

Mechanisms underlying reinforcement learning of motor skills

Présentée le 7 février 2023

À l'école polytechnique fédérale de Lausanne
à la Faculté des sciences de la vie
Unité du Prof. Hummel
Programme doctoral en neurosciences

et À l'Université catholique de Louvain
Institute of Neuroscience (IoNS)

Thèse présentée en vue de l'obtention du grade de doctorat en Sciences médicales

par

Pierre Theopistos VASSILIADIS

Acceptée sur proposition du jury

Prof. D. N. A. Van De Ville, président du jury
Prof. F. C. Hummel, Prof. J. Duqué, directeurs de thèse
Dr. G. Derosiere, co-directeur
Prof. R. Hardwick, rapporteur
Prof. H. Schambra, rapporteuse
Prof. O. Blanke, rapporteur

Members of the UCLouvain accompanying committee:

Prof. Yves Vandermeeren (President of the committee)

Prof. André Mouraux

Prof. Michael Andres

Acknowledgements

This joint PhD performed between Belgium and Switzerland has come with a series of challenges that have transformed me scientifically and humanly. This work could not have been achieved without the help of many people that I would like to thank.

First, I would like to express my gratitude to my three supervisors for their constant support and trust during this amazing journey.

Julie, I would like to thank you for providing me with so many opportunities since the beginning of my work in your lab, back in 2014. Working by your side has been, and is, a constant source of scientific and human inspiration which has shaped the scientist I am today and the PI I hope to become.

Friedhelm, I would like to thank you for your constant trust in my ideas and abilities since my arrival in the lab. You have taught me how important it is to be versatile, ambitious in the questions we aim for, but also strategic in the way we think about science. I am looking forward to continuing this exciting journey with you.

Finally, to you Gerard, I would like to thank you for being there, "on the field" with me since the beginning and helping me so much throughout this PhD. Our endless discussions about neuroscience and life in every corner of the world are the beating heart of this work which has always motivated me to continue whatever the difficulties. You started this PhD as my colleague and co-supervisor and you finish it as one of my best friends. For all these great past and future moments, I would like to thank you my friend.

Then, I would like to thank the CoActions lab and in particular Cecile and Aegryan, great master students who largely contributed to this PhD. Thank you for your fantastic work and dedication. I also give special thanks to Vincent for the shared breaks and sports performed together in Belgium, which always cheered me up, I had a great time with you!

I am also grateful for the wonderful UPHummel team at EPFL, for welcoming me so warmly despite Covid and integrating me in the team. I have learned and continue to learn so much every day with you guys. In particular, I would like to thank you, Elena, for your constant

help in the last year and for always being available to answer my multiple (and annoying) questions on MRI, I am really looking forward for our future common projects!

Last but not least, this work would not have been possible without the unconditional support of my family and best friends during this period.

To “Les blaireaux” and “Team LFB”, I would like to thank you for always making me laugh and for sharing all these great memories. I am so happy to have you guys.

To you Olivia, I would like to thank you for all the great moments spent together in Brussels and Geneva and for supporting me in the recent challenging times I had to face. To cite Lorie, you are truly “Ma meilleure amie”.

To my best friend Mouss, I would just like to thank you for always being here and for sharing this great relationship that means so much. You are like a brother to me.

Finally, to my sisters and parents, I would like to tell you how grateful I am for your unconditional love during the good as well as the difficult times that I faced during these 4 years. This PhD is yours.

Abstract

Motor learning allows animals, including human beings, to acquire skills that are essential for efficient interactions with the environment. This ability to learn new motor skills is of great practical relevance for daily-life activities (such as when learning to drive), but also for motor rehabilitation after a lesion of the nervous system (such as a stroke). For a long time, motor learning has been mainly conceptualized as a process allowing to iteratively correct movements based on sensory information (e.g., visual, somatosensory). Importantly though, in the last years, there has been an increased appreciation that motor learning also results from other mechanisms including reinforcement learning, a process through which appropriate actions are selected through outcome-based feedback (e.g., success or failure). As such, recent evidence shows that reinforcement feedback and motivation can be beneficial for motor learning both in healthy individuals and neurological populations. Despite the potential importance of these findings to improve current rehabilitation protocols, the mechanisms underlying reinforcement-related improvements in motor learning remain largely unexplored. This PhD aimed at providing deeper mechanistic understanding of reinforcement learning of motor skills through behavioral analyses, neuroimaging and non-invasive brain stimulation. In **Study 1**, I found that enhancing motivation (by providing monetary reward for good performance) during a motor training can lead to persistent improvements in performance that are not obtained with reinforcement feedback only, and are related to an increased regulation of motor variability based on previous outcomes. In **Study 2**, I investigated the effect of reward timing (i.e., the delay between the end of movement execution and reward receipt) on motor learning and found that delaying reward by only a few seconds could strongly influence motor learning dynamics and consolidation. Finally, in **Study 3**, I investigated the causal role of the striatum in reinforcement motor learning. Here, I show, by combining an innovative non-invasive deep brain stimulation approach called transcranial electric temporal interference stimulation and neuroimaging, that a specific mechanism relying on striatal high gamma oscillations is causally involved in reinforcement learning of motor skills. Overall, this work characterizes key mechanisms underlying the effect of reinforcement on motor learning, paving the way towards the incorporation of optimized reinforcements in motor rehabilitation protocols.

Keywords

Motor learning, reinforcement learning, reward, motivation, timing, non-invasive brain stimulation, transcranial temporal interference stimulation.

Résumé

L'apprentissage moteur permet aux animaux, y compris aux êtres humains, d'acquérir des compétences essentielles pour interagir efficacement avec l'environnement. Cette capacité à apprendre de nouvelles habiletés motrices est d'une grande importance pratique pour les activités de la vie quotidienne (comme lors de l'apprentissage de la conduite), mais aussi pour la rééducation motrice après une lésion du système nerveux (comme suite à un accident vasculaire cérébral). Pendant longtemps, l'apprentissage moteur a été principalement conceptualisé comme un processus permettant de corriger itérativement des mouvements sur la base d'informations sensorielles (par exemple, visuelles ou somatosensorielles). Cependant, au cours des dernières années, il a été reconnu que l'apprentissage moteur est aussi le fruit d'autres mécanismes, dont l'apprentissage par renforcement, un processus par lequel les actions appropriées sont sélectionnées grâce à des informations sur le résultat des mouvements passés (par exemple succès ou échec). À ce titre, des données récentes montrent que le renforcement et la motivation peuvent être bénéfiques pour l'apprentissage moteur, tant chez les personnes en bonne santé que chez les populations neurologiques. Malgré l'importance potentielle de ces résultats pour améliorer les protocoles de révalidation actuels, les mécanismes qui sous-tendent les améliorations liées au renforcement dans l'apprentissage moteur restent largement inexplorés. Ce doctorat visait à fournir une compréhension mécanistique plus approfondie de l'apprentissage moteur par renforcement par le biais d'analyses comportementales, de neuro-imagerie et de stimulation cérébrale non-invasive. Dans **une première étude**, j'ai découvert que le fait d'augmenter la motivation (en offrant une récompense monétaire pour une bonne performance) pendant un entraînement moteur peut conduire à des améliorations persistantes de la performance qui ne sont pas obtenues avec le feedback de renforcement uniquement, et qui sont liées à une régulation accrue de la variabilité motrice basée sur les résultats des mouvements précédents. Dans **une seconde étude**, j'ai étudié l'effet du timing de la récompense (c'est-à-dire le délai entre la fin de l'exécution du mouvement et la récompense) sur l'apprentissage moteur et j'ai découvert que le fait de retarder le renforcement de quelques secondes seulement pouvait fortement influencer la dynamique et la consolidation de l'apprentissage mo-

teur. Enfin, dans **une troisième étude**, j'ai étudié le rôle causal du striatum dans l'apprentissage moteur par renforcement. Ici, je montre, en combinant une nouvelle approche de stimulation cérébrale non-invasive profonde appelée stimulation électrique transcrânienne par interférence temporelle et la neuro-imagerie, qu'un mécanisme spécifique s'appuyant sur les oscillations haut gamma dans le striatum est impliqué de manière causale dans l'apprentissage moteur par renforcement. Dans l'ensemble, ce travail caractérise des mécanismes clés qui sous-tendent l'effet du renforcement sur l'apprentissage moteur, ouvrant la voie à l'incorporation de renforcements optimisés dans les protocoles de rééducation motrice.

Mots-clés

Apprentissage moteur, apprentissage par renforcement, récompense, motivation, timing, stimulation cérébrale non-invasive, stimulation transcrânienne par interférence temporelle.

Contents

Acknowledgements	2
Abstract	4
Résumé	6
List of figures	11
Preamble	13
1. Introduction.....	15
1.1. The popular fascination for motor skills.	15
1.2. Principles of motor learning	16
1.2.1. Scope and definitions	16
1.2.2. Tasks commonly used to study motor learning	17
1.2.3. Different phases of motor learning	19
1.2.4. Neural correlates of motor learning.....	20
1.3. Sensory and reinforcement signals during motor learning.....	24
1.3.1. Sensory-based motor learning	25
1.3.2. Reinforcement-based motor learning.....	26
1.4. Plan of the PhD	35
Study 1: Reward boosts reinforcement-based motor learning	35
Study 2: Reward timing matters in motor learning	36
Study 3: Non-invasive stimulation of the human striatum disrupts reinforcement learning of motor skills.....	36
2. Study 1: Reward boosts reinforcement-based motor learning.....	39
2.1. Abstract	40
2.2. Introduction.....	41
2.3. Results	45
2.4. Discussion	59
2.5. Materials and Methods	64
2.5.1. Resources table	64
2.5.2. Participants	64
2.5.3. Methods details	64
2.5.4. Data and statistical analyses	67

2.6.	Supplementary materials.....	73
3.	Study 2: Reward timing matters in motor learning.....	75
3.1.	Abstract	76
3.2.	Introduction.....	77
3.3.	Results	79
3.4.	Discussion	89
3.5.	Materials and Methods	95
3.5.1.	Resources table	95
3.5.2.	Participants	95
3.5.3.	Methods details	96
3.5.4.	Data and statistical analyses	98
3.6.	Supplementary materials.....	103
4.	Study 3: Non-invasive stimulation of the human striatum disrupts reinforcement learning of motor skills	109
4.1.	Abstract	111
4.2.	Introduction.....	113
4.3.	Results	118
4.4.	Discussion	134
4.5.	Methods	141
4.5.1.	Participants	141
4.5.2.	Experimental procedures	141
4.5.3.	Data and statistical analyses	151
4.6.	Supplementary material	157
5.	Discussion.....	169
5.1.	Main conclusions	169
5.2.	General discussion.....	171
5.2.1.	Motivational interventions during motor learning	171
5.2.2.	Reward timing: what's next?.....	172
5.2.3.	Causal role of the striatum in reinforcement learning of motor skills	173
5.3.	Perspectives.....	174
5.3.1.	Towards incorporation of reinforcement into motor rehabilitation: optimization of training features and personalization.....	175
5.3.2.	Trancranial electric temporal interference stimulation: a new non-invasive approach for deep brain stimulation	185
5.4.	Conclusion	187
	References.....	188
	Annex 1: Motor training strengthens corticospinal suppression during movement preparation.....	210
	Annex 2: Selecting and executing actions for rewards	240

Annex 3: Beyond motor noise: considering other causes of impaired reinforcement learning in cerebellar patients.....	247
Curriculum Vitae.....	258

List of figures

Chapter 1:

Figure 1. Schematic plan of the PhD

Figure 2. The motor learning network.

Figure 3. Stimulation of deep brain structures with conventional non-invasive brain stimulation vs. tTIS.

Chapter 2:

Figure 1. Experimental protocol.

Figure 2. Learning curves.

Figure 3. Between-trial adjustments in the Error.

Figure 4. Between-trial adjustments in initiation time, amplitude error and variability.

Figure S1. Motor components.

Chapter 3:

Figure 1. Motor skill learning task.

Figure 2. Effect of reward timing on motor skill learning.

Figure 3. Effect of reward timing on the dynamics of learning.

Figure 4. Effect of reward timing on overnight consolidation of the motor skill.

Figure S1. Individual learning curves.

Figure S2. Effect of reward timing on motor skill learning with the inclusion of GroupPastStudy (n = 30).

Figure S3. Additional analysis showing the effect of reward timing on learning rates for early and late phases of practice.

Chapter 4:

Figure 1. Striatal tTIS during reinforcement learning of motor skills in the MRI.

Figure 2. Behavioural results.

Figure 3. Striatal activity.

Figure 4. Striatum to frontal cortex effective connectivity.

Figure 5. Relationship between impulsivity and the neural effects of tTIS_{80Hz}.

Figure S1. Patterns of motion of the target.

Figure S2. tTIS-related sensations.

Figure S3. Whole-brain activity during reinforcement motor learning.

Figure S4. Total signal to noise ratio (tSNR).

Chapter 5:

Figure 1. Main conclusions of the studies presented in this PhD.

Figure 2. Effect of visual uncertainty on reinforcement motor learning.

Figure 3. Effect of sensitivity to reward on reinforcement-related gains in motor learning in stroke patients.

Figure 4. Effect of reward on motor cortex plasticity during motor learning.

Figure 5. Factors to consider for the use of reinforcement feedback in motor rehabilitation.

Preamble

[I]t is the object of appetite which originates movement, this object may be either the real or the apparent good...To the thinking soul images serve as if they were contents of perception...just as if it were seeing, it calculates and deliberates what is to come by reference to what is present; and when it makes a pronouncement, as in the case of sensation it pronounces the object to be pleasant or painful, in this case it avoids or pursues.

Aristotle, On the Soul, Book III, Part 10; 350 BC.

My PhD journey led me to investigate the influence of performance-based feedback – also called reinforcement feedback – on human motor learning from behavioral and neurophysiological points of view. Inspired by recent research showing that motor learning is not purely a sensorimotor process but rather also relies on cognitive functions, I have conducted three studies investigating the mechanisms at play during reinforcement learning of motor skills.

In **Study 1**, I explored the effect of motivation on reinforcement motor learning. Employing a new motor skill learning task allowing to dissociate reinforcement feedback (providing knowledge of performance) and monetary rewards (providing motivation), I found that enhancing motivation during a motor training can lead to persistent improvements in performance that are not obtained with reinforcement feedback only, and are related to an increased regulation of motor variability based on previous outcomes.

In **Study 2**, I investigated the effect of reward timing (i.e., the delay between the end of movement execution and delivery of reward) on motor learning and found that delaying reward by only a few seconds could strongly influence motor learning dynamics and consolidation.

In **Study 3**, I studied the causal role of the striatum in reinforcement motor learning. Here, I show, by combining an innovative non-invasive deep brain stimulation approach called transcranial electric temporal interference stimulation and neuroimaging, that the striatum is causally involved in reinforcement motor learning.

Altogether, the present work characterizes key mechanisms at play during reinforcement learning of motor skills that could pave the way towards the incorporation of optimized reinforcements in motor rehabilitation therapies.

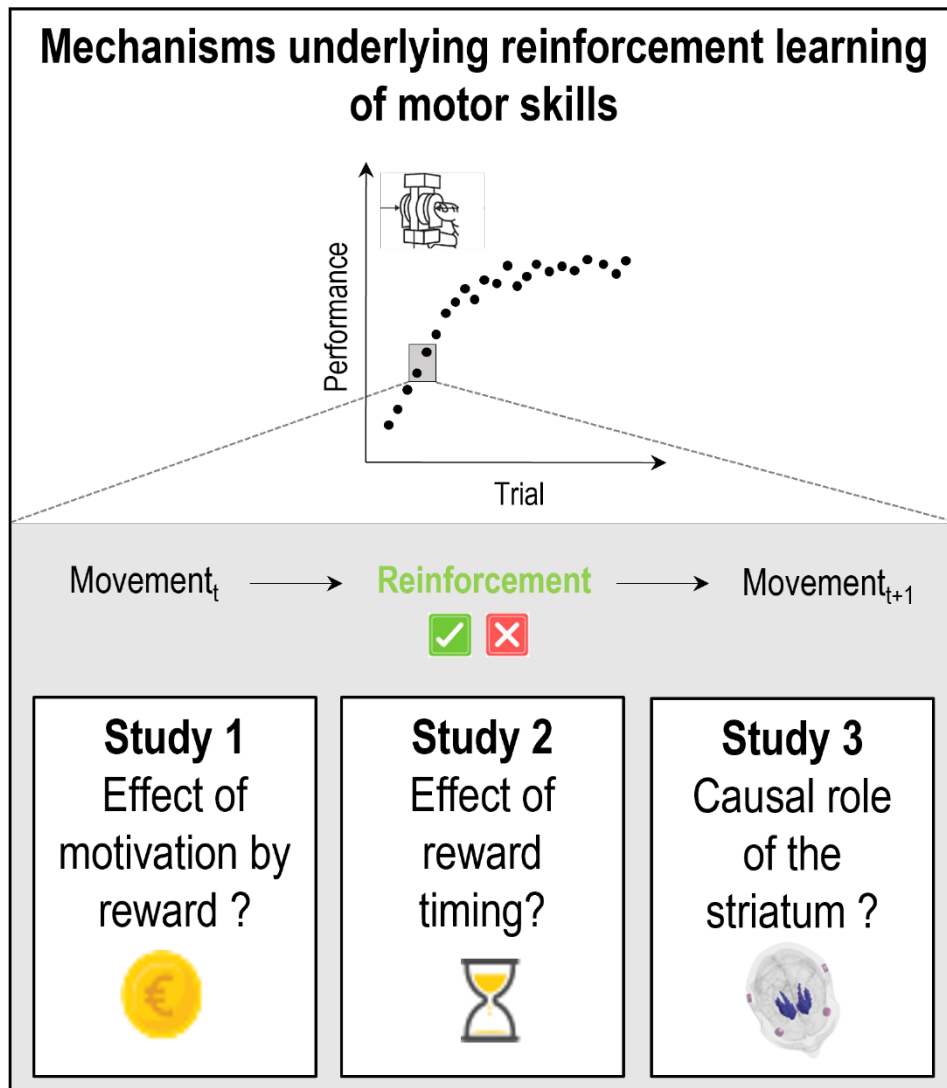


Figure 1. Schematic plan of the PhD

1. Introduction

This chapter describes the main concepts at the basis of my PhD work on reinforcement learning of motor skills. I first describe historical examples showing that displays of motor skills and the process of motor learning have always constituted an object of intense fascination for humans. After this initial part, I define the scope of motor learning and describe experimental tasks commonly used to study this phenomenon. Then, I summarize what we currently know about the two main phases of motor learning, namely acquisition and consolidation and present their putative neural correlates. Next, I focus on two key mechanisms of motor learning that are central to the present work; sensory-based error correction and reinforcement learning, and show how they both contribute to improvement of motor performance. Finally, I present in more details the specific rationale for the three studies presented in this thesis and how they allowed us to uncover key mechanisms at play during reinforcement learning of motor skills.

1.1. The popular fascination for motor skills.

Motor skills are an object of fascination for humans. Since antiquity, across different cultures and civilizations, watching people perform motor skills has constituted a substantial source of entertainment. There is plethora of examples of such popular interest in motor skills across history. Think about the excitement generated by Sumo wrestling in Japan since the 8th century, the effervescence preceding a performance of Wolfgang Amadeus Mozart in the 18th-century Vienna or, more recently, the enthusiasm created each year by the final of the Super Bowl in the United States. This fascination also extends to the performers of such skills, often considered as heroes in their respective societies, with roles and responsibilities spreading far beyond the entertainment domain. The universality of such interest in motor skills is also well represented by the ancient Olympic Games, created in 776 BC in Ancient Greece, and for which people of different (and often opponent) cities

gathered peacefully for several weeks every 4 years, sharing their passion for motor performance and admiration for athletes. More than 2500 years later, the Olympic Games continue to be a major event attracting billions of people from all around the globe¹.

But why are humans of different origins, cultures and epochs so fascinated by displays of motor skills? One possible reason for this may reside in the extensive amount of practice required to acquire motor skills and perform at such high level (Krakauer et al., 2019). Yet, despite such popular captivation for motor skills, the mechanisms underlying motor skill learning have remained unknown for a long time. In the last decades, scientists from different fields of research have become interested in this topic and have started to delineate the behavioral and neurophysiological processes occurring when people learn motor skills.

