Brainard, 2007; Woolley et al., 2014), where reinforcement learning is used to model vocal production. These studies found that juvenile songbirds possess greater vocal variability than adult birds, consistent with the idea of motor exploration in reinforcement learning. The putative neural substrate behind such behavior is thought to be the LMAN - Area X circuit, the output part of the cortico-basal ganglia circuit in songbirds.

4.4. Somatosensory working memory in motor learning

In a typical scenario in which learning involves exploration and trial and error, movements sometimes can of course be successful and not successful. In Chapter 1, it was hypothesized that under these conditions working memory is presumably critical for learning. Chapter 3 (Study-2) assessed the properties of somatosensory working memory and its role in motor learning. In both experiments, it was shown that people with better working memory learn more. With the offline procedure, the capacity of somatic working memory as measured using the current procedures was found to be around two items. Conversely, in a matched visuospatial control task, memory capacity was estimated to be at least four items. Both working memory systems exhibit a performance decay as a function of lag, that is, memory performance of an item drops as a function of the item's sequential position in the list.

The capacity limit of human working memory has been a topic of investigation and discussion for years. Recently, Cowan posited that the capacity of short-term memory is four items (Cowan, 2001) after taking into account feature complexity and the role of attention. In another study by Gallace and colleagues, focused directly on the somatosensory system, the average capacity of tactile working memory was estimated to be between 3 and 5 items, which is similar to Cowan's data (Gallace et al., 2008). It is likely that the discrepancy in the memory capacity found here is due to the nature of the memory task in the offline procedure. Although the smallest difference between each direction in this procedure was 10° [higher than ~3°, the minimum detectable change in angle as measured by Goble and colleagues (Goble et al., 2005)], the presence of multiple items in the memory list and the way they were presented to the subjects might affect this capacity. This thesis has not attempted to resolve the ambiguity in the issue of the actual capacity of the somatosensory working memory. Nevertheless, it has successfully shown that even for potentially difficult somatic memory tasks (which are presumably representative of the memory demands for motor learning), individual differences in memory performance explain the extent of learning.

The involvement of working memory directly during learning was investigated in the online experiment of Chapter 3. In the procedure, the capacity of memory tested during learning was shown to be at least 2, consistent with the offline procedure. This suggests that the role of working memory in reinforcement learning may operate from moment to moment, that is, over a small number of consecutive movements. This is a reminiscent of a similar finding in error-based adaptation where error experienced in one trial predicts adaptation in the following trial (Scheidt et al., 2001; Herzfeld et al., 2014).

Although at first glance the finding that working memory is better for successful movements, the finding by no means implies that working memory is not involved following unsuccessful trials. One limitation of the current study is that we are unable to tease apart the differential roles of working memory during early (high exploratory stage) and late phase (or reward accumulation stage) of learning, as there was neither drastic nor gradual change in the memory performance over the course of learning (Supplementary Materials, Chapter 3). While this result suggests that somatosensory working memory is important throughout learning, the alternative hypothesis may still hold. Working memory may be important at the beginning but its role fades away with learning. Indeed, there is a dynamic interaction between pure reinforcement learning which is reward-driven and working memory which is capacity-limited (Collins and Frank, 2012). Using a mixture model of reinforcement learning and working memory, the authors suggest that with increasing experience the reinforcement learning model wins over the working-memory model.

Findings in Study-2 have some potential limitations that are related to the type of information stored in the working memory, namely somatosensory information. An alternative possibility is that working memory stores a history of motor commands instead. Although the second study did not directly address this question, there is evidence suggesting that the memory stores somatosensory states. In the offline procedure, in which a passive task is used to eliminate any form of an efferent outflow (i.e. motor commands), it is found that somatosensory working memory reliably predicts the amount of motor learning. This is consistent with the idea that somatosensory working memory mechanism is potentially involved in motor learning. What somatosensory states are encoded in the memory during learning also requires further elaboration. For example, it is possible the somatosensory states associated with the whole movement trajectory is stored in the working memory. An alternative hypothesis will be that the memory only stores the limb configuration at the movement end point. At the moment, the current task design is not able to distinguish these alternatives.

4.5. Neural signatures of learning as revealed by neuroimaging

In the study presented in Chapter 2, functional connectivity of the resting brain was used to identify the brain areas involved early in motor learning. Regions of the brain implicated in this type of learning include the sensorimotor network [primary motor and somatosensory cortices, supplementary motor area (SMA), and second somatosensory area (SII)] and the reinforcement network (ventral striatum and ventromedial prefrontal cortex or vmPFC). The finding that plasticity in the ventromedial prefrontal cortex (vmPFC) varied in proportion to the number of successful trials is in line with the idea that more complex processes are involved during learning, that is, factors beyond basic sensorimotor function and reinforcement. Consistent with earlier studies, it can be said that early learning involves encoding stimulus-reward value and decision-making that involves selecting actions that are more rewarding (O'Doherty et al., 2003; Rushworth et al., 2004; Daw et al., 2006; Dayan and Niv, 2008). Plasticity in the associated brain areas may indicate that the value of making successful movements is comparable to getting other types of reward, e.g. monetary gain. As mentioned in Section 4.2, reinforcement-based motor learning in real life does not occur in isolation. Thus, basal ganglia involvement may be complemented by other brain areas such as cerebellum, for example, when a form of motor adaptation is involved. Indeed, both areas can complement each other during motor learning (Doya, 2000)