1.2. Principles of motor learning

1.2.1. Scope and definitions

Motor learning can be defined as any experience-dependent improvement in motor performance (Krakauer et al., 2019). As highlighted in the previous section, motor learning is a crucial and fascinating process to reach the expertise needed to perform sports or play music at a professional level. However, motor learning is also of great practical relevance in more common daily-life situations across the life span, from the baby learning to walk, the kid learning to tie shoelaces, the young adult learning to drive, to the elderly learning to use a tablet. Importantly, motor learning is also relevant for patients suffering from motor disability (e.g., such as after a stroke) and who have to re-learn skills through rehabilitation (Muratori et al., 2013). Finally, from an evolutionary perspective, the ability of individuals to learn and retain practiced motor skills in nature is likely to have favored survival, shaping evolution of the human brain (Boraud et al., 2018). Again, there are plenty of examples highlighting the importance of motor learning in nature, from the ability to hunt successfully, climb trees to escape predators or find fruits, to more advanced skills, such as using tools

¹ <https://olympics.com/ioc/news/olympic-games-tokyo-2020-watched-by-more-than-3-billion-people>

or building shelters. Hence, the scope of motor learning research appears to be extremely large, from understanding one of the most basic brain functions to promoting learning in health and disease.

The definition of motor learning provided above – an experience-dependent improvement in movements - is very broad and encompasses a variety of phenomena and situations (Wolpert et al., 2011; Krakauer et al., 2019). As such, daily-life motor tasks are extremely diverse (e.g., playing tennis versus using pedals when driving) and therefore probably engage a variety of learning mechanisms. A point of general agreement in the field is that motor learning is the process by which the speed and/or accuracy of movements is improved through practice, leading to an overall enhancement of the so-called speed-accuracy tradeoff (Chen et al., 2017). In other words, motor learning relies on improvements occurring at several levels of motor control that can be modulated according to task demands (Vassiliadis and Derosiere, 2020). For instance, a tennis player can improve her performance by increasing the speed of movement initiation (e.g., allowing faster displacement towards the ball and more vigorous strokes) and the accuracy of the movement in terms of selection (e.g., allowing to choose whether to hit a cross court or down the line backhand given the position of the opponent) or execution (e.g., allowing to reliably execute the selected backhand). Moreover, motor learning can also result in other modifications of behavior such as a reduced susceptibility to external perturbations (which can be measured as the relative modification of the movement when an external force is applied, Crevecoeur et al., 2019), or as an ability to perform movements with reduced cognitive load (which can be measured with dual-task paradigms; Wu et al., 2004). Overall, these examples indicate that motor learning involves a variety of phenomena that are likely to be engaged differently depending on the task that has to be performed.

1.2.2. Tasks commonly used to study motor learning

The variety of processes underlying motor learning is also well represented by the multiplicity of laboratory tasks that have been used to study it (Ranganathan et al., 2019). Notably, the field has been dominated by two main paradigms namely motor sequence

learning and motor adaptation (Krakauer et al., 2019). Briefly, motor sequence learning requires participants to type a sequence of key presses on a button box (or keyboard) as quickly and accurately as possible, whereas motor adaptation tasks require people to adapt a well-learned movement (often an upper-limb reaching movement) to a perturbation (often mechanical or visual). These paradigms therefore put emphasis on different aspects of motor learning: sequence tasks probe learning of successive motor elements while adaptation tasks focus on movement correction in response to experimental perturbations. Whereas these tasks have led to extremely important discoveries on the mechanisms and neural substrates of motor learning (Doyon et al., 2018; Morehead and Orban De Xivry, 2021), there is currently a debate in the literature on whether they constitute good models of skill learning (Krakauer et al., 2019). As such, a common feature of both tasks is that they require to improve movement selection (to press a key in a correct order or to counteract a perturbation), but put little emphasis on the correct execution of the movement, after selection. In contrast, real-world skills often also put strong emphasis on the execution process: in the tennis example considered above, selecting the correct stroke without being able to execute it correctly would have little benefit for the player (Du et al., 2022). Other motor learning paradigms can help partially address this issue by putting more emphasis on movement execution.

In this thesis, isometric force production tasks have been used as model of motor skill learning. In this type of tasks, participants are asked to modulate the level of force applied on a force sensor positioned in the hand in order to reach targets as quickly and accurately as possible on a computer screen (Reis et al., 2009; Spampinato and Celnik, 2018). This task therefore contains an execution component, because once a target has been selected, participants have to dose the applied force very carefully in order to reach it and stay within its boundaries. Notably, this type of tasks parallel skills that are relevant for daily-life activities, such as when pushing the brake pedal during driving, or when manipulating fragile objects. Moreover, such tasks were chosen because they previously showed sensitivity to reinforcements (Abe et al., 2011; Mawase et al., 2017), they allow to modulate sensory feedback (Mawase et al., 2017) and they can be performed by patients suffering from motor deficits (Raspovic et al., 2014; Hardwick et al., 2017), therefore enabling future clinical translational studies. Importantly, although isometric force production tasks were used in the present work, in this Introduction, I describe what we currently know about motor learning in

general by covering studies that exploited different types of tasks. Despite this variety of paradigms, a point of similarity is that learning usually develops in different phases, an aspect that I present in the next section.

1.2.3. Different phases of motor learning

Motor learning results from improvements occurring during repeated bouts of practice interspersed by resting periods. Performance improvement occurring during practice is often referred to as skill acquisition or “online” learning. Conversely, improvement occurring during rest periods, in the absence of practice is called skill consolidation or “offline” learning (Dayan and Cohen, 2011). Beyond acquisition and consolidation, it is also important to be able to apply the learned skills in other contexts that are different from the ones experienced during acquisition (Krakauer et al., 2006). For instance, we often learn to drive a car in very particular conditions (e.g., in our hometown, in a particular car). The ability to extend this knowledge to different traffic conditions and vehicles is crucial in order to avoid having to re-learn the skill each time we drive in a different context. This capacity is often referred to as generalization of the skill, and was not specifically investigated in the present work. Here, we will focus on the acquisition and consolidation phases of motor learning.

Motor skill acquisition

On the one hand, a great deal of research has shown that motor skill acquisition relies on various learning mechanisms including error-based movement corrections, reinforcement learning, use-dependent plasticity or the use of explicit strategies (Wolpert et al., 2011; Spampinato and Celnik, 2021; Therrien and Wong, 2022). These learning mechanisms and their putative neural correlates will be discussed in further details in section 1.3 of this Introduction. It is also important to note that these processes may have distinct temporal dynamics. As such, motor skill acquisition can be decomposed in a fast (i.e., developing over a few trials) and a slow (i.e., developing over tens/hundreds of trials) process with different characteristics (Smith et al., 2006). The slow process is characterized by both a low learning rate and a sluggish forgetting of the acquired behavior while the fast process entails both a high learning rate and a quick forgetting of the new behavior (Trewartha et al., 2014; McDougale

et al., 2015). As it will be discussed in the context of Study 2 (Vassiliadis et al., 2022), a current line of research in the field aims at characterizing which learning mechanisms support the fast and slow processes of motor learning (Trewartha et al., 2014; McDougale et al., 2015). Put together, these studies show that motor skill acquisition results from the operation of multiple interacting learning processes that are thought to develop with either fast or slow temporal dynamics (Spampinato and Celnik, 2021).

Motor skill consolidation

On the other hand, consolidation processes allow the strengthening of motor memories in the absence of training and are evidenced at the behavioral level by two phenomena. One is the improvement in performance between practice sessions (often referred to as “offline” gains in performance), the other is the stabilization of motor memories evidenced by a reduction in their susceptibility to interference (Robertson et al., 2004). Consolidation is crucial to ensure that what is learned (e.g., during a rehabilitation session) can be retained in future situations (e.g., when coming back at home). Classically, consolidation has been studied hours or days after the end of training (Doyon et al., 2009), with the implicit assumption that long periods of rest are required for the consolidation process to develop. Contrary to this view, recent work investigating motor sequence learning has demonstrated that consolidation can already occur over very short time spans, in the order of seconds (Bönstrup et al., 2019, 2020). Importantly, this rapid type of consolidation is thought to rely on partially different neural mechanisms from overnight consolidation, which requires sleep (Bönstrup et al., 2019; Jacobacci et al., 2020). Overall, these works start to unveil the multiplicity of mechanisms taking place during both practice and resting periods, which all contribute to the learning of new motor skills.

1.2.4. Neural correlates of motor learning

Multiple lines of research have investigated the neural correlates of motor skill acquisition and consolidation. It is important to note that the brain regions involved in the learning of a particular skill can greatly depend on multiple factors including the type of task (Steel et al., 2019), the learning mechanisms at play (Uehara et al., 2018) and the phase of learning

(Shadmehr and Holcomb, 1997; Doyon et al., 2018). In this section, I will present an overview of what we know about the neural substrates of motor skill acquisition and consolidation, and will finish by describing the plasticity mechanisms associated to motor learning.

Motor skill acquisition

Given the multiplicity of motor learning protocols, delineating a core motor learning network from individual studies can be challenging. To address this issue, a meta-analysis has described the brain regions activated during motor learning by compiling the results of 70 neuroimaging studies (Hardwick et al., 2013). This work has shown consistent activation in the primary motor cortex (M1), supplementary motor area (SMA), dorsal premotor cortex, primary somatosensory cortex, superior parietal lobule, thalamus, putamen (i.e., a part of the striatum), and cerebellum (**Figure 2**, Hardwick et al., 2013). Notably though, the study also found that the specific motor learning task performed influenced the volume of activation of these regions and the potential recruitment of other brain structures. These areas include the hippocampus (particularly in sequence learning tasks: Albouy et al., 2013, but see also McDougle et al., 2022), prefrontal regions (Floyer-Lea and Matthews, 2005; Sidarta et al., 2016; Draaisma et al., 2022) and dopaminergic nuclei of the midbrain such as the ventral tegmental area (VTA) and substantia nigra (Hosp et al., 2011; Hosp and Luft, 2013; Leemburg et al., 2018). Overall, these studies suggest that motor learning engages a core set of regions that are activated irrespective of the particular task performed and other task-specific areas depending on the skill that has to be learned.

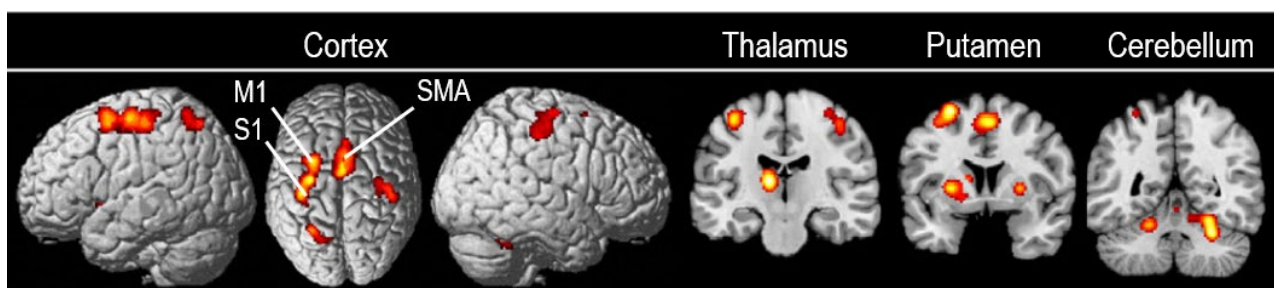


Figure 2. The motor learning network. Functional magnetic resonance imaging pre data from the meta-analysis of 70 experiments. Significant converging activity between studies was found in the left dPMC, pre-SMA, SMA proper, bilateral M1, left S1, left SPL, left thalamus, bilateral putamen, and bilateral/anterior medial cerebellum. Adapted from Hardwick et al., 2013, *NeuroImage*.

Finally, advanced neuroimaging and brain stimulation techniques have also allowed to study how functional interactions between these brain regions could be modulated during motor learning (Doyon et al., 2018). This body of work has revealed that motor skill acquisition is associated to connectivity changes in cortico-cerebellar (Schlerf et al., 2012; Mehrkanoon et al., 2016; Spampinato et al., 2017; Spampinato and Celnik, 2018; Uehara et al., 2018; Schubert et al., 2021), cortico-striatal (Hamzei et al., 2012; Gabbitov et al., 2015; Kupferschmidt et al., 2017; Wolff et al., 2022), and cortico-cortical circuits (Waters et al., 2017; Gabbitov et al., 2019; Mirdamadi and Block, 2020). Notably, interactions between sub-cortical regions are also likely to contribute to motor learning (e.g., striatum-hippocampus interactions during sequence learning; Albouy et al., 2013; Freedberg et al., 2020). Hence, motor training induces changes in functional interactions between brain regions that are again likely to depend on the particular demands of the motor task that is being performed (Doyon and Benali, 2005).

Motor skill consolidation

Motor consolidation is the process by which motor memories are stabilized or enhanced during resting periods (Johnson et al., 2021). As mentioned above, consolidation has been extensively studied in the context of sleep but there is accumulating evidence that offline learning can already occur over much shorter timescales (in the order of seconds; Bönstrup et al., 2019, 2020). Seminal as well as more recent studies have repetitively implicated M1 as a crucial region for consolidation (Muellbacher et al., 2002; Robertson et al., 2005; Reis et al., 2009; Hussain et al., 2021). Notably, there is also evidence for an involvement of the striatum and hippocampus (Fischer et al., 2005; Albouy et al., 2008, 2013b, 2015; Debas et al., 2010), and the fronto-parietal network (Albert et al., 2009; Taubert et al., 2011; Bönstrup et al., 2019) in different aspects of motor consolidation. A plausible mechanism for the consolidation of motor memories is the replay of practice-related neural activity during rest that would allow the strengthening of previously acquired memories. As such, neural replay is a central mechanism for multiple memory systems (Foster and Wilson, 2006; Singer and Frank, 2009; Carr et al., 2011; Ambrose et al., 2016; Liu et al., 2019, 2021; Sterpenich et al., 2021) and has been recently shown to also be at play during consolidation of motor memories in a network encompassing the hippocampus and sensorimotor cortex (Buch et al., 2021; Rubin et al., 2022). Hence, consolidation of motor memories appears to

be a crucial mechanism for motor learning, supporting offline improvement and stabilization of the skill during rest and is likely to rely, at least in part, on neural replay in a distributed brain network.

Brain plasticity induced by motor skill learning

A key property of the brain is its ability to undergo plastic changes. Plasticity is defined as a modification of the brain's structure and/or function in response to internal or external events (Pascual-Leone et al., 2005; Sampaio-Baptista et al., 2018). It has been shown to be a crucial mechanism for motor skill learning (Doyon and Benali, 2005; Dayan and Cohen, 2011) and motor recovery after a lesion of the nervous system, such as a stroke (Raffin and Hummel, 2018; Sampaio-Baptista et al., 2018). Notably, it seems difficult to identify which plastic mechanisms are specifically associated to skill acquisition given that consolidation mechanisms can already take place over very short periods of rest (Bönstrup et al., 2019) and induce micro-structural plasticity (Jacobacci et al., 2020). Hence, in this section, I describe the plasticity processes associated to motor learning in general, without dissociating the mechanisms putatively associated to motor skill acquisition and consolidation.

At the cellular level, motor learning is thought to rely, at least in part, on synaptic plasticity. In this regard, long-term potentiation (LTP) and long-term depression (LTD), cellular mechanisms initially described in the hippocampus (Dudek and Bear, 1992; Bliss, T.V.P. & Collingridge, 1993), modulate synaptic efficacy and contribute to various types of learning (Whitlock et al., 2006; Nabavi et al., 2014). More specifically, LTP and LTD are evoked by simultaneous firing of neurons sharing synapses and lead to a durable increase or decrease of post-synaptic potentials, respectively (Malenka and Bear, 2004). LTP and LTD in areas including M1 (Riolt-Pedotti et al., 1998, 2000), the striatum (Mahon et al., 2004; Dang et al., 2006; Giordano et al., 2018) and the cerebellum (Schonewille et al., 2010, 2021) constitute crucial mechanisms for motor learning which are thought to contribute to learning-related reorganization (Riolt-Pedotti et al., 1998; Butefisch et al., 2000). Hence, one important neuronal mechanism for motor learning-related plasticity is the regulation of synaptic efficacy. It is important to note that other cellular mechanisms such as remodeling of dendritic spines (which can occur during non-rapid eye movement sleep, Yang et al.,

2014) and axonal terminals (Boele et al., 2013), glial hypertrophy (Anderson et al., 1994) and synaptogenesis (Kleim et al., 2004), are also thought to contribute to motor learning.

An important line of research has focused on the cortical reorganization resulting from motor learning. Such cortical plasticity can occur already after very short durations of practice (Pascual-Leone et al., 1995) but such rapid changes often do not persist over time (Classen et al., 1998). In contrast, years of training like that undergone by professional musicians or athletes can lead to persistent changes in cortical representations (Elbert et al., 1995; Pantev et al., 2001; Jäncke et al., 2009). At the level of the motor cortex, training causes changes in the size and excitability of motor cortical maps (Kami et al., 1995; Pascual-Leone et al., 1995; Rosenkranz et al., 2007; Vassiliadis et al., 2020a), in the activity of intracortical circuits (Hummel et al., 2005; Dupont-Hadwen et al., 2018) and in the direction of involuntary movements evoked by transcranial magnetic stimulation (TMS) of M1 (a process often referred to as use-dependent plasticity (UDP); Classen et al., 1998; Duque et al., 2008; Mawase et al., 2017). In a study performed in the beginning of my thesis (Annex 1, Vassiliadis et al., 2020a), we have shown that motor training also involves a modulation of preparatory activity in the motor system (i.e., activity recorded during the preparation of simple finger movements, before initiation) that was related to improvements in the speed of action initiation. Hence, plastic reorganization of the motor cortex appears as a key neurophysiological mechanism supporting improvements of movements with practice. As mentioned above, motor learning also involves modulations of connectivity in cortico-cerebellar, cortico-striatal and cortico-cortical circuits over the course of training. Interestingly, some of these changes persist at rest, after training, suggesting that motor learning leads to a plastic reorganization of these circuits (Doyon and Benali, 2005). Put together, this body of work indicates that multiple structural and functional neuronal changes support the acquisition and consolidation of new motor skills. Interestingly, such neuronal changes are likely to be modulated by the particular learning signals available during training.

1.3. Sensory and reinforcement signals during motor learning

Recent evidence indicates that motor learning is not a unitary process, but rather results from the operation of parallel and interacting learning processes including sensory-

based learning (often also referred to as “error-based learning”) and reinforcement learning, (Vassiliadis et al., 2019; Therrien and Wong, 2022). During my thesis, I mainly focused on these two learning mechanisms, which I therefore present in detail below. It is important to note, however, that other learning processes including use-dependent learning (resulting from the repetition of similar movements; Diedrichsen et al., 2010) and the use of explicit strategies (allowing strategical corrections of movements based on task instructions; Taylor et al., 2014) also contribute to motor learning (Spampinato and Celnik, 2021), and are likely to have operated in the experimental tasks presented in this thesis.

1.3.1. Sensory-based motor learning

Previous work on motor learning has largely focused on how motor commands can be adjusted based on sensory feedback (Shadmehr et al., 2010b). This line of work has mainly employed adaptation paradigms where sensory feedback (e.g., visual, somatosensory) can be manipulated experimentally. As such, in motor adaptation tasks, a perturbation (e.g., visual, mechanical) is imposed while participants are performing movements (e.g., reaching, pointing, eye movements) towards visual targets (Krakauer et al., 2019). The perturbation leads participants to adapt by generating a motor command in the opposite direction, a phenomenon that results from both implicit and explicit processes (Taylor et al., 2014). The implicit process is thought to rely on the adjustments of movements based on sensory prediction errors (SPE), which correspond to the difference between predicted and actual sensory outcome (Tseng et al., 2007; Schlerf and Ivry, 2012; Shadmehr, 2017, 2018). The predicted sensory consequences of movements may be computed through a forward model that exploits an efference copy of the motor command to predict its future sensory consequences (Wolpert et al., 1995; Flanagan and Wing, 1997; Morton and Bastian, 2006; Shadmehr et al., 2010b; but see also Hadjiosif et al., 2021 for a recent challenge of this theory). Notably, this predictive mechanism is crucial to counteract sensorimotor delays inherent to the architecture of our nervous system (Wolpert et al., 2011). As mentioned above, the role of sensory feedback has been mostly examined through adaptation paradigms, yet, it is also a fundamental process in other motor learning tasks including motor sequence learning (Popp et al., 2022) and force modulation tasks (Miall et al., 1985, 2021; Flanagan

and Wing, 1997; Raspopovic et al., 2014). A common aspect in these different types of tasks is that sensory feedback provides directional information on how movements should be adjusted after an error (Spampinato and Celnik, 2021).

At the neural level, sensory-based error correction has been consistently shown to rely on the cerebellum (Tseng et al., 2007; Schlerf and Ivry, 2012; Schlerf et al., 2012; Taylor and Ivry, 2014; Herzfeld et al., 2018; Uehara et al., 2018). At the cellular level, sensory errors cause climbing fibers (a class of afferent fibers arising from the inferior olive, a complex of nuclei located in the brainstem) to generate “complex spikes” (a large depolarizing event) in Purkinje cells of the cerebellum (Ito, 2001). These cells, which are the only efferent neurons of the cerebellar cortex, suppress deep cerebellar nuclei, reducing their excitatory influence on M1 through the thalamus. Hence, sensory errors may generate error corrections through a cerebello-thalamo-cortical pathway (Allen and Tsukahara, 1974). Moreover, there is also evidence that Purkinje cells can undergo synaptic plasticity (including through LTD) that may subserve motor learning (Ito, 2001). Such mechanism has been suggested to play a role in the aforementioned change of cortico-cerebellar connectivity that is observed following motor learning (Jayaram et al., 2011; Spampinato et al., 2017, 2020). Overall, sensory-based motor learning results from the iterative correction of movements based on SPEs, a process that is at play in multiple motor learning tasks and strongly relies on the cerebellum.

1.3.2. Reinforcement-based motor learning

In the last years, there has been an increased appreciation in the literature that motor learning is not only the result of sensory-based error corrections, but rather also relies on other learning processes (Huang et al., 2011; Wolpert et al., 2011; Morehead and Orban De Xivry, 2021; Spampinato and Celnik, 2021). One of these learning mechanisms is reinforcement learning (Schultz, 2015; Sutton and Barto, 2018). Reinforcement learning refers to the process through which appropriate actions are selected through outcome-based feedback (e.g., success or failure). The use of such reinforcement feedback to guide movements is an extremely ancient ability, which has thus been well-conserved throughout evolution

(Cisek, 2019). For instance, foraging in extremely various forms of life (e.g., insects, mollusks, reptiles, birds or human hunter-gatherer) is thought to rely on a basic reinforcement learning system allowing to select patches of food in the environment based on the outcomes of previous searches (Cisek, 2019). It is interesting to note that reinforcement learning appears as a very fundamental function for animals, but is also increasingly used by artificial systems to learn to solve a variety of contemporary problems (Littman, 2015; Neftci and Averbeck, 2019).

In this section, I will introduce basic concepts of reinforcement learning derived from the decision-making field, and then will present how these concepts have inspired recent research on reinforcement learning of motor skills.

1.3.2.1. Reinforcement during decision-making

Reinforcement learning has been initially described and largely investigated in the context of decision-making where participants have to choose between a discrete number of possible choices (Lee et al., 2012; O'Doherty et al., 2017). These works have led to the assumption that reinforcement learning occurs through the computation of reward prediction errors (RPE), which reflect the difference between the expected and obtained outcome following a decision. According to this framework, the values of potential choices are learned based on previous experience, allowing the learner to progressively choose actions yielding positive outcomes (Schultz et al., 1997; Rangel et al., 2008). At the neural level, a large body of literature has shown that RPEs cause phasic firing in dopaminergic neurons in the midbrain (mainly in the ventral tegmental area (VTA); Schultz et al., 1997; Matsumoto and Hikosaka, 2009; Cohen et al., 2012; Schultz, 2015) inducing dopamine release in the striatum (Mohebi et al., 2019; Lee et al., 2020). Interestingly, there is now ample evidence that such dopaminergic responses are not only sensitive to the mismatch between expected and predicted rewards but also to other factors such as the timing at which reward is provided (Fiorillo et al., 2008; Kobayashi and Schultz, 2008; Klein-Flügge et al., 2011; and see also Study 2), availability of rewards (Matsumoto et al., 2016) or uncertainty about the environment (Starkweather et al., 2017). Hence, dopaminergic responses to reinforcements integrate multiple factors to drive learning.

There is also evidence that frontal regions such as the ventro-medial prefrontal cortex (vmPFC), the orbitofrontal cortex (OFC) and the anterior cingulate cortex (ACC) contribute to reinforcement learning (Rangel et al., 2008; Lee et al., 2012; Husain and Roiser, 2018; Bongioanni et al., 2021), partially through the modulation of striatal activity (Haber et al., 2006; Neftci and Averbeck, 2019; Voloh et al., 2020; Averbeck and O'Doherty, 2022). Ultimately, the striatum is thought to integrate these different inputs to drive motor cortex activity through the basal ganglia (Lee et al., 2020). Accordingly, there is accumulating evidence that the motor cortex, a region initially thought to subserve merely motor functions, also encodes information related to reinforcements (Ramkumar et al., 2016; Derosiere et al., 2017c, 2017a; Ramakrishnan et al., 2017; Levy et al., 2020; Lee et al., 2021). Especially, in a previous study to which I contributed, we found that M1 is involved in reinforcement learning and that this involvement depends on the stage of learning (Derosiere et al., 2017a). Overall, this body of work shows that reinforcement learning strongly relies on activity in a distributed brain network centered on midbrain dopaminergic nuclei and fronto-striatal loops.

1.3.2.2. Reinforcement during motor learning

Reinforcement vs. reward in human studies

Before diving into the literature which has investigated reinforcement motor learning, it seems important to explain the conceptual difference between performance-based feedback (called “reinforcement feedback” in the present work) and extrinsic incentives (referred to as “reward”). As we have seen above, reinforcement provides information about the performance at a particular task and therefore can guide a learning process that evaluates outcomes to adjust future actions. Reinforcements can also provide motivation because knowledge of performance can be intrinsically rewarding (Leow et al., 2018). In contrast, additional extrinsic reward, which can be delivered for instance in the form of money (e.g., Wachter et al., 2009) or social praise (e.g., Sugawara et al., 2012), does not provide any additional direct learning signal, but rather only gives motivation to perform well (Berke, 2018). Despite such differences, most previous studies have coupled reinforcement and reward or provided only reinforcement, which did not allow considering the specific effect of motivation on motor learning, an issue that we addressed in Study 1. It is also important to

note that this distinction between the “learning” and “motivation” properties of incentives is less straightforward in the context of animal research, since training animals to perform a laboratory task requires extrinsic rewards. Still, this dissociation can be achieved experimentally, for instance through dynamic manipulations of reward rate (Mohebi et al., 2019).

Distinction between knowledge of performance and knowledge of results

In addition to this distinction between reinforcement and reward, some authors in the literature also distinguish “knowledge of results”, which provides information about the final outcome of a movement and “knowledge of performance”, which gives information about the movements characteristics that led to that outcome (Sharma et al., 2016). According to this terminology, both “knowledge of performance” and “knowledge of results” can be used as reinforcements but the difference is that “knowledge of performance” reinforces the specific kinematics leading to success whereas “knowledge of results” puts emphasis on the final outcome of the movement (Kitago and Krakauer, 2013). Importantly though, in both cases, the learner has to exploit information about task success (either after the completion of the movement for “knowledge of results” or online, during motor control for “knowledge of performance”) to adjust future motor commands. Hence, in the following sections, I do not distinguish between these two ways of providing feedback and consider that reinforcement feedback provides knowledge of performance in general (to refer to the informational properties of reinforcements), regardless of if it is provided after the completion of movements (as in Study 1 and 2) or online, during motor control (as in Study 3).

Evidence for reinforcement effects during motor learning

Research conducted in the last 15 years has shown that providing performance-based reinforcement feedback (often coupled to monetary rewards, see below) during motor training can improve motor learning abilities in various tasks including skill learning (Abe et al., 2011; Dayan et al., 2014a; Mawase et al., 2017), motor adaptation (Huang et al., 2011; Izawa and Shadmehr, 2011; Shmuelof et al., 2012; Wu et al., 2014; Galea et al., 2015a; Cashaback et al., 2017; Codol et al., 2018; Uehara et al., 2018) and sequence learning (Wachter et al., 2009; Sugawara et al., 2012; Sporn et al., 2022). In particular, such training

boosts motor skill acquisition (e.g., Mawase et al., 2017; Sporn et al., 2022) and consolidation (e.g., Abe et al., 2011b; Sugawara et al., 2012; Widmer et al., 2016), suggesting that reinforcements can act on multiple phases of motor learning.

The effect of reinforcement on the acquisition of a skill seems to be particularly prominent in situations where sensory feedback is uncertain (Izawa and Shadmehr, 2011; Bernardi et al., 2015; Nikooyan and Ahmed, 2015; Cashaback et al., 2017), implying that this approach could be promising for rehabilitation, as patients with motor impairments often also exhibit sensory deficits (Connell, 2008; Hepworth et al., 2016). Consistently, reinforcement has been shown to improve motor learning in different populations of patients including cerebellar (Therrien et al., 2016, 2018 but see also Vassiliadis et al., 2019) and stroke patients (Goodman et al., 2014; Quattrocchi et al., 2017; Widmer et al., 2021). Because of its crucial impact on motor learning and promising perspectives regarding rehabilitation, it is important to better characterize the mechanisms through which reinforcement feedback can improve motor learning (Chen et al., 2017; Johnson and Cohen, 2022).

Behavioral mechanisms of reinforcement motor learning

As explained above, reinforcement feedback provides information about previous performance and thereby can guide learning by pushing the agent to repeat actions leading to success and avoid those generating failure. Importantly, reinforcement feedback can also enhance motivation during training as performance feedback can be intrinsically rewarding. As I describe in this section, the informative and motivational properties of reinforcement feedback may influence motor learning through different mechanisms.

First, reinforcement feedback provides information about previous performance that can support motor learning by pushing the agent to repeat actions leading to success and avoid those generating failure. It is important to note that the learning signals used to drive learning are very different in sensory-based and reinforcement learning. Whereas sensory inputs provide directional information on how to correct movements (e.g., through vision), reinforcement feedback only gives information about the outcome of a movement, without informing the learner on how the movement should be corrected. Interestingly, these two types of learning rely on different behavioral (Izawa and Shadmehr, 2011; Nikooyan and

Ahmed, 2015) and neural (Sidarta et al., 2016; Therrien et al., 2016; Mathis et al., 2017; Uehara et al., 2018) mechanisms. Multiple lines of evidence show that reinforcement feedback can be used to explore different movement solutions following failure or to exploit motor commands that previously led to success (Wu et al., 2014; Therrien et al., 2016; Dhawale et al., 2017, 2019; Uehara et al., 2019; van Mastrigt et al., 2021). In motor learning tasks, this exploration/exploitation process can be measured by comparing the extent to which motor commands are modified from one trial to the next depending on the type of feedback obtained (Uehara et al., 2019; van Mastrigt et al., 2021, see also Study 1). By doing so, it is possible to dissociate the between-trial changes in movement parameters that are related to exploration to the one reflecting motor noise (defined as an unwanted source of motor variability; Therrien et al., 2016). Regulation of motor variability is thought to be a fundamental process in motor learning as it allows guiding the motor system toward new patterns of movements and motor activity that improve performance and may reduce costs (Wu et al., 2014; Shadmehr et al., 2016; Dhawale et al., 2017). This idea is also well supported by research investigating vocalization in song birds, which have shown that regulation of motor variability is crucial for learning (Ölveczky et al., 2005; Tumer and Brainard, 2007) but also for adaptation of the song to new contexts (e.g., birds sing with less vocal variability when they are close to potential partners compared to when they are “practicing” alone; Kao et al., 2005). Interestingly, the regulation of motor exploration is thought to play a particularly important role in situations with high uncertainty about the sensory consequences of actions (Izawa and Shadmehr, 2011) or about the environment and task (Bernardi et al., 2015; Sidarta et al., 2016; Dhawale et al., 2019). When there is uncertainty, exploration may be particularly useful to discover the reward landscape of the task, allowing one progressively to favor control policies yielding the best outcomes (Cashaback et al., 2017; Dhawale et al., 2019). Overall, regulation of motor variability appears to be a crucial mechanism underlying the beneficial effects of reinforcements on motor learning.

Regulation of motor variability makes use of the informative properties of reinforcement. Yet, as mentioned above, reinforcement feedback, even in the absence of extrinsic reward may also boost motivation to perform well and thereby influence motor learning (Leow et al., 2018). As such, recent studies investigating the effect of motivation on motor control have shown that when offered monetary rewards, participants are more prone to engage in costly control strategies such as regulation of feedback control gains (Carroll et

al., 2019; Comite et al., 2021), stiffness of the limbs (Codol et al., 2020; Vassiliadis and Derosiere, 2020) and internal motor noise (Manohar et al., 2015, 2019). These control policies allow movements to be simultaneously quicker and more accurate, apparently exceeding the limits of the speed-accuracy trade-off (Manohar et al., 2015, see Annex 2 for a scientific commentary written during my PhD on this topic). Such improvement of performance during training may translate into future non-rewarded contexts, potentially by boosting training-related plasticity (Mawase et al., 2017; Uehara et al., 2018). Another mechanism through which motivation could improve motor learning is by modulating the offline processing of motor memories (Abe et al., 2011). Further arguments for this hypothesis come from the decision-making literature, which has shown that the reward context in which a decision-making task is learned influences post-practice consolidation of the memory both during wakeful rest (Singer and Frank, 2009; Ambrose et al., 2016; Liu et al., 2021) and sleep (Igloi et al., 2015; Sterpenich et al., 2021). More specifically, the presence of reward increases the probability of neural replay of a memory (Igloi et al., 2015; Ambrose et al., 2016; Igata et al., 2021; Liu et al., 2021; Sterpenich et al., 2021) and the directionality of the replay (i.e., whether the replay occurs in the trained “forward” order or in the opposite “reverse” order; Ambrose et al., 2016; Liu et al., 2019; Igata et al., 2021), suggesting a brain mechanism for the prioritization of reward-related memories. Hence, performance of a motor task with reinforcement feedback could upregulate motivation (compared to a situation with only sensory feedback) and thereby boost offline consolidation of the skill. Put together, these elements support the view that the motivational properties of reinforcements may contribute to reinforcement-related benefits in motor learning, independently from the regulation of motor variability.

Neural correlates of reinforcement motor learning

In this part, I present what we currently know about the neural correlates of reinforcement motor learning. As explained above, previous studies have not systematically dissociated the “learning” and “motivational” properties of reinforcement feedback. For this reason, I present here insights into the neural mechanisms of reinforcement motor learning in general, without being able to systematically dissociate the neural correlates of these two aspects.

At the neural level, a key area involved in motor (Hardwick et al., 2013) and reinforcement learning (Schultz, 2015) and which could therefore be well-suited to mediate reinforcement-related benefits in motor learning is the striatum. Anatomically, the striatum can be decomposed into a ventral (nucleus accumbens and olfactory tubercle) and dorsal part (neostriatum: caudate nucleus and putamen). Thanks to its dense connexions with dopaminergic as well as with pre-frontal and motor cortical regions, the striatum is thought to play a key role in translating information about reinforcements into motor behaviors (Nakamura and Hikosaka, 2006; Williams and Eskandar, 2006; Piray et al., 2017; Hori et al., 2019). This idea is in line with previous neuroimaging studies showing reward-related activations of the striatum during motor learning (Wachter et al., 2009; Widmer et al., 2016²). Consistently, reinforcement-dependent motor exploration is disrupted by Parkinson's disease (a model of striatal dysfunction; Pekny et al., 2015), and by dopaminergic antagonists in humans (Galea et al., 2013; but see also Quattrocchi et al., 2018 and Palidis et al., 2021 for null effects of dopaminergic manipulations on reinforcement motor learning). Moreover, studies conducted on songbirds indicate that motor exploration causally relies on a basal ganglia-related circuit that injects variability into motor cortex, and thereby drives behavioral variability necessary for learning (Kao et al., 2005; Ölveczky et al., 2005). Finally, micro-stimulation of the striatum during reinforcement presentation can modulate associative learning in non-human primates (Nakamura and Hikosaka, 2006; Williams and Eskandar, 2006). Put together, these studies suggest that the striatum is causally involved in the generation of motor adjustments based on previous outcomes.

As mentioned earlier, the striatum is densely connected to multiple cortical areas that are likely to play an important role in reinforcement learning of motor skills (Averbeck and O'Doherty, 2022). As such, resting-state functional connectivity of the striatum with secondary motor cortical areas (Sidarta et al., 2016; Steel et al., 2019) and with the vmPFC (Sidarta et al., 2016) is modulated following reinforcement motor learning, but not sensory-based learning, suggesting that reinforcement learning of motor skills leads to plastic modifications in fronto-striatal pathways. Consistent with an involvement of pre-frontal regions in reinforcement motor learning, a previous study found that the propensity to benefit from reinforcement in a motor learning task co-varies with inter-individual variability in grey-matter volume

² Note the interesting increase of striatal activity in the presence of reinforcement feedback paired with reward in this study, compared to the reinforcement only condition, suggesting that increasing motivation during motor training (through extrinsic reward) leads to a boosting of striatal activity.

of the lateral prefrontal cortex (Dayan et al., 2014a). In the same line, non-invasive brain stimulation of the fronto-polar cortex modulates reinforcement motor learning through a change in motor exploration (Herrojo Ruiz et al., 2021). Notably, there is also evidence that providing reinforcements during motor learning can influence M1 activity (Levy et al., 2020; Lee et al., 2021) and plasticity in response to practice (Mawase et al., 2017; Uehara et al., 2018; Spampinato et al., 2019). More specifically, Uehara et al., (2018) found that while sensory-based motor learning led to cerebellar plasticity, reinforcement motor learning generated LTP-like plasticity within M1. This points towards the idea that the presence of reinforcement during a motor training leads to specific forms of plasticity that involve the motor cortex. Such boosting of M1 plasticity, may rely on the activation of dopaminergic neurons in the VTA that project to M1 and are known to be crucial for motor skill acquisition (Hosp et al., 2011, 2015; Leemburg et al., 2018). Interestingly, a very recent study suggested that reinforcement-related M1 plasticity following motor learning also relies on cholinergic neurotransmission (Bowles et al., 2022). Hence, pre-frontal as well as motor cortical regions are specifically involved in reinforcement learning of motor skills, potentially through their interactions with the striatum as well as with dopaminergic and cholinergic populations of neurons.

Finally, another structure that has been largely considered in the context of sensory-based motor learning (Shadmehr et al., 2010b), but less in the framework of reinforcement learning is the cerebellum. Yet, as I have argued in a scientific commentary published in the beginning of my thesis (Vassiliadis et al., 2019; Annex 3), there is accumulating evidence that the cerebellum is involved in multiple cognitive domains including reinforcement learning (Swain et al., 2011). Anatomically, this idea is supported by the presence of bidirectional connexions between the cerebellum and dopaminergic basal ganglia structures including the striatum (Bostan and Strick, 2018). Functionally, recent works provide evidence that rewards and reward predictions influences the firing rate of cerebellar cells in rodents (Ohmae and Medina, 2015; Wagner et al., 2017; Heffley et al., 2018; Kostadinov et al., 2019; Kostadinov and Häusser, 2022) and cerebellar activity in humans (O'Doherty, 2004; Ramnani et al., 2004; Seymour et al., 2004; Tanaka et al., 2004; Tobler and O'Doherty, 2006; Garrison et al., 2013), suggesting that this structure is involved in processing reinforcement feedback. Consistently, cerebellar patients exhibit altered reinforcement learning

in a decision-making task requiring very simple movements (Thoma et al., 2008). Interestingly, it appears that these patients can use reinforcement feedback to adapt to a visuomotor perturbation, but to a lesser degree than age-matched controls (Therrien et al., 2016). The results of a follow-up study then suggested that this impairment could not be completely accounted for by an increase in motor noise in patients (defined as an unwanted source of motor variability, Therrien et al., 2018), but may rather result from a true deficit in reinforcement motor learning following cerebellar damage (Vassiliadis et al., 2019). Hence, the cerebellum, known to be involved in error corrections based on sensory feedback may participate in the integration of sensory and reinforcement information to guide motor learning.

Overall, these studies indicate that learning motor skills through reinforcements engages partially different patterns of neural activity than pure sensory-based motor learning. More specifically, this type of learning is thought to strongly rely on fronto-striatal interactions and may also involve communication between the cerebellum and dopaminergic circuits (Carta et al., 2019).