Changes in functional connectivity that were associated with the extent of exploration were also examined. One of the changes was found in the sensorimotor network involving SMA and subcortical network (basal ganglia and thalamus). Connectivity change was found in proportion to the average trial-by-trial change in direction (Δm) regardless whether the outcome of the previous trial was successful or unsuccessful. Subjects that displayed greater amounts of exploration showed larger reduction in functional connectivity. In turns, strengthening in connectivity was observed between SII and primary sensorimotor cortices that varied in proportion with the amount of exploration. It has been shown that SMA receives anatomical projections from the output of basal ganglia (Akkal et al., 2007). It is plausible that exploration-related decrease in functional connectivity observed here is linked to the role of basal ganglia in modulating movement variability, a phenomenon that has been demonstrated in song birds (see Section 4.2 above). More specifically, the way basal ganglia introduces variability to the motor system is through a decrease in functional coupling between putamen and motor cortices. At the same time, increases in functional coupling between SII and other sensorimotor cortices suggest that this network accompanies exploratory behavior, probably in order to construct somatosensory targets. It should be noted, however, that due to the nature of the rs-fMRI methodology, it is impossible to assess trial-to-trial changes in activity during the task performance.

Changes in connectivity were also observed between a seed in left SII and a somatic network which includes the right ventrolateral prefrontal cortex (BA 9/46v) and right supramarginal gyrus. The result indicated that people with greater changes in connectivity had a smaller trial-to-trial change in direction (a measure of exploration) following an unsuccessful trial. There is evidence in monkey literature that suggests the involvement of this similar network (Romo et al., 1999; Romo et al., 2002), with exception supramarginal gyrus (Area 40) which was observed to have learning related connectivity change in the present study With functional MRI in humans, prior evidence revealed that in addition to lateral prefrontal and SII, supramarginal gyrus is also implicated in the tactile recognition and discrimination tasks (Stoeckel et al., 2003; Kitada et al., 2005; Reed et al., 2005). Owing to the fact that this network is somatic and has been associated with tactile discrimination and decision-making tasks, it is reasonable to hypothesize that the network involving SII, right lateral prefrontal cortex, and right supramarginal gyrus is the putative neural substrate of somatosensory working memory in reinforcement-based motor learning.

4.6. Conclusion and Future Work

The work presented here studied motor learning mechanism using reinforcement-based learning. Plasticity associated with this type of learning was found to involve two separate networks, a sensorimotor network and a reward-related network. The involvement of somatosensory working memory in learning was assessed in two experiments. It was found that people with better memory were able to improve more after motor learning and obtained more reinforced trials during training. Somatosensory working memory contains quite a limited capacity (two items in the present studies), which is consistent with the idea that it contributes to learning over a short time scale. Lastly, it has been shown that reinforced trials during training are better remembered than non-reinforced trials.

The findings described in this thesis require more investigation. In Chapter 2, it was shown that the right lateral prefrontal cortex was implicated in the motor learning task. In Chapter 3, somatosensory working memory was shown to play a key role in reinforcement learning. This statement faces two limitations. As mentioned in Section 4.4, it is necessary to examine whether somatosensory working memory only encodes certain aspects of the movement (end position, whole movement trajectory, etc). Another limitation is the behavioral significance of the somatic network involving right BA 9/46v in trial-to-trial movement exploration during reinforcement learning. Therefore, it is necessary to further examine the neural substrates of somatosensory memory in reinforcement-based motor learning directly. One may use task-based functional MRI with an appropriate behavioral manipulation to test this. Another simpler way is to modulate the activity of neurons in putative working memory areas such as right Area

9/46v, SII and SMA using, transcranial magnetic stimulation (TMS). It has been previously shown that a TMS protocol called *continuous theta burst stimulation* is able to create reversible neural inhibition for at least 90 minutes (Goldsworthy et al., 2012). The hypothesis is that inhibiting neurons in the above-mentioned areas would have an adverse effect on memory and reduce trial-to-trial movement variability. This eventually will lead to a disruption in motor learning.

References

Abe M, Schambra H, Wassermann EM, Luckenbaugh D, Schweighofer N, Cohen LG (2011) Reward improves long-term retention of a motor memory through induction of offline memory gains. Current Biology 21:557-562.

Akkal D, Dum RP, Strick PL (2007) Supplementary motor area and presupplementary motor area: targets of basal ganglia and cerebellar output. Journal of Neuroscience 27:10659-10673.

Albert NB, Robertson EM, Miall RC (2009) The resting human brain and motor learning. Current Biology 19:1023-1027.

Anguera JA, Reuter-Lorenz PA, Willingham DT, Seidler RD (2010) Contributions of spatial working memory to visuomotor learning. Journal of cognitive neuroscience 22:1917-1930.

Atkinson RC, Shiffrin RM (1968) Human memory: A proposed system and its control processes1. In: Psychology of learning and motivation, pp 89-195: Elsevier.

Baddeley A (1992) Working memory. Science 255:556-559.

Baddeley A, Gathercole S, Papagno C (1998) The phonological loop as a language learning device. Psychological review 105:158.

Baddeley AD, Hitch G (1974) Working memory. In: Psychology of learning and motivation, pp 47-89: Elsevier.

Balédent O, Idy-peretti I (2001) Cerebrospinal fluid dynamics and relation with blood flow: a magnetic resonance study with semiautomated cerebrospinal fluid segmentation. Investigative radiology 36:368-377.