1.4. Plan of the PhD

Study 1: Reward boosts reinforcement-based motor learning

As mentioned above, multiple studies have shown that reinforcement feedback (often coupled to monetary rewards) could improve several aspects of motor learning. In this behavioral study, we aimed at dissociating the contributions of reinforcement feedback (providing knowledge of performance) and reward (providing motivation) during motor skill learning in a large sample of healthy participants. Our experimental paradigm also allowed us to evaluate the effect of reinforcement and reward on regulation of motor variability on a trial-by-trial basis. Overall, our results show that motivation by reward can have persistent effects on motor learning and that these effects are related to an enhanced regulation of motor variability based on previous outcomes that lasts after reward removal.

Study 2: Reward timing matters in motor learning

In this work, we asked if reward timing (i.e., the delay between the end of a movement and the ensuing feedback), which is known to strongly influence dopaminergic responses and learning in classical conditioning paradigms (Fiorillo et al., 2008; Kobayashi and Schultz, 2008) could also impact motor skill learning. Importantly, based on the results of Study 1, in this work, we have coupled reinforcement feedback to a monetary reward, allowing us to specifically test the effect of reward delay in a situation where participants would learn the task. Hence, in this study, we refer to “reward” to designate the combination of reinforcement feedback coupled to a monetary reward. As described in Chapter 3, we found that delaying reward delivery by only a few seconds strongly impacts the dynamics of learning and consolidation 24h later.

Study 3: Non-invasive stimulation of the human striatum disrupts reinforcement learning of motor skills

Here, we investigated the causal contribution of the striatum in reinforcement motor learning. Indeed, previous evidence implicating the striatum in reinforcement motor learning was based on correlational approaches showing a relationship between striatal activity and the presence of reinforcement (Wachter et al., 2009; Widmer et al., 2016), but causal evidence for a role of the striatum in this process was lacking. In this combined brain stimulation and neuroimaging study we leverage a new type of non-invasive brain stimulation, called transcranial temporal electric interference stimulation (tTIS), that allows to target deep brain structures without stimulation of the overlying cortex (Grossman et al., 2017). Conventional brain stimulation techniques, such as transcranial electrical current stimulation (tES) or TMS (used in some of my previous work, e.g., see Annex 1) mostly activate the cortex, because the strength of the electric/magnetic field strongly decays as a function of the distance to stimulation (Figure 3A, B). Notably, some TMS protocols have been shown to engage deep brain structures (Strafella et al., 2003), but in these cases stimulation also causes significant concomitant engagement of the cortex (Mas-Herrero et al., 2021). During my PhD, I have been involved in the validation of tTIS stimulation in humans (Wessel et al., 2021) and used it to test my own research questions in the context of Study 3. Below, I describe the basic

principles of tTIS stimulation to provide some background on this new brain stimulation approach.

Note on transcranial electric temporal interference stimulation

During tTIS stimulation, two high frequency electrical currents oscillating at f_1 and $f_1 + \Delta f$ are applied via two pairs of electrodes connected to independent stimulators. These high-frequency currents are thought to be too high to directly modulate neuronal activity (Grossman et al., 2017; Grossman, 2018). The two currents result in a modulated electric field deep in the brain with the envelope oscillating at Δf (Figure 3C). By adjusting electrode positioning, the modulated electric field can be steered towards a target location. Thanks to an ongoing collaboration with the team of Prof. Esra Neufeld, we determined via simulations on realistic head models the optimal electrode montage to target the striatum in humans (F3-F4, TP7-TP8; Figure 3D). We applied tTIS in healthy participants via surface electrodes applying a low-intensity protocol respecting the currently accepted cut-offs and safety guidelines for transcranial electric stimulation in humans (Antal et al., 2017; Figure 3E).

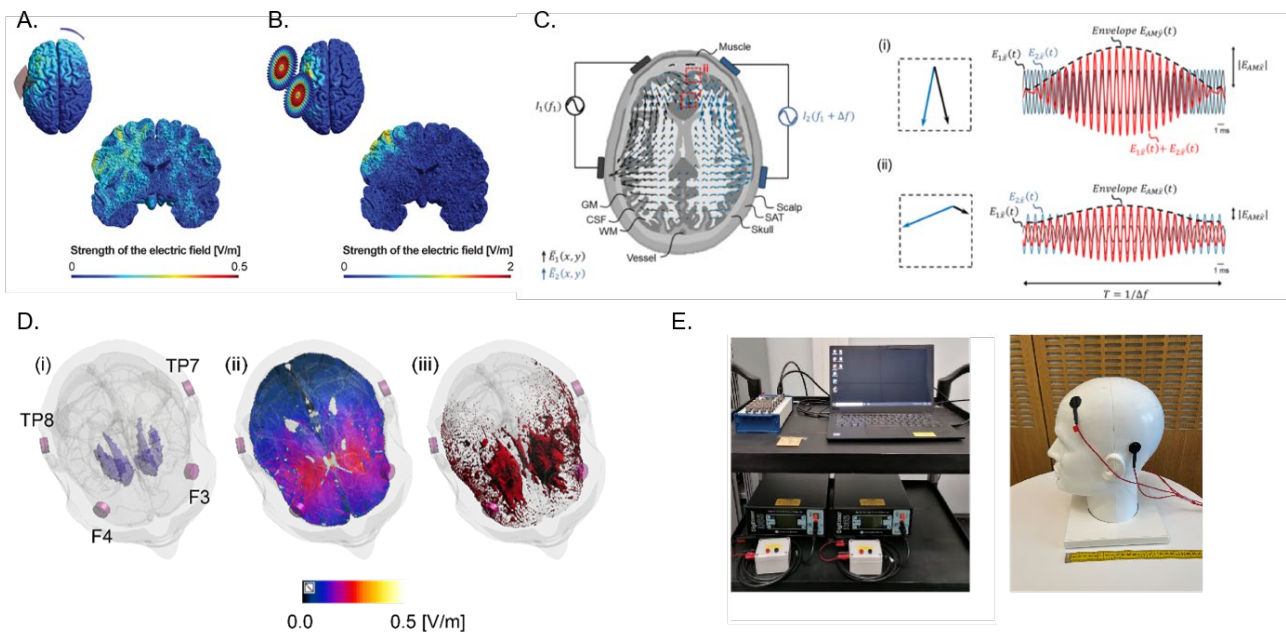


Figure 3. Stimulation of deep brain structures with conventional non-invasive brain stimulation vs. tTIS. Electric field simulations for an exemplary tDCS (A) and TMS (B) protocol implemented in the open source software package SimNIBS (analysis provided by Dr. Maximilian Wessel). TMS coil were centered over the left sensorimotor region. Common stimulation intensities were used as

input for the models: 1 mA (A) and 1×10^6 A/s (B). The electric field maxima are located in proximity to the brain cortex. The induced electric field does not reach a magnitude required for modulation of neuronal activity deeper in the brain. **(C)** tTIS strategy (illustration provided by Prof. Grossman, Imperial College London). Two sinusoidal electric fields in the kHz-range (I1 and I2) are generated via isolated current sources and are applied transcranially through scalp electrodes. When the fields are applied with a slight difference in frequencies (Δf), the superposition of the fields can lead to a slowly “beating” envelope oscillating at Δf . The peak amplitude of this envelope modulation can be focused towards deeper brain targets and drive their neuronal activity. **(D)** Electrode montage optimization for striatal tTIS (Prof. Neufeld, ETZ). The selected montage was the one displaying the best trade-off between stimulation strength in the striatal target and focality of stimulation. Notably, simulations were run for 1 mA stimulation whereas stimulation intensities used in the current work were 2 mA. The expected electrical field in the striatal target with our stimulation protocol are therefore in the range of 0.5-0.6 V/m. **(E)** tTIS hardware and electrode montage for striatal stimulation in humans used in Study 3.

In Study 3, we tested the causal contribution of the striatum to reinforcement motor learning by comparing learning abilities in a force-tracking task depending on the presence of reinforcement and the type of tTIS applied during training (tTIS_{Sham}, tTIS_{20Hz} or tTIS_{80Hz}). Overall, we find that high gamma, not beta, tTIS of the striatum disrupts the benefits of reinforcement on motor learning. This effect was associated to changes of activity in the striatum, as recorded with functional magnetic resonance imaging. Moreover, we also show that tTIS_{80Hz}, but not tTIS_{20Hz}, increases effective connectivity between the striatum and the frontal cortex. Overall, these results show for the first time that tTIS is able to target the striatum with relevant behavioral and neural effects on reinforcement learning.

2. Study 1: Reward boosts reinforcement-based motor learning

Pierre Vassiliadis^{1,2*}, Gerard Derosiere¹, Cecile Dubuc¹, Aegryan Lete¹, Frederic Crevecœur^{1,3}, Friedhelm C Hummel^{2,4,5}, Julie Duque¹

¹ Institute of Neuroscience, Université Catholique de Louvain, 1200, Brussels, Belgium.

² Defitech Chair for Clinical Neuroengineering, Center for Neuroprosthetics (CNP) and Brain Mind Institute (BMI), Swiss Federal Institute of Technology (EPFL), 1202, Geneva, Switzerland

³ Institute of Information and Communication Technologies, Electronics and Applied Mathematics, Université Catholique de Louvain, Louvain-la-Neuve 1348, Belgium.

⁴ Defitech Chair for Clinical Neuroengineering, Center for Neuroprosthetics (CNP) and Brain Mind Institute (BMI), Swiss Federal Institute of Technology Sion (EPFL), 1951, Sion, Switzerland

⁵ Clinical Neuroscience, University of Geneva Medical School (HUG), 1202, Geneva, Switzerland

Reference: Vassiliadis P, Derosiere G, Dubuc C, Lete A, Crevecœur C, Hummel FC, Duque J (2021). Reward boosts reinforcement-based motor learning. iScience, 24:102821.

Personal contribution: Conceptualization, Methodology, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Funding acquisition.

* Correspondence

Pierre Vassiliadis

Institute of Neuroscience

Université catholique de Louvain

53, Avenue Mounier

COSY- B1.53.04

B-1200 BRUSSELS

Belgium

Email: pierre.vassiliadis@uclouvain.be

Web: <http://www.coactionslab.com>

2.1. Abstract

Besides relying heavily on sensory and reinforcement feedback, motor skill learning may also depend on the level of motivation experienced during training. Yet, how motivation by reward modulates motor learning remains unclear. In 90 healthy subjects, we investigated the net effect of motivation by reward on motor learning while controlling for the sensory and reinforcement feedback received by the participants. Reward improved motor skill learning beyond performance-based reinforcement feedback. Importantly, the beneficial effect of reward involved a specific potentiation of reinforcement-related adjustments in motor commands, which concerned primarily the most relevant motor component for task success and persisted on the following day in the absence of reward. We propose that the long-lasting effects of motivation on motor learning may entail a form of associative learning resulting from the repetitive pairing of the reinforcement feedback and reward during training, a mechanism that may be exploited in future rehabilitation protocols.

2.2. Introduction

Motor skill learning is the process by which the speed and accuracy of movements improve with practice (Krakauer et al., 2019). A significant amount of research has since long demonstrated that motor learning relies on sensory feedback, which allows reducing movement errors (*e.g.*, (Tseng et al., 2007; Shadmehr et al., 2010b)). More recently, some studies have shown that reinforcement feedback, allowing the adjustment of movements based on knowledge of performance, also plays a role in motor learning (Wachter et al., 2009; Palminteri et al., 2011; Bernardi et al., 2015; Galea et al., 2015b; Therrien et al., 2016; Mawase et al., 2017). The contribution of reinforcement feedback to motor learning seems to be particularly important when the quality of the available sensory feedback is low (Izawa and Shadmehr, 2011; Cashaback et al., 2017). These observations suggest that reinforcement feedback may be critical for motor rehabilitation (Quattrocchi et al., 2017; Roemmich and Bastian, 2018), where patients often exhibit sensory impairments in addition to their motor disability (Connell, 2008; Hepworth et al., 2016). However, before clinical translation can occur, significant research is required to characterize the optimal conditions in which sensory and reinforcement feedback can improve motor learning.

One key factor that may influence sensory- and reinforcement-based motor learning is motivation (Lewthwaite and Wulf, 2017). This idea is in line with an ethological perspective: in nature, animals are motivated to learn efficiently motor behaviors that have been repetitively associated with rewarding outcomes, in order to increase the likelihood of reaching these outcomes again in the future (Barron et al., 2010; Yamazaki et al., 2016). Whereas past research on motivation has traditionally focused on the impact of reward on decision-making (Shima and Tanji, 1998; Bush et al., 2002; O'Doherty, 2004; Dayan and Niv, 2008; Balleine and O'Doherty, 2010; Hare et al., 2011; Padoa-Schioppa, 2011; Schultz, 2015; Derosiere et al., 2017c, 2017a; Gershman and Daw, 2017), there has been a recent rise in interest regarding its influence on motor learning (Therrien et al., 2016, Mawase et al., 2017, Uehara et al., 2019; Vassiliadis et al., 2019; Chen et al., 2017; Sporn et al., 2020, Vassiliadis and Derosiere, 2020, Holland et al., 2019).

To tackle this issue, previous studies have investigated motor skill learning with different types of reinforcement and reward. This research showed that the combination of reinforcement (providing knowledge of performance) and reward (providing motivation) can influence motor skill learning (e.g., (Wachter et al., 2009; Abe et al., 2011; Wilkinson et al., 2015; Steel et al., 2016, 2019)). A key aspect of the aforementioned studies is that they considered reinforcement and reward in a bonded way, with the rewarded participants being also the ones receiving performance-based reinforcement feedback. The assumption underlying this approach is that receiving knowledge of performance (e.g., points or binary feedback) provides a form of intrinsic reward that by itself increases motivation to perform well (Leow et al., 2018). However, in addition to the intrinsically rewarding properties of reinforcement, knowledge of performance also provides a learning signal to the motor system, that can influence motor learning (Huang et al., 2011; Shmuelof et al., 2012; Bernardi et al., 2015; Galea et al., 2015b; Nikooyan et al., 2015; Therrien et al., 2016; Mawase et al., 2017; Leow et al., 2018; Uehara et al., 2018; Kim et al., 2019). In contrast, extrinsic reward increases motivation to perform well, without conveying any additional learning signal (Berke, 2018). In accordance with a dissociable role of reinforcement and reward in motor learning, past research has shown that certain subpopulations of neurons in the motor cortex (*i.e.*, a key region of the motor learning network; (Krakauer et al., 2019)) are responsive to the outcome of previous movements irrespective of reward (Levy et al., 2020), while others respond to reward regardless of the previous outcome (Ramkumar et al., 2016). Put together, these elements indicate that estimating the net impact of reward on motor learning requires controlling for the effect of the reinforcement feedback on the learning process. Based on these elements, we experimentally uncoupled knowledge of performance from reward to test the hypothesis that reward induces a specific improvement in motor learning and maintenance.

Another important question relates to how, at the single-trial level, motivation by reward may affect motor skill learning and maintenance. As such, computational models of motor learning posit that movement errors can be corrected based on sensory and reinforcement feedbacks on a trial-by-trial basis (Cashaback et al., 2017), with possible interactions between these two processes (Izawa and Shadmehr, 2011). Sensory-based motor learning relies on the ability to produce motor commands that match predicted sensory consequences (e.g., visual, somatosensory consequences; (Sidarta et al., 2016, Bernardi et al.,

2015). Conversely, reinforcement-based motor learning is thought to depend on the ability to efficiently regulate between-trial motor variability based on previous outcomes (Wu et al., 2014; Pekny et al., 2015; Sidarta et al., 2016; Therrien et al., 2016; Dhawale and Smith, 2017; Uehara et al., 2019). Importantly, in this framework, reward may have a global influence, enhancing both sensory- and reinforcement-based adjustments from one trial to another, or could have a more specific effect, boosting only one of the two learning systems (Galea et al., 2015b; Kim et al., 2019). Here, we investigated the impact of reward on sensory- and reinforcement-based motor adjustments during motor skill learning at the single-trial level, in a situation where they can both contribute to the learning process.

Healthy participants ($n = 90$) trained on a pinch-grip force reproduction task with limited sensory feedback over two consecutive days, while we manipulated the reinforcement feedback and reward on Day 1. By removing visual feedback on most trials, we ensured that the learning process would largely depend on the integration of somatosensory and reinforcement feedbacks (Izawa and Shadmehr, 2011; Bernardi et al., 2015; Sidarta et al., 2018). Moreover, subjects were distributed in three groups where training involved sensory (S) feedback only (Group-S; $n = 30$), sensory and reinforcement (SR) feedback (Group-SR; $n=30$), or both feedbacks and a reward (SRR, Group-SRR; $n=30$). Monetary gains were used as they are known to strongly modulate the motivation to engage in various tasks (Grogan et al., 2020; Manohar et al., 2015; Schultz, 2015; Shadmehr et al., 2019). We investigated how participants learned and maintained the skill depending on the type of feedback experienced during training. We found that while sensory and reinforcement feedbacks were not sufficient for the participants to learn the task in the present study, adding reward during training markedly improved motor performance. Reward-related gains in motor learning were maintained on Day 2, even if subjects were no longer receiving a reward on that day. Importantly, single-trial analyses showed that reward specifically increased reinforcement-related adjustments in motor commands, with this effect being maintained on Day 2, in the absence of reward. The pinch-grip force task used here also allowed considering adjustments separately for the speed of force initiation, and the accuracy of the performed force, both in terms of variability and amplitude. Importantly, we found that reward did not affect the control of all motor components in the same way, with the amplitude component turning out to be the more strongly influenced by the presence of reward.

Altogether, the present results provide evidence that motivation by reward can improve motor skill learning and maintenance even when the task is performed with the same knowledge of performance. More importantly, this effect seems to entail a specific potentiation of reinforcement-related adjustments in the motor command at the single-trial level. These behavioral results are important to characterize the mechanisms by which reward can improve motor learning and may guide future motivational interventions for rehabilitation (McGrane et al., 2015).

2.3. Results

Ninety healthy participants practiced a pinch-grip force task over two consecutive days. The task required participants to hold a pinch grip sensor in their right hand and to squeeze it as quickly as possible in order to move a cursor displayed on a computer screen in front of them, from an initial position to a fixed target (**Figure 1A**). The force required to reach the target ($\text{Target}_{\text{Force}}$) corresponded to 10 % of the individual maximum voluntary contraction (MVC). In most of the trials (90 %), participants practiced the task with very limited sensory feedback: the cursor disappeared when the generated force reached half of the $\text{Target}_{\text{Force}}$. In the remaining trials (10 %; not considered in the analyses), full vision of the cursor allowed participants to be visually guided towards the $\text{Target}_{\text{Force}}$ and therefore to be reminded of the somatosensory sensation corresponding to the $\text{Target}_{\text{Force}}$. Hence, in this task, learning relied mostly on the successful reproduction of the $\text{Target}_{\text{Force}}$ based on somatosensory feedback (Raspovic et al., 2014), with the target somatosensory sensation being regularly reminded to the participants through the full vision trials. To learn the task, subjects were provided with six training blocks (40 trials each; *i.e.*, total of 240 training trials; **Figure 1B**), during which Group-S subjects trained with sensory feedback only (Block-S), Group-SR subjects trained with sensory and reinforcement feedback (Block-SR), and Group-SRR subjects trained with both feedbacks and a monetary reward (Block-SRR). Notably, the groups were not significantly different for a variety of features including age, gender, $\text{Target}_{\text{Force}}$, difficulty of the task, muscular fatigue and final monetary gains (see Materials and Methods, **Table 1**). Beside the training blocks, all participants performed the task in a Block-SR setting so that the familiarization, the pre- and post-training assessments on Day 1, as well as Re-test on Day 2, occurred in the same conditions in the three experimental groups. This design allowed us to investigate the effect of reinforcement and reward, both on learning and on maintenance of the learned motor skill. Importantly, in Block-SR and Block-SRR, the binary reinforcement feedback depended on the Error, estimated as the absolute difference between the $\text{Target}_{\text{Force}}$ and the exerted force over the whole trial (excluding the first 150 ms, **Figure 1C**; (Abe et al., 2011; Steel et al., 2016). Hence, to be successful, participants had to reduce the Error by approximating the $\text{Target}_{\text{Force}}$ as quickly and accurately as possible.

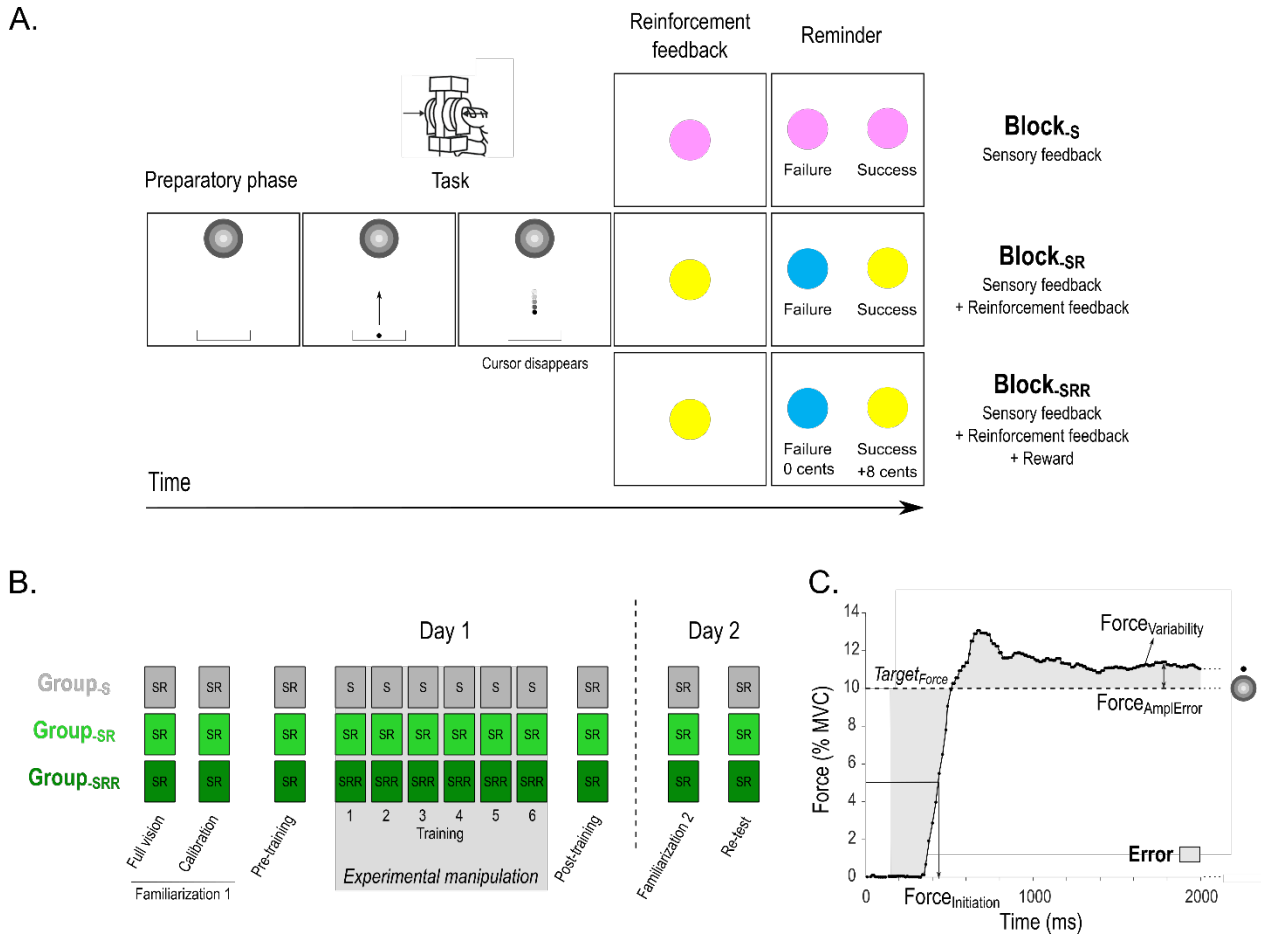


Figure 1. Experimental protocol. A) Time course of a trial in the motor skill learning task. Each trial started with the appearance of a sidebar and a target. After a variable preparatory phase (800-1000ms), a cursor appeared in the sidebar, playing the role of a “Go” signal. At this moment, participants were required to pinch the force transducer to bring the cursor into the target as quickly as possible and maintain it there until the end of the task (2000ms). Notably, on most trials, the cursor disappeared halfway towards the target (as displayed here). Then, a reinforcement feedback was provided in the form of a colored circle for 1000ms and provided binary knowledge of performance (Success or Failure in Block-SR and Block-SRR) or was non-informative (Block-S). The reinforcement feedback was determined based on the comparison between the Error on the trial and the individual success threshold (computed in the Calibration block, see Materials and Methods). Finally, each trial ended with a reminder of the color/feedback association and potential reward associated to good performance (1500ms). **B) Experimental procedure.** On Day 1, all participants performed two familiarization blocks in a Block-SR condition. The first one involved full vision of the cursor while the second one provided only partial vision and served to calibrate the difficulty of the task on an individual basis (See Materials and Methods). Then, Pre- and Post-training Block-SR assessments were separated by 6 blocks of training in the condition corresponding to each individual group (Block-S for Group-S, Block-SR for Group-SR and Block-SRR for Group-SRR). Day 2 involved a Familiarization block (with partial vision) followed by a Re-test assessment (4 Block-SR pooled together). There was no recalibration on Day 2. **C) Example of a force profile.** Force applied (in % of MVC) during the task. Participants were asked to approximate the $Target_{Force}$ as quickly and accurately as possible to minimize the Error (grey shaded area). As shown on the Fig, this Error depended on the speed of force initiation ($Force_{Initiation}$) and on the accuracy of the maintained force, as reflected by its amplitude with respect to the $Target_{Force}$ ($Force_{AmplError}$) and its variability ($Force_{Variability}$). Note that the first 150ms of each trial were not considered for the computation of the Error.

Reward improves motor skill learning

Participants' initial performance was not significantly different in the different groups: the Error in the Pre-training block equaled 3.14 ± 0.18 % MVC in Group-S, 3.33 ± 0.17 % MVC in Group-SR and 3.30 ± 0.15 % MVC in Group-SRR (one-way ANOVA: $F_{(2,87)} = 0.37$, $p = 0.69$, partial $\eta^2 = 0.0084$; **Figure 2A**). In contrast, skill learning, estimated as the training-related reduction in Error on Day 1 (Normalized Error = Post-training Error expressed in % of Pre-training Error) varied as a function of the group (**Figure 2B**). As such, learning was stronger in the Group-SRR compared to the two other groups (ANOVA: $F_{(2,87)} = 4.41$, $p = 0.015$, partial $\eta^2 = 0.092$; post-hocs: Group-SRR vs. Group-SR: $p = 0.014$, Cohen's $d = 0.60$; Group-SRR vs. Group-S: $p = 0.010$, $d = 0.98$), with no significant difference between Group-S and Group-SR ($p = 0.91$, $d = 0.025$). This was confirmed by a subsequent analysis showing that learning was significant in the Group-SRR (Post-training = 80.7 ± 3.5 % of Pre-training; single-sample t-test against 100 %: $t_{(29)} = -5.49$, $p < 0.00001$, $d = -1.42$), but not in Group-S (Post-training = 103.9 ± 5.15 % of Pre-training; $t_{(29)} = 0.75$, $p = 0.46$, $d = 0.19$) or in Group-SR (Post-training = 102.9 ± 8.80 % of Pre-training; $t_{(29)} = 0.33$, $p = 0.745$, $d = 0.085$). Skill maintenance on Day 2, estimated as the Error at Re-test in percentage of Pre-training, was not significantly different between the groups ($F_{(2,87)} = 1.96$, $p = 0.15$, partial $\eta^2 = 0.043$; **Figure 2C**). However, in Group-SRR, we found that the Error at Re-test remained lower than at Pre-training (Re-test = 85.6 ± 5.01 % of Pre-training; single-sample t-test against 100 %: $t_{(29)} = -2.88$, $p < 0.0073$, $d = -0.74$) demonstrating that the skill was maintained, while this effect was not significant in the two other groups (Group-S: Re-test = 100.5 ± 4.63 % of Pre-training; $t_{(29)} = 0.11$, $p = 0.92$, $d = -0.12$, Group-SR: Re-test = 97.0 ± 6.82 % of Pre-training; $t_{(29)} = -0.45$, $p = 0.66$, $d = -0.028$). Hence, while reinforcement alone did not contribute to reduce the Error in this task, its combination with reward successfully helped participants to learn and maintain the skill, as also evident when considering the averaged success rates (**Figure 2D**) and individual force profiles (**Figure 2E**).

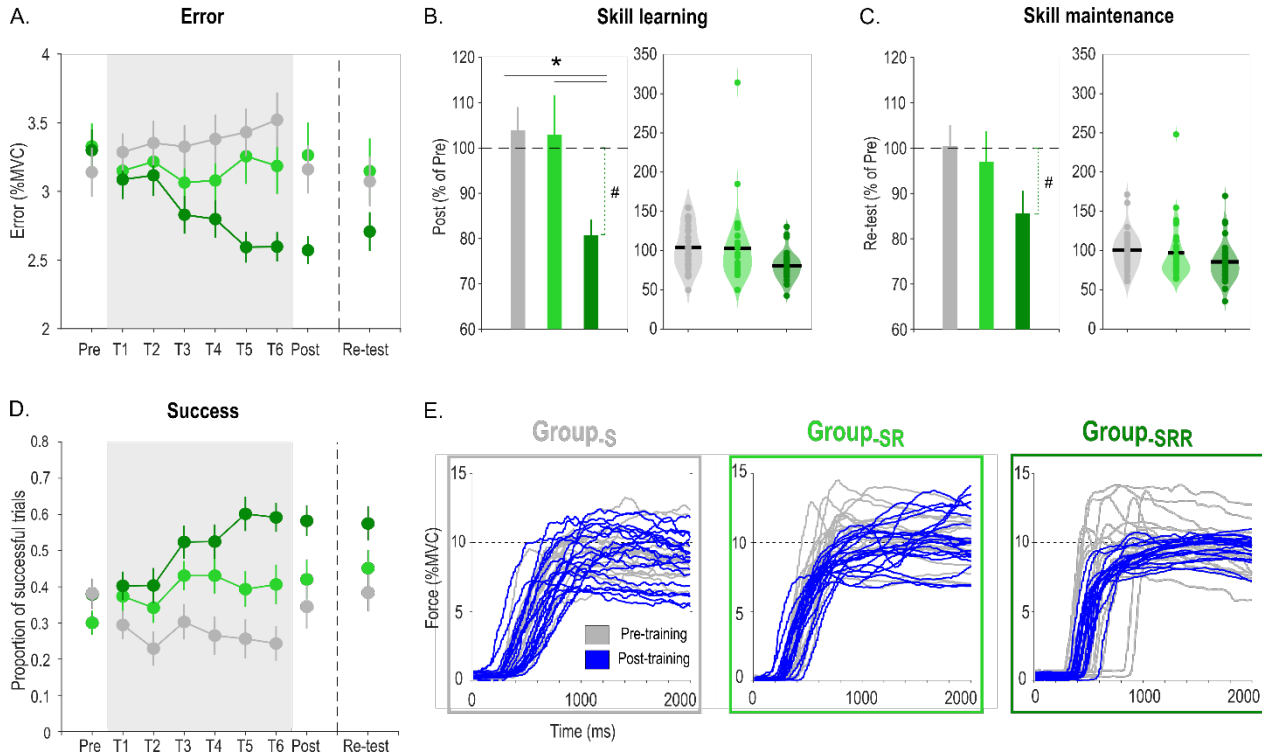


Figure 2. Learning curves. A) Error. Average Error is represented across practice for the three experimental groups (grey: Group-S, light green: Group-SR, dark green: Group-SRR). The grey shaded area highlights the blocks concerned by the reinforcement manipulation. The remaining blocks were performed with knowledge of performance only (*i.e.*, in a Block_{SR} setting). **B) Skill learning.** Bar plot (left) and violin plot (right, each dot = one subject) representing skill learning (quantified as the Error in Post-training blocks expressed in percentage of Pre-training blocks) in the three experimental groups. Skill learning was significantly enhanced in Group-SRR compared to the two other groups. This result remained significant when removing the subject showing an extreme value in the Group-SR (ANOVA: $F_{(2,86)} = 6.44$, $p = 0.0025$, partial $\eta^2 = 0.13$; post-hocs; Group-SRR vs. Group-SR: $p = 0.027$; Group-SRR vs. Group-S: $p = 0.00064$; Group-SR vs. Group-S: $p = 0.21$). **C) Skill maintenance.** Bar plot (left) and violin plot (right) representing skill maintenance quantified as the Error in Re-test blocks expressed in percentage of Pre-training blocks) in the three experimental groups. **D) Success.** Proportion of successful trials for each block. **E) Force profiles.** Individual force profiles of one representative subject of Group-S (left), Group-SR (middle) and Group-SRR (right) in the Pre- (grey) and Post-training blocks (blue). Note the better approximation of the Target_{Force} and the reduced inter-trial variability at Post-training in the exemplar subject of Group-SRR. *: significant difference between groups ($p < 0.05$). #: significant difference within a group between normalized Post-training Error and a constant value of 100% ($p < 0.017$ to account for multiple comparisons).

Reward boosts reinforcement-related adjustments during motor skill learning

To identify the mechanisms at the basis of the effect of reward on motor learning, we quantified how much participants adjusted motor commands based on reinforcement or sensory feedback at the single-trial level. This allowed us to estimate how subjects relied on

each type of feedback on a trial-by-trial basis and how this behaviour was affected by reward. In order to investigate reinforcement-related adjustments in motor commands, we computed the absolute between-trial change (BTC) in Error ($\text{Error}_{\text{BTC}} = |\text{Error}_{n+1} - \text{Error}_n|$) following successful or failed trial_n of similar Error in the three groups (Materials and Methods, see also (Pekny et al., 2015; Uehara et al., 2019) for similar approaches in reaching tasks). Comparing $\text{Error}_{\text{BTC}}$ depending on the Outcome of the previous trial (Success or Failure) allowed us to estimate how much participants modified their force profile based on the reinforcement feedback. Notably, considering changes in the Error in absolute terms allowed us to explore the effect of reward on the magnitude of the adjustments in the different groups, regardless of their directionality (increase or decrease in the Error). We found that $\text{Error}_{\text{BTC}}$ was generally higher after failed than successful trials (two-way ANOVA; main effect of Outcome: $F_{(1,84)} = 8.66$, $p = 0.0042$, partial $\eta^2 = 0.093$; **Figure 3A**), consistent with an exploration process following failed trials (Uehara et al., 2019, Pekny et al., 2015). Interestingly, this difference between failed and successful trials was modulated by Group_{TYPE} (Outcome x Group_{TYPE}: $F_{(2,84)} = 11.47$, $p < 0.001$, partial $\eta^2 = 0.21$): while it was significant in Group-SR and Group-SRR (post-Success vs. post-Failure: $p = 0.028$, $d = -0.54$ and $p < 0.001$, $d = -0.92$ respectively), this effect was only at the trend level for Group-S ($p = 0.060$, $d = 0.29$). Relatedly, post-hoc tests revealed that post-Success $\text{Error}_{\text{BTC}}$ was significantly lower in Group-SRR than in Group-S ($p = 0.026$, $d = -0.54$), but not different between Group-SR and Group-S ($p = 0.24$, $d = -0.23$) and between Group-SR and Group-SRR ($p = 0.30$, $d = 0.29$). Besides, post-Failure $\text{Error}_{\text{BTC}}$ was significantly higher in Group-SRR than in Group-S ($p = 0.040$, $d = 0.65$). Yet, it was not different between Group-SR and Group-S ($p = 0.14$, $d = 0.48$) and between Group-SR and Group-SRR ($p = 0.58$, $d = -0.16$). Hence, providing a reward on top of reinforcement feedback led to a particularly low $\text{Error}_{\text{BTC}}$ following successful trials and a particularly high $\text{Error}_{\text{BTC}}$ following failed trials. This analysis suggests that reward modulated between-trial changes in behavior in response to the reinforcement feedback, regardless of whether reinforcement was positive or negative. To further confirm this, we directly compared the magnitude of reinforcement-based adjustments between the three groups, by expressing the $\text{Error}_{\text{BTC}}$ following failed trials relative to the $\text{Error}_{\text{BTC}}$ following successful trials. Doing so, we found a significant effect of the Group_{TYPE} during Day 1 training ($F_{(2,84)} = 10.27$, $p < 0.001$, partial $\eta^2 = 0.20$; **Figure 3A**). As expected, participants of the Group-SR adjusted their force profile depending on the reinforcement feedback, while participants of the Group-S were unable to do so (post-hocs; Group-S vs. Group-SR: $p = 0.022$, $d = -0.68$). Interestingly,

this ability to adjust motor commands based on the reinforcement was amplified by reward (Group-SR vs. Group-SRR: $p = 0.036$, $d = -0.57$). This result suggests that one mechanism through which reward improves motor learning is the potentiation of reinforcement-related adjustments in motor commands. To further test this idea, we evaluated the relationship between the magnitude of reinforcement-based changes in motor commands and the average success rate in the following trial across all subjects. Consistently, we found that the magnitude of reinforcement-related adjustments was strongly associated to the probability of success ($R^2 = 0.62$; $p = 1.5 \times 10^{-19}$; **Figure 3B**): the more participants adjusted their behavior based on the reinforcement feedback in a given trial (e.g., by reducing Error_{BTC} following Success and/or by increasing it following a Failure), the more they were likely to be successful in the following trial, supporting the view that these adjustments were relevant in the present task. Hence, these data suggest that the effect of reward on motor skill learning relies on the ability to adjust movements based on the reinforcement feedback.

In a second step, we asked whether such single-trial effects were maintained on Day 2, while all participants performed the task with sensory and reinforcement feedback, but in the absence of reward (*i.e.*, in a Block-SR setting). Interestingly, there was also an Outcome \times Group_{TYPE} interaction: $F_{(2,78)} = 3.75$, $p = 0.027$, partial $\eta^2 = 0.088$) demonstrating differences in the way participants relied on the reinforcement feedback on Day 2 based on the type of training experienced on Day 1 (**Figure 3C**). All groups displayed a larger Error_{BTC} following a failed compared to a successful trial (Group-S: $p < 0.018$, $d = 0.46$; Group-SR: $p < 0.001$, $d = 0.75$; Group-SRR: $p < 0.001$, $d = 1.13$). Notably though, post-hoc tests did not identify any group difference in post-Success Error_{BTC} (Group-SR vs. Group-SRR: $p = 0.13$, $d = 0.72$; Group-SR vs. Group-S: $p = 0.96$, $d = -0.019$; Group-SRR vs. Group-S: $p = 0.12$, $d = -0.56$) nor did it in post-Failure Error_{BTC} (Group-SR vs. Group-SRR: $p = 0.33$, $d = -0.20$; Group-SR vs. Group-S: $p = 0.35$, $d = 0.25$; Group-SRR vs. Group-S: $p = 0.058$, $d = 0.39$). Yet, when expressing Error_{BTC} in Post-Failure relative to Post-Success trials, we found that participants who had previously received reward on Day 1 in Group-SRR adjusted more their movements according to the reinforcement feedback compared to Group-S and Group-SR ($F_{(2,78)} = 3.53$, $p = 0.034$, partial $\eta^2 = 0.083$; post-hocs; Group-S vs. Group-SRR: $p = 0.017$, $d = -0.66$, Group-SR vs. Group-SRR: $p = 0.039$, $d = -0.56$, **Figure 3C**). There was no difference between Group-S and Group-SR ($p = 0.72$, $d = -0.10$). Here again, the magnitude of reinforcement-based adjustments correlated with the success in the next trial ($R^2 = 0.51$; $p = 5.5 \times 10^{-14}$; **Figure**

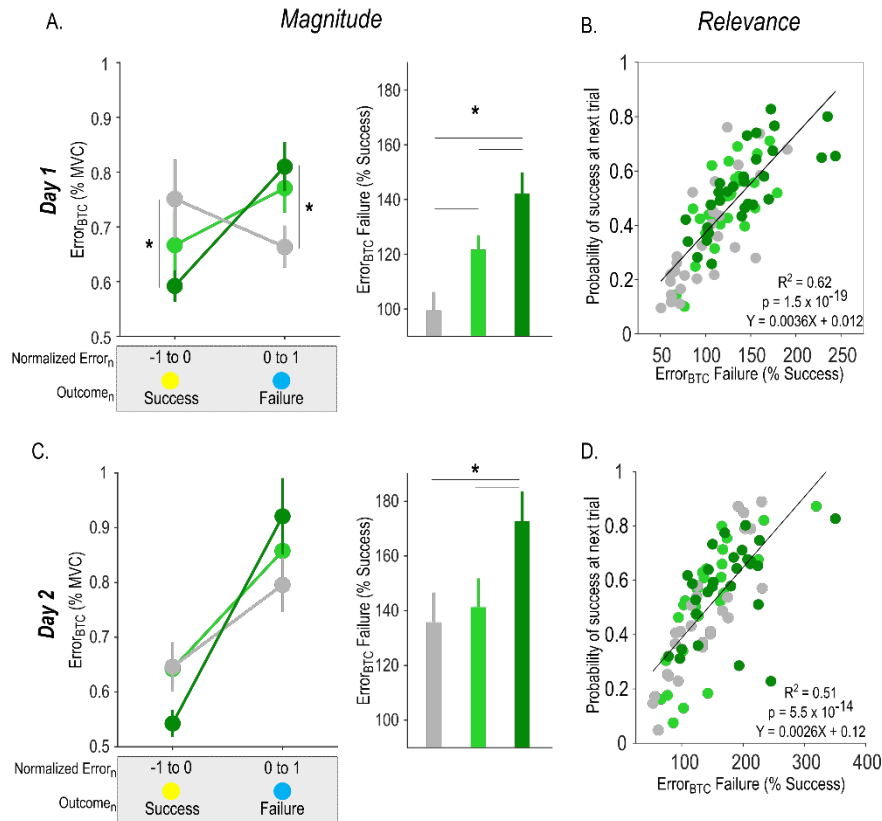
3D). Hence, the effect of reward on reinforcement-based adjustments can persist on a subsequent session of training, even after reward removal.