Balslev D, Miall RC, Cole J (2007) Proprioceptive deafferentation slows down the processing of visual hand feedback. Journal of Vision 7:12-12.

Beckmann CF, Smith SM (2004) Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Transact, Medical Imaging 23:137-152.

Bernardi NF, Darainy M, Ostry DJ (2015) Somatosensory contribution to the initial stages of human motor learning. Journal of neuroscience 35:14316-14326.

Bernier P-M, Grafton ST (2010) Human posterior parietal cortex flexibly determines reference frames for reaching based on sensory context. Neuron 68:776-788.

Berns GS, McClure SM, Pagnoni G, Montague PR (2001) Predictability modulates human brain response to reward. The Journal of Neuroscience 21:2793-2798.

Bhadelia R, Bogdan A, Kaplan R, Wolpert S (1997) Cerebrospinal fluid pulsation amplitude and its quantitative relationship to cerebral blood flow pulsations: a phase-contrast MR flow imaging study. Neuroradiology 39:258-264.

Birn RM, Diamond JB, Smith MA, Bandettini PA (2006) Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. Neuroimage 31:1536-1548.

Bischoff-Grethe A, Hazeltine E, Bergren L, Ivry RB, Grafton ST (2009) The influence of feedback valence in associative learning. Neuroimage 44:243-251.

Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine 34:537-541.

Bo J, Seidler RD (2009) Visuospatial working memory capacity predicts the organization of acquired explicit motor sequences. Journal of neurophysiology 101:3116-3125.

Bo J, Jennett S, Seidler R (2011) Working memory capacity correlates with implicit serial reaction time task performance. Experimental brain research 214:73.

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

Bull R, Espy KA, Wiebe SA (2008) Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. Developmental neuropsychology 33:205-228.

Burton H, Fabri M (1995) Ipsilateral intracortical connections of physiologically defined cutaneous representations in areas 3b and 1 of macaque monkeys: projections in the vicinity of the central sulcus. Journal of Comparative Neurology 355:508-538.

Bütefisch CM, Davis BC, Wise SP, Sawaki L, Kopylev L, Classen J, Cohen LG (2000) Mechanisms of use-dependent plasticity in the human motor cortex. Proceedings of the national academy of sciences 97:3661-3665.

Carel C, Loubinoux I, Boulanouar K, Manelfe C, Rascol O, Celsis P, Chollet F (2000) Neural substrate for the effects of passive training on sensorimotor cortical representation: a study with functional magnetic resonance imaging in healthy subjects. Journal of Cerebral Blood Flow & Metabolism 20:478-484.

Carp J (2013) Optimizing the order of operations for movement scrubbing: Comment on Power et al. Neuroimage 76:436-438.

Cashaback JG, McGregor HR, Mohatarem A, Gribble PL (2017) Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. PLoS computational biology 13:e1005623.

Chapman C, Heath M, Westwood D, Roy E (2001) Memory for kinesthetically defined target location: evidence for manual asymmetries. Brain and cognition 46:62-66.

Christou AI, Miall RC, McNab F, Galea JM (2016) Individual differences in explicit and implicit visuomotor learning and working memory capacity. Scientific reports 6:36633.

Cisek P (2007) Cortical mechanisms of action selection: the affordance competition hypothesis. Philosophical Transactions of the Royal Society B: Biological Sciences 362:1585-1599.

Codol O, Holland PJ, Galea JM (2018) The relationship between reinforcement and explicit control during visuomotor adaptation. Scientific reports 8.

Cohen D, Prud'homme MJ, Kalaska JF (1994) Tactile activity in primate primary somatosensory cortex during active arm movements: correlation with receptive field properties. Journal of Neurophysiology 71:161-172.

Colder B (2015) The basal ganglia select the expected sensory input used for predictive coding. Frontiers in computational neuroscience 9:119.

Collins AG, Frank MJ (2012) How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. European Journal of Neuroscience 35:1024-1035.

Cordes D, Haughton VM, Arfanakis K, Wendt GJ, Turski PA, Moritz CH, Quigley MA, Meyerand ME (2000) Mapping functionally related regions of brain with functional connectivity MR imaging. American Journal of Neuroradiology 21:1636-1644.

Cordes D, Haughton VM, Arfanakis K, Carew JD, Turski PA, Moritz CH, Quigley MA, Meyerand ME (2001) Frequencies contributing to functional connectivity in the cerebral cortex in “resting-state” data. American Journal of Neuroradiology 22:1326-1333.

Cordo P, Gurfinkel VS, Bevan L, Kerr GK (1995) Proprioceptive consequences of tendon vibration during movement. Journal of Neurophysiology 74:1675-1688.

Cowan N (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. The Behavioral and brain sciences 24:87-114; discussion 114-185.

Criscimagna-Hemminger SE, Bastian AJ, Shadmehr R (2010) Size of error affects cerebellar contributions to motor learning. Journal of neurophysiology 103:2275-2284.

Cuppone AV, Squeri V, Semprini M, Masia L, Konczak J (2016) Robot-assisted proprioceptive training with added vibro-tactile feedback enhances somatosensory and motor performance. PloS one 11:e0164511.

d'Esposito M, Aguirre G, Zarahn E, Ballard D, Shin R, Lease J (1998) Functional MRI studies of spatial and nonspatial working memory. Cognitive Brain Research 7:1-13.

Darainy M, Vahdat S, Ostry DJ (2013) Perceptual learning in sensorimotor adaptation. Journal of neurophysiology 110:2152-2162.

Darian-Smith C, Darian-Smith I, Burman K, Ratcliffe N (1993) Ipsilateral cortical projections to areas 3a, 3b, and 4 in the macaque monkey. Journal of Comparative Neurology 335:200-213.

Darshan R, Leblois A, Hansel D (2014) Interference and shaping in sensorimotor adaptations with rewards. PLoS Comput Biol 10:e1003377.

Daw ND, Doya K (2006) The computational neurobiology of learning and reward. Current opinion in neurobiology 16:199-204.

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

Dayan E, Hamann JM, Averbeck BB, Cohen LG (2014) Brain structural substrates of reward dependence during behavioral performance. The Journal of Neuroscience 34:16433-16441.

Dayan P, Niv Y (2008) Reinforcement learning: the good, the bad and the ugly. Current opinion in neurobiology 18:185-196.

Della-Maggiore V, Villalta JI, Kovacevic N, McIntosh AR (2017) Functional evidence for memory stabilization in sensorimotor adaptation: a 24-h resting-state fMRI study. Cerebral Cortex 27:1748-1757.

Desjardins AE, Kiehl KA, Liddle PF (2001) Removal of Confounding Effects of Global Signal in Functional MRI Analyses. NeuroImage 13:751-758.

Diedrichsen J, Hashambhoy Y, Rane T, Shadmehr R (2005) Neural correlates of reach errors. The Journal of neuroscience : the official journal of the Society for Neuroscience 25:9919-9931.

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

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

Doyon J, Song AW, Karni A, Lalonde F, Adams MM, Ungerleider LG (2002) Experience-dependent changes in cerebellar contributions to motor sequence learning. Proceedings of the National Academy of Sciences 99:1017-1022.

Edin BB (2004) Quantitative analyses of dynamic strain sensitivity in human skin mechanoreceptors. Journal of neurophysiology 92:3233-3243.

Fiehler K, Burke M, Engel A, Bien S, Rösler F (2007) Kinesthetic working memory and action control within the dorsal stream. Cerebral Cortex 18:243-253.

Fine MS, Thoroughman KA (2007) Trial-by-trial transformation of error into sensorimotor adaptation changes with environmental dynamics. Journal of neurophysiology 98:1392-1404.

Friston KJ, Jezzard P, Turner R (1994) Analysis of functional MRI time-series. Human brain mapping 1:153-171.

Friston KJ, Williams S, Howard R, Frackowiak RS, Turner R (1996) Movement-related effects in fMRI time-series. Magnetic resonance in medicine 35:346-355.

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

Gallace A, Spence C (2009) The cognitive and neural correlates of tactile memory. Psychological bulletin 135:380.

Gallace A, Tan HZ, Haggard P, Spence C (2008) Short term memory for tactile stimuli. Brain research 1190:132-142.

Gaspar P, Stepniewska I, Kaas J (1992) Topography and collateralization of the dopaminergic projections to motor and lateral prefrontal cortex in owl monkeys. Journal of Comparative Neurology 325:1-21.

Gathercole SE, Service E, Hitch GJ, Adams AM, Martin AJ (1999) Phonological short-term memory and vocabulary development: further evidence on the nature of the relationship. Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition 13:65-77.

Gerbella M, Borra E, Tonelli S, Rozzi S, Luppino G (2012) Connectional heterogeneity of the ventral part of the macaque area 46. Cerebral Cortex 23:967-987.

Gershman SJ, Pesaran B, Daw ND (2009) Human reinforcement learning subdivides structured action spaces by learning effector-specific values. Journal of Neuroscience 29:13524-13531.

Ghez C, Gordon J, Ghilardi MF (1995) Impairments of reaching movements in patients without proprioception. II. Effects of visual information on accuracy. Journal of neurophysiology 73:361-372.

Gilbert CD, Wiesel TN (1992) Receptive field dynamics in adult primary visual cortex. Nature 356:150.

Gläscher J, Hampton AN, O'doherty JP (2008) Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. Cerebral cortex 19:483-495.

Glover GH, Li TQ, Ress D (2000) Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magnetic resonance in medicine 44:162-167.

Gobet F, Lane PC, Croker S, Cheng PC, Jones G, Oliver I, Pine JM (2001) Chunking mechanisms in human learning. Trends in cognitive sciences 5:236-243.

Goble DJ, Lewis CA, Brown SH (2006) Upper limb asymmetries in the utilization of proprioceptive feedback. Experimental brain research 168:307-311.

Goble DJ, Lewis CA, Hurvitz EA, Brown SH (2005) Development of upper limb proprioceptive accuracy in children and adolescents. Human movement science 24:155-170.

Goldman-Rakic PS (1996) Regional and cellular fractionation of working memory. Proceedings of the National Academy of Sciences 93:13473-13480.

Goldsworthy MR, Pitcher JB, Ridding MC (2012) The application of spaced theta burst protocols induces long-lasting neuroplastic changes in the human motor cortex. European Journal of Neuroscience 35:125-134.

Gong M, Li S (2014) Learned reward association improves visual working memory. Journal of Experimental Psychology: Human Perception and Performance 40:841.

Graybiel AM (2005) The basal ganglia: learning new tricks and loving it. Current Opinion in Neurobiology 15:638-644.