As explained above, we evaluated reinforcement-based adjustments by comparing $\text{Error}_{\text{BTC}}$ following successful or failed trials. However, by definition, successful and failed trials did not only differ with respect to the reinforcement feedback obtained at trial_n, but also regarding the experienced sensory feedback. Hence, the reward effect reported above could be specific to reinforcement-based adjustments, or may reflect a different reliance on the sensory feedback (or a combination of both). To disentangle these possibilities, we reasoned that the extent to which participants relied on the somatosensory feedback to adjust their movements could be estimated by computing $\text{Error}_{\text{BTC}}$ following failed trials of different Error magnitudes (*i.e.*, small or large Failure). In other words, we contrasted $\text{Error}_{\text{BTC}}$ following trials with the same reinforcement feedback (*i.e.*, Failure) but with different somatosensory experiences (*i.e.*, resulting from Small or Large Failures). Here too, we found a significant Outcome x Group_{TYPE} interaction on the $\text{Error}_{\text{BTC}}$ ($F_{(2,76)} = 5.15$, $p = 0.0080$, partial $\eta^2 = 0.12$; **Figure 3E**). As such, adjustments were greater after Large than after Small Failures in Group-SR and Group-SRR ($p < 0.001$, $d = 0.65$ and $p < 0.001$, $d = 1.20$, respectively), but not in Group-S ($p = 0.50$, $d = 0.19$). Post-hoc tests also indicated that adjustments after a Large Failure were greater in Group-SR and Group-SRR than in Group-S ($p < 0.001$, $d = 1.15$ and $p < 0.001$, $d = 1.31$, respectively), but not different between Group-SR and Group-SRR ($p = 0.26$, $p < 0.001$, $d = 0.25$). After Small Failures, $\text{Error}_{\text{BTC}}$ was also larger in Group-SR than in Group-S ($p = 0.042$, $d = 0.56$), but not different between Group-SR and Group-SRR ($p = 0.31$, $d = 0.30$) and between Group-SRR and Group-S ($p = 0.34$, $d = 0.44$). This indicates that while subjects of the Group-SR and Group-SRR adjusted the Error depending on the sensory feedback, participants of the Group-S were not able to do so, suggesting that training with reinforcement feedback allowed participants to be more sensitive to the sensory feedback (Galea et al., 2015; Bernardi et al., 2015), regardless of whether they received reward or not. Consistently, we found a Group_{TYPE} effect ($F_{(2,76)} = 5.05$, $p = 0.0087$, partial $\eta^2 = 0.12$; **Figure 3E**) on the magnitude of sensory-based adjustments ($\text{Error}_{\text{BTC}}$ following Large Failures expressed relative to $\text{Error}_{\text{BTC}}$ following Small Failures), which was driven by differences between Group-S and the two other groups (post-hocs; Group-S vs. Group-SR: $p = 0.0056$, $d = -0.73$, Group-S vs. Group-SRR: $p = 0.011$, $d = -0.85$). Importantly, we did not find any difference

between Group-SR and Group-SRR ($p = 0.90$, $d = -0.033$). Then, similarly as for the reinforcement-based changes, we found that the magnitude of sensory-based adjustments correlated with the subsequent probability of success ($R^2 = 0.34$, $p = 1.8 \times 10^{-8}$; **Figure 3F**), demonstrating that these adjustments were also relevant in the learning process.

On Day 2, the effect of Outcome persisted ($F_{(1,68)} = 15.20$, $p < 0.001$, partial $\eta^2 = 0.18$) with a trend for a Group_{TYPE} effect ($F_{(2,68)} = 3.12$, $p = 0.051$, partial $\eta^2 = 0.084$) but no Outcome \times Group_{TYPE} interaction ($F_{(2,68)} = 3.12$, $p = 0.46$, partial $\eta^2 = 0.013$,). Consistently, the magnitude of sensory-based adjustments was not different between the Group_{TYPES} ($F_{(2,68)} = 0.41$, $p = 0.67$, partial $\eta^2 = 0.012$, **Figure 3G**). Note though that similarly to Day 1, sensory-based adjustments significantly correlated with the probability of success on Day 2 ($R^2 = 0.13$, $p = 0.0022$; **Figure 3H**). Hence, the absence of reward effects on sensory-based adjustments on Day 1 and 2 cannot be explained by the fact that participants did not rely on this type of feedback.

Reinforcement-based adjustments in motor commands



Sensory-based adjustments in motor commands

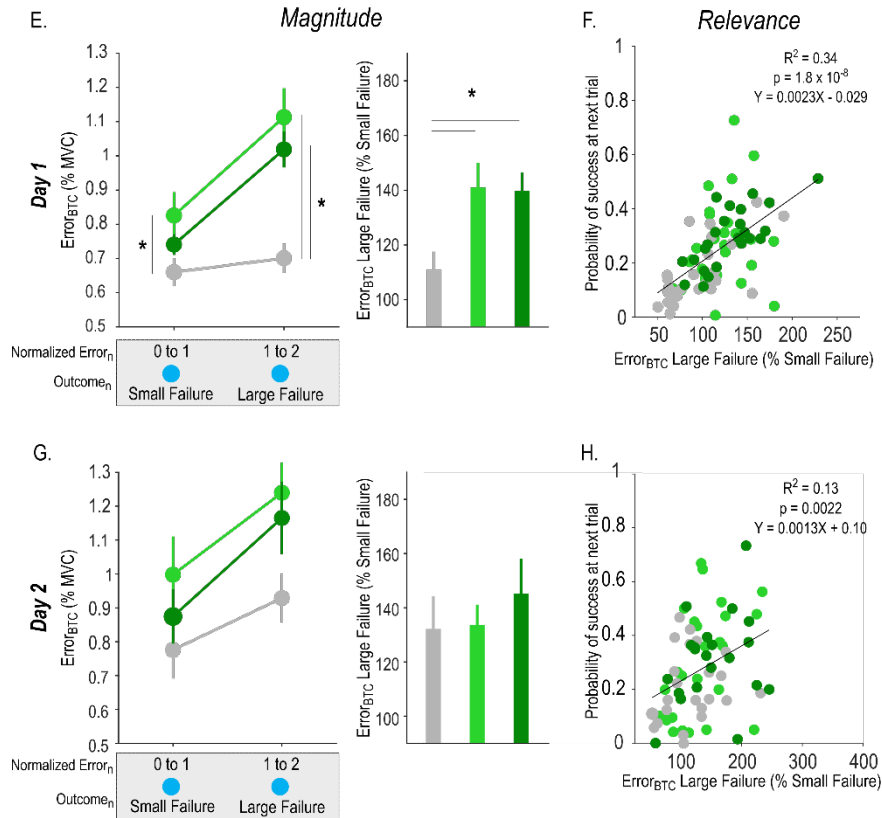


Figure 3. Between-trial adjustments in the Error. (A) Reinforcement-based adjustments in the Error during Day 1 training. Absolute between-trial adjustments in the Error ($\text{Error}_{\text{BTC}} = |\text{Error}_{n+1} - \text{Error}_n|$) according to the reinforcement feedback (*i.e.*, Success or Failure) encountered at trial_n in the three Group_{TYPE}S (grey: Group_{-S}, light green: Group_{-SR}, dark green: Group_{-SRR}). Notably, these bins of trials were constituted based on the success threshold-normalized Error at trial_n in order to compare adjustments in motor commands following trials of similar Error in the three groups. Stars denote significant group differences in $\text{Error}_{\text{BTC}}$ for a given outcome (left panel, see Materials and Methods). Reinforcement-based adjustments ($\text{Error}_{\text{BTC}}$ after Failure in percentage of $\text{Error}_{\text{BTC}}$ after Success) were compared in the three Group_{TYPE}S (right panel). **(B)** Correlations between the magnitude of reinforcement-based adjustments in the Error and the average success rate on the next trial, showing the relevance of these adjustments in the present task. Each dot represents a subject. **(C, D)** Same for Day 2 training. Note that reinforcement-based adjustments in motor commands remained amplified in Group_{SRR}, despite the absence of reward on Day 2. **(E)** Sensory-based adjustments in the Error during Day 1 training. $\text{Error}_{\text{BTC}}$ following trials_n with Failures of different Error magnitudes (left panel). Sensory-based adjustments ($\text{Error}_{\text{BTC}}$ after Large Failure in percentage of $\text{Error}_{\text{BTC}}$ after Small Failure) were compared in the three Group_{TYPE}S (right panel). **(F)** Correlations between the magnitude of sensory-based adjustments in the Error and the probability of success on the next trial, showing the relevance of these adjustments for task success. **(G, H)** Same for Day 2 training. *: $p < 0.05$.

The single-trial analyses on $\text{Error}_{\text{BTC}}$ revealed significant differences in the way participants of each group adjusted their motor commands based on the reinforcement and the sensory feedback. However, the distribution of the Error data could have contributed to these single-trial effects. Indeed, even for random adjustments in motor commands (*e.g.*, based on a Gaussian process), adjustments following small or large Error_n (*i.e.*, in the tails of the Error distribution) would be larger than adjustments following Error_n close to the mean of the distribution. Hence, to ensure that group differences in Error distribution did not contribute to our single-trial results, we ran a control analysis in which we shuffled the Error data for each subject (with 10000 permutations), and then re-computed reinforcement and sensory-based adjustments exactly as in the main analysis. Importantly, we did not find any Group_{TYPE} effect on these shuffled data neither for reinforcement- (Day 1: $F_{(2,84)} = 1.6$, $p = 0.21$, partial $\eta^2 = 0.04$; Day 2: $F_{(2,78)} = 0.89$, $p = 0.41$, partial $\eta^2 = 0.02$) nor for sensory-based adjustments (Day 1: $F_{(2,76)} = 0.02$, $p = 0.98$, partial $\eta^2 = 0.0006$; Day 2: $F_{(2,78)} = 0.20$, $p = 0.82$, partial $\eta^2 = 0.006$). This analysis indicates that the differences in single-trial adjustments reported here were not related to a sampling bias.

Reward boosts reinforcement-based adjustments at a specific level of motor control

As a last step, we asked whether the effect of reward on between-trial adjustments in motor commands concerned all aspects of force control, or only some specific motor components. To do so, we investigated how reinforcement and sensory feedback shaped adjustments in the speed and accuracy of force production in the three Group_{TYPE}s by dissecting each force profile into three separate components (**Figure 1B**). To evaluate the speed at which the force was generated, we computed the time required for force initiation (*i.e.*, the time required to reach half of the Target_{Force}: Force_{Initiation}). To assess the accuracy of the force, we computed the force difference between the average amplitude of the generated force and the Target_{Force} (Force_{AmplError}), and the variability (standard deviation/mean) of the maintained force (Force_{Variability}). Notably, both indicators of force accuracy were computed in the second half of the trial (*i.e.*, the last 1000 ms), well after force initiation, when participants maintained a stable level of force. Importantly, in a first analysis we found that the presence of reward during training allowed improvements at all levels of force control (**Figure S1**), in line with previous studies (Manohar et al., 2015; Codol et al., 2020).

We compared between-trial changes in Force_{Initiation} (Force_{Initiation-BTC}), Force_{AmplError} (Force_{AmplError-BTC}), and Force_{Variability} (Force_{Variability-BTC}) following Success or Failure trials of similar Error magnitude in the three groups. The ANOVA run on the Force_{Initiation-BTC} data revealed a significant Outcome x Group_{TYPE} interaction ($F_{(2,84)} = 7.62$, $p < 0.001$, partial $\eta^2 = 0.15$) that was driven by the fact that post-Success and post-Failure Force_{Initiation-BTC} were different in Group-SR and Group-SRR ($p < 0.001$, $d = -1.04$ and $p < 0.001$, $d = -1.59$, respectively) but not in Group-S ($p = 0.10$, $d = -0.27$). Moreover, post-Success changes in Force_{Initiation} were smaller in Group-SRR than in Group-S ($p = 0.023$, $d = -0.56$); it also tended to be smaller in Group-SR than in Group-S ($p = 0.071$, $d = -0.49$), while it was comparable in Group-SR and Group-SRR ($p = 0.65$, $d = 0.14$). Corroborating these results, we found that reinforcement feedback impacted the modulation of initiation speed (expressed as Force_{Initiation-BTC} following a Failure in percentage of Force_{Initiation-BTC} following a Success; $F_{(2,84)} = 8.50$, $p < 0.001$, partial $\eta^2 = 0.17$; post-hocs: Group-S vs. Group-SR: $p = 0.0011$, $d = -0.84$, Group-S vs. Group-SRR: $p < 0.001$, $d = -1.09$; **Figure 4A**). Interestingly though, we did not find any effect of reward on the reinforcement-based adjustment of speed (Group-SR vs. Group-SRR: $p = 0.78$, $d = -0.072$). At the level of Force_{AmplError}, we found again a Outcome x Group_{TYPE} interaction ($F_{(2,84)} = 14.07$, $p < 0.001$, partial $\eta^2 = 0.25$; **Figure 4B**) that was driven by the fact that post-Success and post-Failure Force_{AmplError-BTC} were different in Group-SR and Group-

SRR ($p < 0.001$, $d = -0.97$ and $p = 0.0034$, $d = -1.44$, respectively) but not in Group-S ($p = 0.99$, $d = -0.0002$). Group comparisons at post-Success and post-Failure did not evidence any significant difference in Force_{AmplError-BTC}. Notably though, there was a trend for the post-Success Force_{AmplError-BTC} to be smaller in Group-SRR than in Group-S ($p = 0.066$, $d = -0.42$). Interestingly, direct comparison of reinforcement-related changes in Force_{AmplError-BTC} (post-Failure vs. post-Success) revealed a significant effect of reward ($F_{(2,84)} = 9.54$, $p < 0.001$, partial $\eta^2 = 0.19$; **Figure 4B**). As such, participants of the Group-SRR modulated more the Force_{AmplError} according to the reinforcement feedback than subjects of the two other groups (Group-S vs. Group-SRR: $p < 0.001$, $d = -1.04$, Group-SR vs. Group-SRR: $p = 0.018$, $d = -0.70$). Notably, there was also a trend for Group-SR to be different from Group-S ($p = 0.064$, $d = -0.50$). Finally, analysis of Force_{Var-BTC} did not reveal any Outcome x Group_{TYPE} interaction ($F_{(2,84)} = 0.79$, $p = 0.46$, partial $\eta^2 = 0.018$; **Figure 4C**), neither did it show a Group_{TYPE} effect ($F_{(2,84)} = 0.81$, $p = 0.45$, partial $\eta^2 = 0.020$; **Figure 4C**) on reinforcement-based adjustments (Force_{Var-BTC} post-Failure vs. post-Success). Hence, while reward strongly influenced reinforcement-based adjustments of force amplitude, it did not modulate the between-trial regulation of the speed at which the force was initiated or the variability of the maintained force. This suggests that the effect of reward on reinforcement-related adjustments was not global (*i.e.*, affecting all aspects of the movement) but rather specific to force amplitude.

We also considered the effect of the sensory feedback on between-trial adjustments by comparing Force_{Initiation-BTC}, Force_{AmplError-BTC}, and Force_{Variability-BTC} following failed trials of different Error magnitudes (*i.e.*, small or large Failure). Contrary to the global Error_{BTC} index, we did not find any Outcome x Group_{TYPE} interaction neither for Force_{Initiation-BTC} ($F_{(2,76)} = 0.54$, $p = 0.59$, partial $\eta^2 = 0.014$), nor for Force_{AmplError-BTC} ($F_{(2,76)} = 2.80$, $p = 0.067$, partial $\eta^2 = 0.069$) or Force_{Variability-BTC} ($F_{(2,76)} = 1.25$, $p = 0.29$, partial $\eta^2 = 0.032$). Consistently, we did not find any significant difference in the way participants from the different groups adjusted individual motor components depending on the size of the preceding Failure (Large vs. Small Failure on **Figure 4D, 4E, 4F**; Force_{Initiation-BTC}: $F_{(2,76)} = 0.10$, $p = 0.90$, partial $\eta^2 = 0.0026$; Force_{AmplError-BTC}: $F_{(2,76)} = 2.57$, $p = 0.083$, partial $\eta^2 = 0.063$; Force_{Variability-BTC}: $F_{(2,76)} = 2.46$, $p = 0.092$, partial $\eta^2 = 0.061$). This analysis supports the idea that reward did not increase the sensitivity to the sensory feedback, but rather boosted specific adjustments in motor commands in response to the reinforcement feedback.

Finally, as a control analysis, we characterized the respective influence of each motor component in the Error, which determined task success. As such, in addition to representing different levels of force control (*i.e.*, initiation, amplitude and variability), the motor components evaluated here may also bear different relevance for task success (van der Kooij et al., 2021). For each participant, we ran separate partial linear regressions on the Error data with $\text{Force}_{\text{Initiation}}$, $\text{Force}_{\text{AmplError}}$ or $\text{Force}_{\text{Variability}}$ as predictors. Notably, we used partial regressions here to assess the relationship between the Error and each motor component, while controlling for the effect of the other motor components in the correlation. Interestingly, we found that $\text{Force}_{\text{AmplError}}$ explained the largest part of variance in the Error ($r = 0.96 \pm 0.003$; $p < 0.05$ in 90/90 subjects). $\text{Force}_{\text{Initiation}}$ also explained a large part of variance in the Error ($r = 0.81 \pm 0.01$; $p < 0.05$ in 90/90 subjects), while $\text{Force}_{\text{Variability}}$ explained a smaller, yet significant in most subjects, part of variance ($r = 0.22 \pm 0.03$; $p < 0.05$ in 68/90 subjects). Hence, although all motor parameters were relevant for task success, the $\text{Force}_{\text{AmplError}}$ was the most influential factor.

Altogether, our results demonstrate that reward potentiates reinforcement-based adjustments in motor commands and that this effect persists even after reward removal, on the subsequent day. The data also show that this effect does not concern all components of the movement but specifically the amplitude of the force which was the most relevant factor for task success.