Greicius MD, Krasnow B, Reiss AL, Menon V (2003) Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. Proceedings of the National Academy of Sciences of the United States of America 100:253-258.

Greve DN, Fischl B (2009) Accurate and robust brain image alignment using boundary-based registration. Neuroimage 48:63-72.

Gupta P, Tisdale J (2009) Does phonological short-term memory causally determine vocabulary learning? Toward a computational resolution of the debate. Journal of memory and language 61:481-502.

Haber S, Kunishio K, Mizobuchi M, Lynd-Balta E (1995) The orbital and medial prefrontal circuit through the primate basal ganglia. The Journal of neuroscience 15:4851-4867.

Haber SN (2003) The primate basal ganglia: parallel and integrative networks. Journal of chemical neuroanatomy 26:317-330.

Haber SN, Fudge JL, McFarland NR (2000) Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. The Journal of neuroscience 20:2369-2382.

Harris JA, Miniussi C, Harris IM, Diamond ME (2002) Transient storage of a tactile memory trace in primary somatosensory cortex. Journal of Neuroscience 22:8720-8725.

Haruno M, Kawato M (2006) Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. Journal of neurophysiology 95:948-959.

Herzfeld DJ, Vaswani PA, Marko M, Shadmehr R (2014) A memory of errors in sensorimotor learning. Science:1253138.

Hikosaka O, Nakamura K, Sakai K, Nakahara H (2002) Central mechanisms of motor skill learning. Current opinion in neurobiology 12:217-222.

Holland P, Codol O, Galea JM (2018) Contribution of explicit processes to reinforcement-based motor learning. Journal of neurophysiology 119:2241-2255.

Horak F, Nashner L, Diener H (1990) Postural strategies associated with somatosensory and vestibular loss. Experimental brain research 82:167-177.

Huang VS, Haith A, Mazzoni P, Krakauer JW (2011) Rethinking motor learning and savings in adaptation paradigms: model-free memory for successful actions combines with internal models. Neuron 70:787-801.

Iriki A, Pavlides C, Keller A, Asanuma H (1989) Long-term potentiation in the motor cortex. Science 245:1385-1387.

Izawa J, Shadmehr R (2011) Learning from sensory and reward prediction errors during motor adaptation. PLoS computational biology 7:e1002012.

Jenkins WM, Merzenich MM, Ochs MT, Allard T, Guic-Robles E (1990) Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. Journal of neurophysiology 63:82-104.

Jenkinson M, Bannister P, Brady M, Smith S (2002) Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17:825-841.

Jo HJ, Saad ZS, Simmons WK, Milbury LA, Cox RW (2010) Mapping sources of correlation in resting state FMRI, with artifact detection and removal. Neuroimage 52:571-582.

Jones SA, Henriques DY (2010) Memory for proprioceptive and multisensory targets is partially coded relative to gaze. Neuropsychologia 48:3782-3792.

Jordan MI, Rumelhart DE (1992) Forward models: Supervised learning with a distal teacher. Cognitive science 16:307-354.

Kaas AL, Van Mier H, Goebel R (2006) The neural correlates of human working memory for haptically explored object orientations. Cerebral Cortex 17:1637-1649.

Kane MJ, Hambrick DZ, Tuholski SW, Wilhelm O, Payne TW, Engle RW (2004) The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. Journal of Experimental Psychology: General 133:189.

Kapogiannis D, Campion P, Grafman J, Wassermann EM (2008) Reward-related activity in the human motor cortex. European Journal of Neuroscience 27:1836-1842.

Kawato M (1999) Internal models for motor control and trajectory planning. Current opinion in neurobiology 9:718-727.

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

Keller A, Iriki A, Asanuma H (1990) Identification of neurons producing long-term potentiation in the cat motor cortex: Intracellular recordings and labeling. Journal of comparative neurology 300:47-60.

Kelly RE, Alexopoulos GS, Wang Z, Gunning FM, Murphy CF, Morimoto SS, Kanellopoulos D, Jia Z, Lim KO, Hoptman MJ (2010) Visual inspection of independent components: Defining a procedure for artifact removal from fMRI data. Journal of Neuroscience Methods 189:233-245.

Kennerley SW, Wallis JD (2009) Reward-dependent modulation of working memory in lateral prefrontal cortex. Journal of Neuroscience 29:3259-3270.

Kitada R, Hashimoto T, Kochiyama T, Kito T, Okada T, Matsumura M, Lederman SJ, Sadato N (2005) Tactile estimation of the roughness of gratings yields a graded response in the human brain: an fMRI study. NeuroImage 25:90-100.

Kleim JA, Barbay S, Nudo RJ (1998) Functional reorganization of the rat motor cortex following motor skill learning. Journal of neurophysiology 80:3321-3325.

Klink PC, Jeurissen D, Theeuwes J, Denys D, Roelfsema PR (2017) Working memory accuracy for multiple targets is driven by reward expectation and stimulus contrast with different time-courses. Scientific reports 7:9082.

Knutson B, Fong GW, Adams CM, Varner JL, Hommer D (2001) Dissociation of reward anticipation and outcome with event-related fMRI. Neuroreport 12:3683-3687.

Koch MA, Norris DG, Hund-Georgiadis M (2002) An investigation of functional and anatomical connectivity using magnetic resonance imaging. Neuroimage 16:241-250.

Kostopoulos P, Albanese M-C, Petrides M (2007) Ventrolateral prefrontal cortex and tactile memory disambiguation in the human brain. Proceedings of the National Academy of Sciences 104:10223-10228.

Krakauer JW, Ghilardi M-F, Ghez C (1999) Independent learning of internal models for kinematic and dynamic control of reaching. Nature neuroscience 2:1026.

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

Krawczyk DC, Gazzaley A, D'Esposito M (2007) Reward modulation of prefrontal and visual association cortex during an incentive working memory task. Brain research 1141:168-177.

Lederman SJ, Klatzky RL (2009) Haptic perception: A tutorial. Attention, Perception, & Psychophysics 71:1439-1459.

Lehéricy S, Benali H, Van de Moortele P-F, Pélégrini-Issac M, Waechter T, Ugurbil K, Doyon J (2005) Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. Proceedings of the National Academy of Sciences 102:12566-12571.

Leon MI, Shadlen MN (1999) Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. Neuron 24:415-425.

Logothetis NK (2002) The neural basis of the blood–oxygen–level–dependent functional magnetic resonance imaging signal. Philosophical Transactions of the Royal Society of London B: Biological Sciences 357:1003-1037.

Lu MT, Preston JB, Strick PL (1994) Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. Journal of Comparative Neurology 341:375-392.

Luppino G, Matelli M, Camarda R, Rizzolatti G (1993) Corticocortical connections of area F3 (SMA-proper) and area F6 (pre-SMA) in the macaque monkey. Journal of Comparative Neurology 338:114-140.

Madelain L, Paeye C, Wallman J (2011) Modification of saccadic gain by reinforcement. Journal of neurophysiology 106:219-232.

Malfait N, Ostry DJ (2004) Is interlimb transfer of force-field adaptation a cognitive response to the sudden introduction of load? Journal of Neuroscience 24:8084-8089.

Manley H, Dayan P, Diedrichsen J (2014) When money is not enough: awareness, success, and variability in motor learning. PLoS One 9:e86580.

Mathis MW, Mathis A, Uchida N (2017) Somatosensory cortex plays an essential role in forelimb motor adaptation in mice. Neuron 93:1493-1503. e1496.

McDougle SD, Boggess MJ, Crossley MJ, Parvin D, Ivry RB, Taylor JA (2016) Credit assignment in movement-dependent reinforcement learning. Proceedings of the National Academy of Sciences 113:6797-6802.

McKeown MJ, Hansen LK, Sejnowsk TJ (2003) Independent component analysis of functional MRI: what is signal and what is noise? Current opinion in neurobiology 13:620-629.

Mehler DMA, Reichenbach A, Klein J, Diedrichsen J (2017) Minimizing endpoint variability through reinforcement learning during reaching movements involving shoulder, elbow and wrist. PloS one 12:e0180803.

Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. Annual review of neuroscience 24:167-202.

Miller GA (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological review 63:81.

Miller GA, Galanter E, Pribram KH (1960) Plans and the structure of behavior.

Mink JW (1996) The basal ganglia: focused selection and inhibition of competing motor programs. Progress in neurobiology 50:381-425.

Moeller S, Yacoub E, Olman CA, Auerbach E, Strupp J, Harel N, Uğurbil K (2010) Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. Magnetic Resonance in Medicine 63:1144-1153.

Mogenson GJ, Jones DL, Yim CY (1980) From motivation to action: functional interface between the limbic system and the motor system. Progress in neurobiology 14:69-97.

Murphy K, Birn RM, Bandettini PA (2013) Resting-state fMRI confounds and cleanup. Neuroimage 80:349-359.

Newell KM (1991) Motor skill acquisition. Annual review of psychology 42:213-237.

Nissen MJ, Bullemer P (1987) Attentional requirements of learning: Evidence from performance measures. Cognitive Psychology 19:1-32.

Nudo R, Milliken G, Jenkins W, Merzenich M (1996) Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys. The Journal of Neuroscience.

O'Doherty J, Critchley H, Deichmann R, Dolan RJ (2003) Dissociating Valence of Outcome from Behavioral Control in Human Orbital and Ventral Prefrontal Cortices. The Journal of Neuroscience 23:7931-7939.

O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 304:452-454.

Ölveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. PLoS biology 3:e153.

Owen AM, McMillan KM, Laird AR, Bullmore E (2005) N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. Human brain mapping 25:46-59.

Pavlides C, Miyashita E, Asanuma H (1993) Projection from the sensory to the motor cortex is important in learning motor skills in the monkey. J Neurophysiol 70:733-741.

Pekny SE, Izawa J, Shadmehr R (2015) Reward-dependent modulation of movement variability. Journal of Neuroscience 35:4015-4024.

Penhune V, Doyon J (2005) Cerebellum and M1 interaction during early learning of timed motor sequences. Neuroimage 26:801-812.

Petrides M, Pandya DN (1984) Projections to the frontal cortex from the posterior parietal region in the rhesus monkey. Journal of Comparative Neurology 228:105-116.

Petrides M, Pandya D (2002) Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. European Journal of Neuroscience 16:291-310.

Pons T, Kaas J (1986) Corticocortical connections of area 2 of somatosensory cortex in macaque monkeys: a correlative anatomical and electrophysiological study. Journal of Comparative Neurology 248:313-335.

Power JD, Schlaggar BL, Petersen SE (2014) Studying brain organization via spontaneous fMRI signal. Neuron 84:681-696.