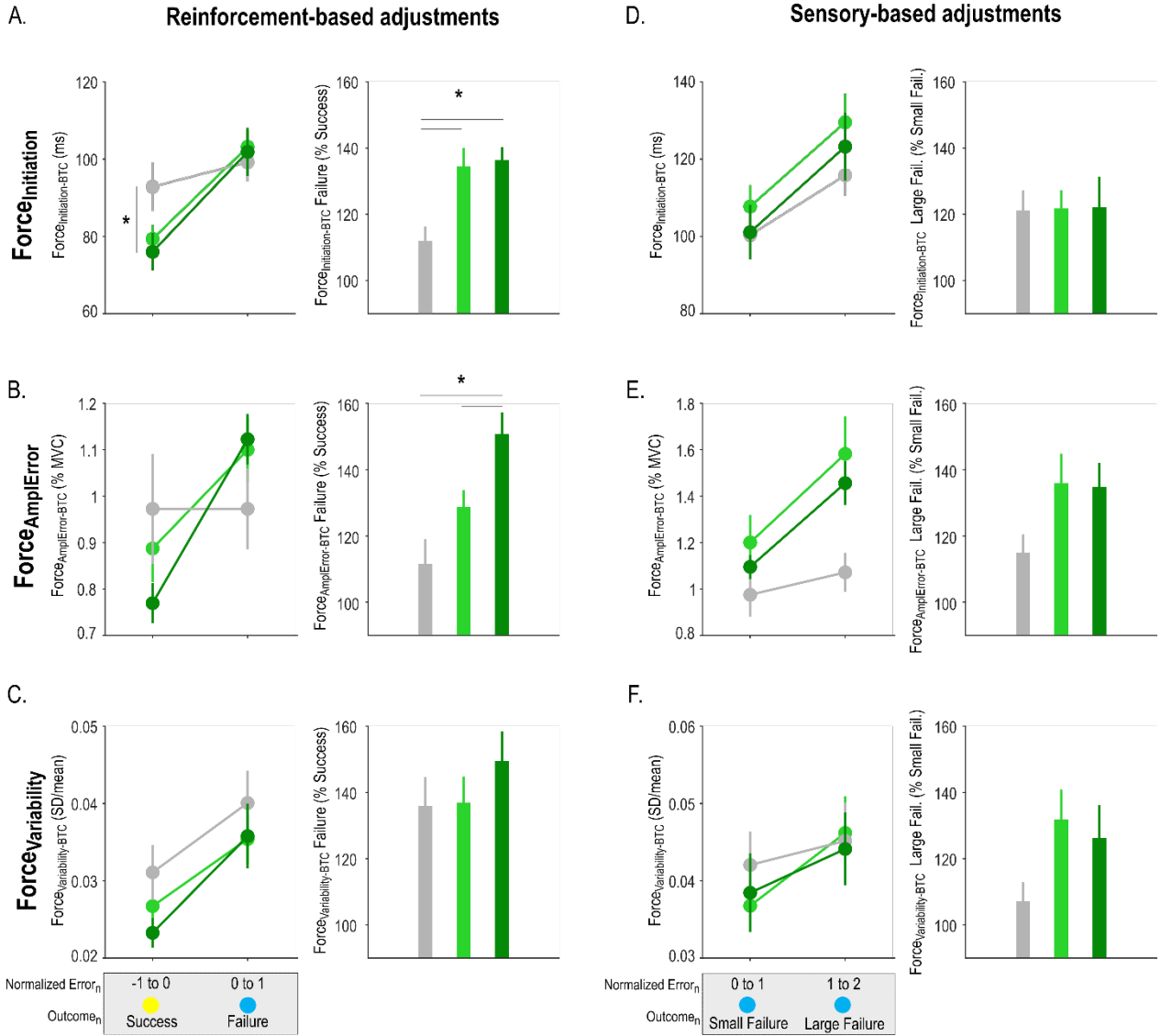


Figure 4. Between-trial adjustments in initiation time, amplitude error and variability. **Reinforcement-based adjustments in the $\text{Force}_{\text{Initiation}}$ (A), $\text{Force}_{\text{AmplError}}$ (B) and $\text{Force}_{\text{Variability}}$ (C).** Absolute between-trial changes (BTC) for each motor component ($\text{Force}_{\text{BTC}} = |\text{Force}_{n+1} - \text{Force}_n|$) according to the reinforcement feedback (*i.e.*, Success or Failure) encountered at trial_n in the three Group_{TYPES} (grey: Group_S, light green: Group_{SR}, dark green: Group_{SRR}). Notably, these bins of trials were constituted based on the success threshold-normalized Error at trial_n. Stars denote significant group differences in Error_{BTC} for a given outcome (left panel). **Reinforcement-based adjustments ($\text{Force}_{\text{BTC}}$ after Failure in percentage of $\text{Force}_{\text{BTC}}$ after Success) in the three Group_{TYPES} (right panel).** **Sensory-based adjustments in the $\text{Force}_{\text{Initiation}}$ (D), $\text{Force}_{\text{AmplError}}$ (E) and $\text{Force}_{\text{Variability}}$ (F).** $\text{Force}_{\text{BTC}}$ following trials_n with Failures of different Error magnitudes (left panel). **Sensory-based adjustments ($\text{Force}_{\text{BTC}}$ after Large Failure in percentage of $\text{Force}_{\text{BTC}}$ after Small Failure) in the three Group_{TYPES} (right panel).** *: $p < 0.05$.

2.4. Discussion

In this study, we investigated the net effect of reward on motor learning while controlling for the reinforcement feedback received by the participants. Our results provide evidence that reward can improve motor skill learning and that this effect is related to a specific potentiation of reinforcement-related adjustments in motor commands. Strikingly, the potentiation of such adjustments persisted on a subsequent day in the absence of reward. Moreover, such boosting of reinforcement-based adjustments did not concern all components of force production but only the amplitude, which was the most relevant one for task success. These findings shed light on the mechanisms through which reward can durably enhance motor performance. They also lay the groundwork for future rehabilitation strategies involving optimized sensory and reinforcement feedbacks.

A main goal of the present study was to explore the net effect of reward on motor skill learning by experimentally dissociating it from the reinforcement feedback. As such, previous motor learning studies have often coupled reinforcement and reward (*e.g.*, (Wachter et al., 2009; Abe et al., 2011; Wilkinson et al., 2015; Steel et al., 2016, 2019)), based on the underlying assumption that receiving knowledge of performance (*e.g.*, points or binary feedback) provides a form of intrinsic reward that can by itself increase motivation to perform well (Leow et al., 2018). However, in addition to providing some form of intrinsic reward, reinforcement feedback also provides a learning signal to the motor system, that can influence motor learning (Huang et al., 2011; Shmuelof et al., 2012; Bernardi et al., 2015; Galea et al., 2015b; Nikooyan et al., 2015; Therrien et al., 2016; Mawase et al., 2017; Leow et al., 2018; Uehara et al., 2018; Kim et al., 2019). In order to assess the net effect of motivation on motor learning, we therefore compared groups of participants trained with different monetary rewards but with the exact same reinforcement feedback. We found that motivation by reward allowed marked improvements in motor performance that were maintained after reward removal and even 24 hours later (**Figure 2**). Notably, this was the case despite the fact that reinforcement alone was not sufficient to influence motor learning in our task. This demonstrates that the motivational context experienced during training can by itself strongly influence motor skill learning, beyond performance-based reinforcement feedback.

The prospect of obtaining rewards for good performance enhances motivation but does not provide any additional learning signal to the motor system (Berke, 2018). Yet, it may boost the reliance on sensory and/or reinforcement feedbacks (Kim et al., 2019). To

explore this possibility, we developed an analysis allowing us to investigate how participants adjusted their motor commands based on sensory or reinforcement feedbacks while controlling for differences in performance between the groups (see Materials and Methods for more details). Interestingly, we found that reward specifically boosted reinforcement-based adjustments, following both positive and negative feedbacks, while sensory-based adjustments remained unaffected by reward (**Figure 3**). This suggests that reward boosted both the reproduction of successful behaviour (exploitation) and correction of motor commands after failure (exploration; (Dhawale et al., 2017)). This was the case despite the fact that both types of feedback were relevant to improve motor performance at the single-trial level (**Figure 3B, D, F, H**). This result suggests that reward increases the reliance on reinforcement information during the learning process, with less effect on sensory-based adjustments. Interestingly, this finding may explain why tasks that strongly emphasize sensory-based learning (over reinforcement-based learning; (Izawa and Shadmehr, 2011; Cashaback et al., 2017)), often show less sensitivity to motivation. Accordingly, monetary reward shows little impact on sensorimotor adaptation (Galea et al., 2015b; Hill et al., 2020) and on motor skill acquisition in tasks that strongly rely on sensory feedback (e.g., (Abe et al., 2011; Steel et al., 2016; Widmer et al., 2016)). The differential effect of reward on sensory and reinforcement-based adjustments may be due to the qualitatively different learning processes that are driven by these two types of feedbacks (Cashaback et al., 2017; Uehara et al., 2018). As such, while sensory feedback promotes error correction by providing directional feedback (Shadmehr et al., 2010b), reinforcement can guide motor exploration based on binary feedback about task success (Therrien et al., 2016). Our results, along with the observation that monetary rewards are less effective in tasks where learning is dominated by sensory feedback, suggest that the potential of reward to improve motor learning relies on the boosting of a reinforcement learning mechanism. Based on this, we propose that the susceptibility of a given motor learning task to reward may depend on the relative contribution of sensory and reinforcement feedbacks in the learning process. Characterizing what type of motor tasks can benefit from motivational interventions is an important line of future work to translate fundamental motor control research into innovative rehabilitation procedures.

The finding of a reward-dependent boosting of reinforcement-based adjustments is in line with previous neuroimaging results showing that reward increases reinforcement-related activity in the striatum in the context of motor learning (Widmer et al., 2016). This reward-driven increase in striatal activity is reduced after a stroke (even when the striatum is unlesioned), a process that may contribute to the motor learning deficits observed in these patients (Widmer et al., 2019). Moreover, such reward-dependent modulation of motor adjustments has been shown to rely on dopamine (Galea et al., 2013; Pekny et al., 2015), a key neurotransmitter of the striatal circuitry. Based on these elements and on the causal role of the striatum in reinforcement-based adjustments in motor commands (Nakamura and Hikosaka, 2006; Williams and Eskandar, 2006), we suspect that this region may be crucial for the beneficial effect of reward observed in the present study. Notably, the cerebellum (Wagner et al., 2017; Heffley et al., 2018; Carta et al., 2019; Vassiliadis et al., 2019; Sendhilnathan et al., 2020) and frontal areas (Dayan et al., 2014b, 2018; Sidarta et al., 2016; Ramakrishnan et al., 2017; Hamel et al., 2018; Palidis et al., 2019) are also likely to contribute to reward-based motor learning. Further investigations are required to better delineate the neurophysiological bases of reward-related improvements in motor learning.

The beneficial effect of reward on single-trial adjustments was maintained on day 2, even after reward removal. As in day 1 training, reinforcement-based adjustments were boosted while sensory-based adjustments remained unchanged by reward. This persistent change in the specific reaction to the reinforcement feedback after reward removal is suggestive of an associative learning process. In associative learning, presentation of a neutral stimulus (*i.e.*, a conditioned stimulus) that has been consistently paired with a rewarding stimulus (*i.e.*, an unconditioned stimulus) during a training period elicits a behaviour that was initially only generated in reaction to the reward (Pavlov, 1927; Rescorla and Wagner, 1972). Following this framework, it is possible that the repetitive pairing of the reinforcement feedback with the reward during training induced an implicit association between the two events that remained evident when the reward was removed. This could explain why strong reinforcement-specific adjustments were maintained on day 2 in the reward group, even though no rewards were at stake anymore. Such associative learning processes are known to strongly influence autonomic responses (Pool et al., 2019), inhibitory control (Verbruggen et al., 2014; Lindström et al., 2019; Avraham et al., 2020), decision making (Lindström et al., 2019) and even sensorimotor adaptation (Avraham et al., 2020) in humans. We propose

that associative learning may also contribute to the durable influence of motivation on motor skill learning (Abe et al., 2011; Sporn et al., 2020).

In order to better characterize the effect of reward on motor learning, we considered separately the different components of the movement and found that force amplitude was the most strongly affected, while the speed of initiation and force variability remained largely insensitive to reward. This suggests that reward can have a selective influence on the regulation of a specific component of motor control. Importantly, an estimation of the respective influence of each motor component on task success also showed that force amplitude was the most relevant component for the task. Notably, the specificity of the effect of reward on the regulation of one motor component is in accordance with the idea that multidimensional motor tasks (*i.e.*, requiring the control of multiple motor components) can be decomposed in subtasks that are learned separately in the motor system (Ghahramani and Wolpert, 1997). In this framework, learning of the different motor components may depend on their respective relevance for task success (Ghahramani and Wolpert, 1997; van der Kooij et al., 2021). Such task relevance may be estimated based on a priori knowledge of the task (*e.g.*, following instructions; (Popp et al., 2020)) and through the reliance on a credit assignment system allowing to estimate the particular influence of each motor component on task success through trial and error (McDougle et al., 2016; Parvin et al., 2018). Based on this idea, we believe that the strong relationship between the amplitude of the force and task success in the present task pushed participants of the reward group to largely modulate this component based on the reinforcement feedback. If this is the case, this would suggest that it is possible to affect the training of specific motor abilities by modulating the weight of individual motor components in the computation of the reinforcement feedback, an aspect that could be exploited in future rehabilitation protocols. Alternatively, reward might have specifically modulated the amplitude of the force, independently of the relevance of this parameter. Although the present study cannot rule out this hypothesis, we believe that such interpretation is unlikely given previous demonstration that reward can improve several aspects of motor control concomitantly (Manohar et al., 2015; Codol et al., 2020). Another possibility is that reinforcement feedback alone was sufficient to maximally modulate initiation time and variability in this task, precluding us from observing a difference with the reward-based training because of some form of ceiling effect. Further studies are required to disentangle these

potentially co-existing interpretations to guide the development of component-specific rehabilitation therapies (Norman et al., 2017).

Limitations of the study

Our findings suggest that extrinsic reward can improve the acquisition and maintenance of a motor skill by boosting reinforcement-based adjustments in motor commands. However, it should be noted that here we focused on a very simple unimanual task in which performance relied on the ability to modulate a 1-degree of freedom force. While our analysis of the different motor components suggests that reward may also improve the learning of more complex tasks (by selectively boosting the adjustment of the most relevant dimensions for task success), future studies should address the generalizability of our results by using tasks engaging more complex skills.

Besides, our single-trial analysis suggests that reward affects differently sensory and reinforcement-based adjustments in motor commands. Yet, sensory and reinforcement feedbacks were always coupled in the present task. We did so on purpose, to avoid inducing conflict in the learning process (e.g., resulting from rewarding “bad performance”). Notably though, the reward effect we report here could be influenced by the relationship between these feedbacks. Hence, follow-up investigations should assess the effect of reward on sensory and reinforcement-based adjustments in situations where both feedback types are dissociated (Cashaback et al., 2017).

2.5. Materials and Methods

2.5.1. Resources table

RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Motor learning data	This paper	https://osf.io/5pjem/
Subjects characteristics	This paper	https://osf.io/5pjem/
Software and algorithms		
Matlab vR2007 7.5 and R2008a	Mathworks	MATLAB - MathWorks - MATLAB & Simulink
Statistica 10	StatSoft Inc.	https://www.statsoft.de/en/software/tibco-statisticatm
Psychophysics Toolbox	Psychtoolbox.org	http://psychtoolbox.org/

2.5.2. Participants

A total of ninety right-handed healthy volunteers participated in the present study (58 women, 23.7 ± 0.3 years old; mean \pm SE). Handedness was determined via a shortened version of the Edinburgh Handedness inventory Oldfield (Oldfield, 1971). None of the participants suffered from any neurological or psychiatric disorder, nor were they taking any centrally-acting medication. All participants gave their written informed consent in accordance with the Ethics Committee of the Université Catholique de Louvain (approval number: 2018/22MAI/219) and the principles of the Declaration of Helsinki. Subjects were financially compensated for their participation. Finally, all participants were asked to fill out a French adaptation of the Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ; (Torrubia et al., 2001; Lardi et al., 2008)).

2.5.3. Methods details

2.5.3.1. General aspects

Participants were seated approximately 60 cm in front of a computer screen (refresh rate = 100 Hz) with their right forearm positioned at a right angle of the table. The task was developed on Matlab 7.5 (the Mathworks, Natick, Massachusetts, USA) exploiting the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) and consisted in an adaptation of previously used motor learning tasks (Abe et al., 2011; Steel et al., 2016; Mawase et al.,

2017). The task required participants to squeeze a force transducer (Arsalis, Belgium) between the index and the thumb to control the one-dimension motion of a cursor displayed on the screen. Increasing force resulted in the cursor moving vertically and upward. Each trial started with a preparatory phase in which a sidebar appeared at the bottom of the screen and a target at the top (**Figure 1A**). After a variable time interval, a cursor popped up in the sidebar and participants had to pinch the transducer to move the cursor as quickly as possible from the sidebar to the target and maintain it there for the rest of the task. The level of force required to reach the target ($\text{Target}_{\text{Force}}$) was individualized for each participant and set at 10% of maximum voluntary contraction (MVC). Notably, squeezing the transducer before the appearance of the cursor was considered as an anticipation and therefore led to an interruption of the trial. Such trials were discarded from further analyses. At the end of each trial, a binary reinforcement feedback represented by a colored circle was provided to the subject followed by a reminder of the color/feedback association and potential monetary reward associated to good performance (see Reinforcement feedback section below).

2.5.3.2. Sensory feedback

We provided only limited visual feedback to the participants (Mawase et al., 2017). As such, on most trials (90%), the cursor disappeared shortly after the subject started to squeeze the force transducer (partial vision trials): it became invisible as soon as the generated force became larger than half of the $\text{Target}_{\text{Force}}$ (i.e., 5% of MVC). Conversely, the remaining trials (10%) provided a continuous vision of the cursor (full vision trials). Therefore, on most trials, participants had limited visual information and had to rely exclusively on somatosensory feedback to generate the $\text{Target}_{\text{Force}}$. Importantly, full vision trials were not considered in the analyses.

2.5.3.3. Reinforcement feedback

At the end of each trial, subjects were presented with a binary reinforcement feedback indicating performance. Success on the task was determined based on the Error; that is, the absolute force difference between the $\text{Target}_{\text{Force}}$ and the exerted force (**Figure 1B**; (Abe et al., 2011; Steel et al., 2016)). The Error was computed for each frame refresh (i.e., at 100Hz) from 150 ms to the end of the trial and then averaged for each trial (Steel et al., 2016) and

expressed in percentage of MVC. This indicator of performance allowed us to classify a trial as successful or not based on an individualized success threshold (see below). When the Error on a given trial was below the threshold (negative normalized Error), the trial was considered as successful, and when it was above the threshold (positive normalized Error), the trial was considered as failed. Hence, task success depended on the ability to reduce the Error by approximating the $\text{Target}_{\text{Force}}$ as quickly and accurately as possible. Importantly, participants were told explicitly that both speed and accuracy were taken into account to determine task success. In summary, to be successful, participants knew that they had to quickly initiate the force and be as accurate as possible in reproducing the $\text{Target}_{\text{Force}}$.

In different blocks of trials, we manipulated the reinforcement feedback and reward provided during training. In Block-S, the reinforcement feedback was non-informative (magenta circle regardless of performance), and participants could only rely on somatosensory feedback to perform the task. In Block-SR, the reinforcement feedback consisted in a yellow (representing a successful trial) or blue circle (representing a failed trial), providing knowledge of performance (**Figure 1A**). In Block-SRR, this knowledge of performance was associated to a monetary reward (+8 cents or 0 cent for Success or Failure, respectively). Therefore, contrarily to Block-S, Block-SR and Block-SRR provided knowledge of performance and this feedback was associated to a monetary reward in Block-SRR.

2.5.3.4. Experimental procedure

Subjects' performance was tested for two consecutive days (Day 1 and Day 2; **Figure 1C**). On Day 1, we first measured the individual MVC to calculate the $\text{Target}_{\text{Force}}$. Notably, MVC was measured before and after both sessions to assess potential muscle fatigue related to the training (see 4.4.3). Participants then performed 2 blocks of Familiarization. In a first block, participants performed 20 full vision trials; it served to familiarize the subjects with the task in a Block-SR setting (Full vision block). Subsequently, all blocks were composed of a mixture of full vision trials (10 % of total trials) and partial vision trials (90 % of total trials). The second Familiarization block consisted in 20 trials and allowed us to determine baseline performance to individualize the difficulty of the task for the rest of the experiment (Calibration block). For every subject, each partial vision trial of the Calibration block was classified in terms of Error from the lowest to the greatest in percentage of MVC. We took the 35th percentile of the Error to determine the individual success threshold. Success thresholds

were constrained between 2 and 3.5 % of MVC by asking participants to repeat the Calibration block when the computed threshold was outside these boundaries. Those parameters were determined based on pilot data to obtain coherent learning curves among individuals.