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59:2142-2154.

Preuschhof C, Heekeren HR, Taskin B, Schubert T, Villringer A (2006) Neural correlates of vibrotactile working memory in the human brain. Journal of Neuroscience 26:13231-13239.

Preuss TM, Goldman-Rakic PS (1989) Connections of the ventral granular frontal cortex of macaques with perisylvian premotor and somatosensory areas: anatomical evidence for somatic representation in primate frontal association cortex. Journal of Comparative Neurology 282:293-316.

Prud'Homme M, Kalaska JF (1994) Proprioceptive activity in primate primary somatosensory cortex during active arm reaching movements. Journal of neurophysiology 72:2280-2301.

Pruim RH, Mennes M, van Rooij D, Llera A, Buitelaar JK, Beckmann CF (2015) ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. Neuroimage 112:267-277.

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of brain function. Proceedings of the National Academy of Sciences 98:676-682.

Raj D, Anderson AW, Gore JC (2001) Respiratory effects in human functional magnetic resonance imaging due to bulk susceptibility changes. Physics in Medicine & Biology 46:3331.

Ramnani N, Miall R (2003) Instructed delay activity in the human prefrontal cortex is modulated by monetary reward expectation. Cerebral Cortex 13:318-327.

Reed CL, Klatzky RL, Halgren E (2005) What vs. where in touch: an fMRI study. Neuroimage 25:718-726.

Robertson EM (2007) The Serial Reaction Time Task: Implicit Motor Skill Learning? The Journal of Neuroscience 27:10073-10075.

Roesch MR, Olson CR (2003) Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. Journal of Neurophysiology 90:1766-1789.

Romo R, Hernández A, Zainos A (2004) Neuronal correlates of a perceptual decision in ventral premotor cortex. Neuron 41:165-173.

Romo R, Brody CD, Hernández A, Lemus L (1999) Neuronal correlates of parametric working memory in the prefrontal cortex. Nature 399:470.

Romo R, Hernández A, Zainos A, Lemus L, Brody CD (2002) Neuronal correlates of decision-making in secondary somatosensory cortex. Nature neuroscience 5:1217.

Rushworth MF, Walton ME, Kennerley SW, Bannerman DM (2004) Action sets and decisions in the medial frontal cortex. Trends Cogn Sci 8:410-417.

Saad ZS, Gotts SJ, Murphy K, Chen G, Jo HJ, Martin A, Cox RW (2012) Trouble at rest: how correlation patterns and group differences become distorted after global signal regression. Brain connectivity 2:25-32.

Sainburg RL, Ghilardi MF, Poizner H, Ghez C (1995) Control of limb dynamics in normal subjects and patients without proprioception. Journal of neurophysiology 73:820-835.

Sakamoto T, Arissian K, Asanuma H (1989) Functional role of the sensory cortex in learning motor skills in cats. Brain Res 503:258-264.

Salimi-Khorshidi G, Douaud G, Beckmann CF, Glasser MF, Griffanti L, Smith SM (2014) Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. Neuroimage 90:449-468.

Salway AF, Logie RH (1995) Visuospatial working memory, movement control and executive demands. British Journal of Psychology 86:253-269.

Samejima K, Ueda Y, Doya K, Kimura M (2005) Representation of action-specific reward values in the striatum. Science 310:1337-1340.

Sami S, Robertson EM, Miall RC (2014) The time course of task-specific memory consolidation effects in resting state networks. Journal of Neuroscience 34:3982-3992.

Sarlegna FR, Gauthier GM, Bourdin C, Vercher J-L, Blouin J (2006) Internally driven control of reaching movements: a study on a proprioceptively deafferented subject. Brain research bulletin 69:404-415.

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

Schultz W, Apicella P, Ljungberg T (1993) Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. Journal of neuroscience 13:900-913.

Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275:1593-1599.

Schultz W, Tremblay L, Hollerman JR (2000) Reward processing in primate orbitofrontal cortex and basal ganglia. Cerebral Cortex 10:272-283.

Seidler RD, Bo J, Anguera JA (2012) Neurocognitive contributions to motor skill learning: the role of working memory. Journal of motor behavior 44:445-453.

Selemon L, Goldman-Rakic P (1985) Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. The Journal of Neuroscience 5:776-794.

Setsompop K, Gagoski BA, Polimeni JR, Witzel T, Wedeen VJ, Wald LL (2012) Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. Magnetic Resonance in Medicine 67:1210-1224.

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

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

Shadmehr R, Smith MA, Krakauer JW (2010) Error correction, sensory prediction, and adaptation in motor control. Annual review of neuroscience 33:89-108.

Shima K, Tanji J (1998) Role for Cingulate Motor Area Cells in Voluntary Movement Selection Based on Reward. Science 282:1335-1338.

Shmuelof L, Huang VS, Haith AM, Delnicki RJ, Mazzoni P, Krakauer JW (2012) Overcoming motor “forgetting” through reinforcement of learned actions. Journal of Neuroscience 32:14617-14621a.

Sidarta A, Vahdat S, Bernardi NF, Ostry DJ (2016) Somatic and reinforcement-based plasticity in the initial stages of human motor learning. Journal of Neuroscience 36:11682-11692.

Skinner BF (1963) Operant behavior. American Psychologist 18:503.

Skinner BF (1965) Science and human behavior: Simon and Schuster.

Smith DV, Hayden BY, Truong T-K, Song AW, Platt ML, Huettel SA (2010) Distinct value signals in anterior and posterior ventromedial prefrontal cortex. The Journal of neuroscience 30:2490-2495.

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

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR, Beckmann CF (2009) Correspondence of the brain's functional architecture during activation and rest. Proceedings of the National Academy of Sciences 106:13040-13045.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23 Suppl 1:S208-219.

Smyth MM, Pendleton LR (1990) Space and movement in working memory. The Quarterly Journal of Experimental Psychology Section A 42:291-304.

Sperling G (1960) The information available in brief visual presentations. Psychological monographs: General and applied 74:1.

Srinivasan MA, LaMotte RH (1995) Tactual discrimination of softness. Journal of Neurophysiology 73:88-101.

Steele CJ, Penhune VB (2010) Specific increases within global decreases: a functional magnetic resonance imaging investigation of five days of motor sequence learning. Journal of Neuroscience 30:8332-8341.

Stoeckel MC, Weder B, Binkofski F, Buccino G, Shah NJ, Seitz RJ (2003) A fronto-parietal circuit for tactile object discrimination:: an event-related fmri study. Neuroimage 19:1103-1114.

Sul JH, Jo S, Lee D, Jung MW (2011) Role of rodent secondary motor cortex in value-based action selection. Nature neuroscience 14:1202-1208.

Sun FT, Miller LM, Rao AA, D'Esposito M (2007) Functional connectivity of cortical networks involved in bimanual motor sequence learning. Cerebral cortex 17:1227-1234.

Sutton RS, Barto AG (1998) Reinforcement learning: An introduction: MIT press Cambridge.

Takikawa Y, Kawagoe R, Itoh H, Nakahara H, Hikosaka O (2002) Modulation of saccadic eye movements by predicted reward outcome. Experimental brain research 142:284-291.

Tanné-Gariépy J, Rouiller EM, Boussaoud D (2002) Parietal inputs to dorsal versus ventral premotor areas in the macaque monkey: evidence for largely segregated visuomotor pathways. Experimental Brain Research 145:91-103.

Therrien AS, Wolpert DM, Bastian AJ (2015) Effective reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. Brain 139:101-114.

Tong C, Wolpert DM, Flanagan JR (2002) Kinematics and dynamics are not represented independently in motor working memory: evidence from an interference study. Journal of Neuroscience 22:1108-1113.

Tremblay L, Schultz W (2000) Reward-related neuronal activity during go-nogo task performance in primate orbitofrontal cortex. Journal of neurophysiology 83:1864-1876.

Tsujimoto S, Sawaguchi T (2004) Neuronal representation of response–outcome in the primate prefrontal cortex. Cerebral Cortex 14:47-55.

Tumer EC, Brainard MS (2007) Performance variability enables adaptive plasticity of ‘crystallized’adult birdsong. Nature 450:1240.

Tziortzi AC, Haber SN, Searle GE, Tsoumpas C, Long CJ, Shotbolt P, Douaud G, Jbabdi S, Behrens TEJ, Rabiner EA, Jenkinson M, Gunn RN (2014) Connectivity-Based Functional Analysis of Dopamine Release in the Striatum Using Diffusion-Weighted MRI and Positron Emission Tomography. Cerebral Cortex 24:1165-1177.

Vahdat S, Darainy M, Ostry DJ (2014) Structure of plasticity in human sensory and motor networks due to perceptual learning. Journal of Neuroscience 34:2451-2463.

Vahdat S, Darainy M, Milner TE, Ostry DJ (2011) Functionally specific changes in resting-state sensorimotor networks following motor learning. The Journal of Neuroscience 31:16907-16915.

Van Dijk KR, Sabuncu MR, Buckner RL (2012) The influence of head motion on intrinsic functional connectivity MRI. Neuroimage 59:431-438.

Vidoni E, Acerra N, Dao E, Meehan S, Boyd L (2010) Role of the primary somatosensory cortex in motor learning: An rTMS study. Neurobiology of learning and memory 93:532-539.

Wechsler D (1999) Wechsler adult intelligence scale–Third Edition (WAIS–III). San Antonio, Texas: Psychological Corporation.

Williams SM, Goldman-Rakic P (1998) Widespread origin of the primate mesofrontal dopamine system. Cerebral cortex (New York, NY: 1991) 8:321-345.

Wise RA, Rompre P-P (1989) Brain dopamine and reward. Annual review of psychology 40:191-225.

Wolpert DM, Ghahramani Z, Jordan MI (1995) An internal model for sensorimotor integration. Science 269:1880-1882.

Wong JD, Kistemaker DA, Chin A, Gribble PL (2012) Can proprioceptive training improve motor learning? Journal of neurophysiology 108:3313-3321.

Woolley SC, Rajan R, Joshua M, Doupe AJ (2014) Emergence of context-dependent variability across a basal ganglia network. Neuron 82:208-223.

Wu HG, Miyamoto YR, Castro LNG, Ölveczky BP, Smith MA (2014) Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. Nature neuroscience 17:312.

Xiong J, Parsons LM, Gao JH, Fox PT (1999) Interregional connectivity to primary motor cortex revealed using MRI resting state images. Human brain mapping 8:151-156.

Zatorre RJ, Salimpoor VN (2013) From perception to pleasure: music and its neural substrates. Proceedings of the National Academy of Sciences 110:10430-10437.