After the Familiarization and Calibration blocks, the first experimental session consisted in 280 trials divided in 8 blocks. All subjects started with a Block-SR of 20 trials to evaluate the performance at Pre-training and similarly ended the session with a Post-training assessment of 20 trials. In between, 6 Training blocks of 40 trials were performed by the participants (**Figure 1B**). During this Training period, individuals were split into 3 separate groups (Group_{TYPE}: Group-S, Group-SR or Group-SRR) depending on the type of blocks they performed during training. As such, Group-S completed Block-S, Group-SR performed Block-SR and Group-SRR trained under Block-SRR condition. Contrasting performance in the Pre- and Post-training blocks allowed us to evaluate learning of the skill under the three training conditions. 24 hours later, subjects performed the task again with the same Target_{Force} and success threshold. After a 20 trials Familiarization used to remind the task to participants, they performed 140 trials split in 4 blocks; all were performed in a Block-SR setting. This Re-test session allowed us to assess skill maintenance 24h after training.

2.5.4. Data and statistical analyses

Statistical analyses were carried out with Matlab 2018a (the Mathworks, Natick, Massachusetts, USA) and Statistica 10 (StatSoft Inc., Tulsa, Oklahoma, USA). Post-hoc comparisons were always conducted using the Fisher's LSD procedure. The significance level was set at $p \leq 0.05$, except in the case of correction for multiple comparisons (see below).

2.5.4.1. Motor skill learning and maintenance

The main aim of the present study was to evaluate the effect of reward on motor skill learning and maintenance. To assess skill learning, we expressed the median Error at Post-training in percentage of the value obtained at Pre-training. To evaluate skill maintenance, we expressed the median Error during the Re-test session in percentage of Pre-training. First, compared skill learning and maintenance between the groups through one-way ANOVAs with the factor Group_{TYPE}. Then, we also explored the significance of skill learning and maintenance within each group by conducting Bonferroni-corrected single sample t-tests on

these percentage data against a constant value of 100% (*i.e.*, corresponding to the Pre-training level).

As explained above, task performance depended on both the speed and the accuracy of the produced force (**Figure 1B**). We characterized the effect of reward on these different levels of force control, by evaluating separately the speed of force initiation and the accuracy of the maintained force. To evaluate the speed of force initiation, we measured the force initiation time ($\text{Force}_{\text{Initiation}}$) which was defined as the delay between the appearance of the cursor and the moment where the applied force reached 5% of MVC (*i.e.*, corresponding to half of the $\text{Target}_{\text{Force}}$). Force accuracy was evaluated in the second half of the trial (*i.e.*, the last 1000 ms), through two different parameters. First, we computed the Amplitude Error of the force ($\text{Force}_{\text{AmplError}}$), defined as the absolute difference between the mean force exerted in the last 1000 ms of the trial and the $\text{Target}_{\text{Force}}$. It reflected how much the amplitude of the maintained force differed from the $\text{Target}_{\text{Force}}$. Second, force accuracy was also characterized by considering the variability of the maintained force, with high levels of variability causing increases in the Error. To assess force variability ($\text{Force}_{\text{Variability}}$), we computed the coefficient of variation of the force in the second half of the trial (*i.e.*, standard deviation of force/mean force). In summary, to be successful, participants had to quickly initiate the force (*i.e.*, low $\text{Force}_{\text{Initiation}}$) and be as accurate as possible (*i.e.*, low $\text{Force}_{\text{AmplError}}$ and $\text{Force}_{\text{Variability}}$).

As a control, we verified that the three motor components described above (*i.e.*, $\text{Force}_{\text{Initiation}}$, $\text{Force}_{\text{AmplError}}$ and $\text{Force}_{\text{Variability}}$) were closely related to the Error, and therefore were relevant for task success. To do so, we ran partial linear regressions on the Error data with $\text{Force}_{\text{Initiation}}$, $\text{Force}_{\text{AmplError}}$ and $\text{Force}_{\text{Variability}}$ as predictors to estimate the respective influence of each motor component on the Error, while controlling for the effect of the other components. Interestingly, we found that $\text{Force}_{\text{AmplError}}$ explained the largest part of variance in the Error ($r = 0.96 \pm 0.003$; $p < 0.05$ in 90/90 subjects). $\text{Force}_{\text{Initiation}}$ also explained a large part of variance in the Error ($r = 0.81 \pm 0.01$; $p < 0.05$ in 90/90 subjects), while $\text{Force}_{\text{Variability}}$ explained a smaller, yet significant in most subjects, part of variance ($r = 0.22 \pm 0.03$; $p < 0.05$ in 68/90 subjects). Hence, although all motor parameters were relevant for task success, the $\text{Force}_{\text{AmplError}}$ was the most influential factor.

2.5.4.2. Between-trial adjustments in motor commands

A second goal of the present study was to assess the effect of reward on between-trial adjustments in motor commands. Specifically, we aimed at evaluating how motor commands were adjusted based on reinforcement and sensory feedback in our three experimental groups.

To do so, for each trial_n we computed the absolute between-trial change (BTC) in Error (Error_{BTC}; see (Pekny et al., 2015; Uehara et al., 2019) for similar approaches in reaching tasks).

$$BTC\ Error = |Error(n + 1) - Error(n)|$$

In order to study how much motor commands were adjusted based on previous experience, we compared adjustments in motor commands following trials of different Error magnitudes. To do so, we first subtracted each subject's individual success threshold to the Error data. Hence, normalized Errors below 0 corresponded to successful trials and normalized Errors above 0 corresponded to failed trials. Then, we split the Error data in consecutive bins of 1 % of MVC and averaged the corresponding Error_{BTC}. This allowed us to compare Error_{BTC} following trials of similar Error_n across the groups.

As a first step, to better understand how motor commands were adjusted based on the reinforcement feedback, we compared Error_{BTC} following bins of Success or Failure trials of neighboring Error magnitudes (Bin_{Success}: -1% < Error_n < 0% MVC; Bin_{Failure}: 0% < Error_n < 1% MVC). Fixing the boundaries of Bin_{Success} and Bin_{Failure} allowed us to compare reinforcement-related adjustments between the groups while controlling for the magnitude of Error_n; an aspect that might directly influence between-trial adjustments. First, we performed a two-way ANOVA with the factors Outcome (Success or Failure) and Group_{TYPE}. We then computed reinforcement-based adjustments as the percentage change in Error_{BTC} in Bin_{Failure} compared to Bin_{Success}. This index allowed us to determine in a single measure how participants from the different groups adjusted their behavior based on the reinforcement obtained in the previous trial.

$$Reinforcement-based\ adjustments\ in\ Error_{BTC} = 100 \times$$

These analyses were conducted separately on the Day 1 and Day 2 data. We had to exclude 3 and 9 participants for Day 1 and Day 2 analyses, respectively, because they had less than 7 trials in at least one of the two bins (remaining subjects on Day 1: Group-s = 29;

Group-SR = 28; Group-SRR = 30; Day 2: Group-S = 26; Group-SR = 27; Group-SRR = 28). For the remaining participants, an average of 56 ± 3 and 39 ± 2 trials were included for each bin for Day 1 and Day 2 analyses, respectively. Reinforcement-based changes in $\text{Error}_{\text{BTC}}$ were compared between the groups through one-way ANOVAs with the factor $\text{Group}_{\text{TYPE}}$.

As a second step, we evaluated how participants adjusted movements when they could only rely on the sensory feedback. We compared $\text{Error}_{\text{BTC}}$ following bins of Failure trials of different Error magnitudes ($\text{Bin}_{\text{Small-Failure}}$: $0\% < \text{Error}_n < 1\%$ MVC; $\text{Bin}_{\text{Large-Failure}}$: $1\% < \text{Error}_n < 2\%$ MVC). In this case, the reinforcement feedback was the same in the two bins and the only difference between the trials consisted in the magnitude of the Error experienced at trial_n . Again, we first performed a two-way ANOVA with the factors Outcome (Small or Large Failure) and $\text{Group}_{\text{TYPE}}$. We then computed sensory-based adjustments as the percentage change in $\text{Error}_{\text{BTC}}$ in $\text{Bin}_{\text{Large-Failure}}$ compared to $\text{Bin}_{\text{Small-Failure}}$. This index allowed us to determine how participants adjusted their behavior based on the previous somatosensory experience, in the absence of any difference in the reinforcement feedback obtained.

$$\text{Sensory-based adjustments in } \text{Error}_{\text{BTC}} = 100 \times \frac{\text{Error}_{\text{BTC}} (\text{Bin}_{\text{Large-Failure}})}{\text{Error}_{\text{BTC}} (\text{Bin}_{\text{Small-Failure}})}$$

This analysis was first run on the Day 1 data. We had to exclude 12 participants because they had less than 7 trials in at least one of the two bins (remaining subjects: Group-S = 27; Group-SR = 28; Group-SRR = 24). For Day 2, applying the same procedure led to the exclusion of 29 subjects with a lower number of participants in the Group-SRR (15 subjects). For this reason, we ran another analysis where we exceptionally excluded participants only if they had less than 5 trials in one bin. This allowed us to keep a reasonable number of participants in each group (19 subjects excluded; remaining subjects: Group-S = 24; Group-SR = 26; Group-SRR = 21). Notably, both analyses (*i.e.*, with 7-trials or 5-trials cutoff) gave similar results and we only present the latter in the Results section. For the remaining participants, an average of 47 ± 3 and 29 ± 2 trials were included for each bin for Day 1 and Day 2 analyses, respectively. Sensory-based changes in $\text{Error}_{\text{BTC}}$ were compared between the groups through a one-way ANOVA with the factor $\text{Group}_{\text{TYPE}}$.

As a last step, we asked whether the effect of reward on between-trial adjustments in motor commands concerned all aspects of force control, or only specific motor components. To do so, we investigated reinforcement-based and sensory-based adjustments in $\text{Force}_{\text{Initiation}}$, $\text{Force}_{\text{Amplitude}}$ and $\text{Force}_{\text{Variability}}$, using the same method described above for the average

Error. We first performed two-way ANOVAs with the factors Outcome (reinforcement-based analysis: Success or Failure sensory-based analysis: Small or Large Failure) and Group_{TYPE}. Then, to assess reinforcement-based adjustments, we contrasted between-trial changes in Force_{Initiation} (Force_{Initiation-BTC}), Force_{AmplError} (Force_{AmplError-BTC}) and in Force_{Variability} (Force_{Variability-BTC}) following Bin_{Success} and Bin_{Failure}. Sensory-based adjustments were computed by contrasting Force_{Initiation-BTC}, Force_{AmplError-BTC} and Force_{Variability-BTC} following Bin_{Small-Failure} and Bin_{Large-Failure}. These data were compared between the groups through one-way ANOVAs with the factor Group_{TYPE}.

2.5.4.3. Group features, muscle fatigue and monetary gains

As a control, we verified that our 3 groups were comparable in terms of age, success threshold, Target_{Force} and Sensitivity to Reward and to Punishment (*i.e.*, as assessed by the SPSRQ questionnaire). As displayed in **Table 1**, one-way ANOVAs on these data did not reveal any significant differences between the groups.

We also assessed muscle fatigue on Day 1 and Day 2 (Derosiere and Perrey, 2012; Derosière et al., 2014) by expressing the MVC obtained after each session (MVC_{POST}) in percentage of the MVC measured initially (MVC_{PRE}). The relative change of MVC was not different according to the Group_{TYPE} (Day 1, $F_{(2,87)} = 0.51$, $p = 0.60$; Day 2, $F_{(2,87)} = 0.60$, $p = 0.55$; **Table 1**). As an additional safety check, we wanted to make sure that the decrements in MVC caused by the training period of Day 1 could not impair performance. To test this, we compared MVC_{POST} (expressed in % of MVC_{PRE}) with a fixed value of 10% of MVC_{PRE} (*i.e.*, corresponding to the Target_{Force}) through Bonferroni-corrected single sample t-tests. This analysis revealed that MVC_{POST} levels were always significantly above the Target_{Force} (Group-S: $t_{(29)} = 35.84$, $p < 0.001$; Group-SR: $t_{(29)} = 34.14$, $p < 0.001$ and Group-SRR: $t_{(29)} = 34.44$, $p < 0.001$). Hence, force decrements caused by the training were comparable between groups and are unlikely to have limited task performance.

In a final step, we checked that the monetary gains obtained at the end of the experiment were similar between groups. Subjects received a fixed show-up fee corresponding to 10 euros/hour of experiment. In addition, participants also gained a monetary bonus. This bonus was set at 10 euros for subjects in Group-SR and Group-S while it was variable from 0 to 20 euros according to the Group-SRR performance (gain of 8 cents per successful trial in

Block-SRR). Importantly, this bonus procedure in Block-SRR was determined to match that obtained in the other groups; it corresponded to 10.4 ± 0.67 euros. A t-test revealed that the total ending remuneration was similar across the different Group_{TYPE}S ($t_{(29)} = 0.57$; $p = 0.57$).

	Group-S (n=30)	Group-SR (n=30)	Group-SRR (n=30)	F _(2,87)	p
Age (years)	23.9 \pm 0.67	23.3 \pm 0.50	23.9 \pm 0.43	0.34	0.71
Gender (number of females)	19	19	20	/	/
Success Threshold (% MVC)	2.8 \pm 0.01	2.8 \pm 0.01	2.9 \pm 0.01	0.13	0.88
Target _{Force} (Newtons)	5.3 \pm 0.30	4.7 \pm 0.25	5.1 \pm 0.21	1.16	0.31
Sensitivity to reward (score)	37.1 \pm 1.18	35.6 \pm 1.11	37.5 \pm 1.14	0.79	0.46
Sensitivity to punishment (score)	42.3 \pm 1.59	42.0 \pm 1.59	40.9 \pm 1.46	0.22	0.81
Muscle fatigue – Day 1 (MVC _{POST} in % of MVC _{PRE})	95.7 \pm 2.39	99.0 \pm 2.61	98.6 \pm 2.57	0.51	0.60
Muscle fatigue – Day 2 (MVC _{POST} in % of MVC _{PRE})	94.6 \pm 2.74	93.9 \pm 1.52	97.0 \pm 1.89	0.60	0.55

Table 1. Group features and muscle fatigue in the three experimental groups (mean \pm SE). The 2 last columns provide the results of one-way ANOVAs' ran with the factor Group_{TYPE}.

2.6. Supplementary materials

Effect of reinforcement and reward on different levels of force control

As a control, we examined performance at the different levels of force control examined in Figure 4. This analysis revealed that in contrast to the controls, on average, subjects of the Group-SRR improved their performance at all three levels of force control (normalized Post-training performance below 100%). This result suggests that the beneficial effect of reward on motor learning results from a true improvement in the speed-accuracy trade-off of movements with motor commands being faster but also more accurate, in line with previous studies on motor control (Manohar et al., 2015; Codol et al., 2020).

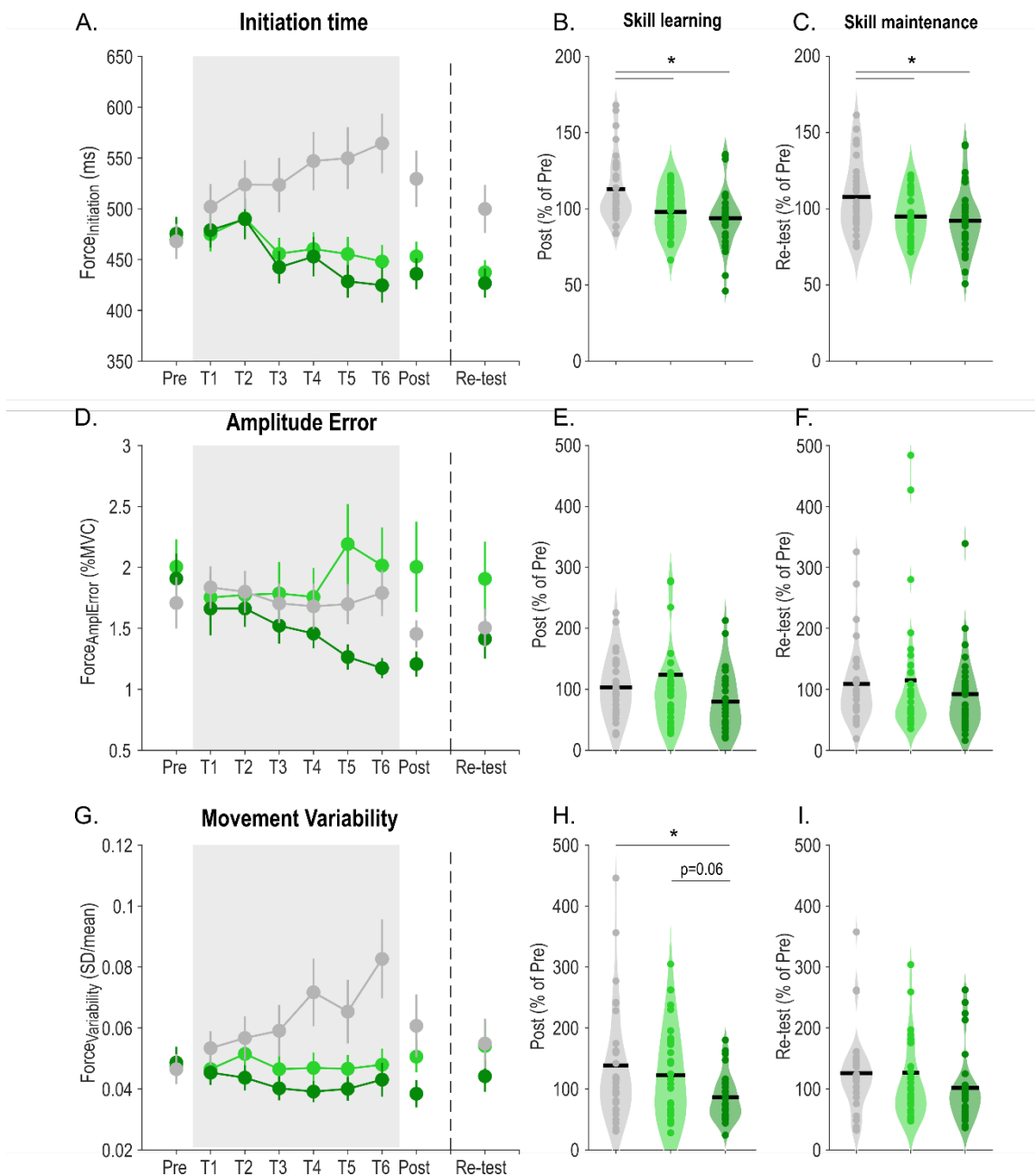


Figure S1. Motor components. **A) Initiation time.** Average initiation time (in ms) is represented across practice for the three experimental groups (grey: Group-S, light green: Group-SR, dark green: Group-SRR). The grey shaded area highlights the blocks concerned by the reinforcement manipulation. The remaining blocks were performed with knowledge of performance only (i.e., in a Block-SR setting). **Changes in Initiation time associated with Skill learning (B) and maintenance (C)** in the three experimental groups. Horizontal black lines represent group average. Each dot represents a single subject. **D) Amplitude Error.** Average amplitude error (in % MVC) across training. **Changes in Amplitude error associated with Skill learning (E) and maintenance (F)** in the three experimental groups. Note that one extreme value is not represented in E) for clarity purposes ($\text{Force}_{\text{AmplitudeError}}=778.1\%$). **G) Variability.** Average amplitude error (in % MVC) across training. **Changes in Variability associated with Skill learning (H) and maintenance (I)** in the three experimental groups. Note that one extreme value is not represented in I) for clarity purposes ($\text{Force}_{\text{Variability}}=595.6\%$).

3. Study 2: Reward timing matters in motor learning

Pierre Vassiliadis^{1,2*}, Aegryan Lete¹, Julie Duque¹, Gerard Derosiere¹

¹ Institute of Neuroscience, Université Catholique de Louvain, 1200, Brussels, Belgium.

² Defitech Chair for Clinical Neuroengineering, Center for Neuroprosthetics (CNP) and Brain Mind Institute (BMI), Swiss Federal Institute of Technology (EPFL), 1202, Geneva, Switzerland

Reference: Vassiliadis P, Lete A, Duque J, Derosiere G (2022). Reward timing matters in motor learning. iScience, 25: 104290.

Personal contribution: Conceptualization, Methodology, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Funding acquisition.

* Correspondence

Pierre Vassiliadis

Institute of Neuroscience

Université catholique de Louvain

53, Avenue Mounier

COSY- B1.53.04

B-1200 BRUSSELS

Belgium

Email: pierre.vassiliadis@uclouvain.be

Web: <http://www.coactionslab.com>

3.1. Abstract

Reward can improve motor learning and the consolidation of motor memories. Identifying the features of reward feedback that are critical for motor learning is a necessary step for successful integration into rehabilitation programs. One central feature of reward feedback that may affect motor learning is its timing – that is, the delay after which reward is delivered following movement execution. In fact, research on associative learning has shown that short and long reward delays (*e.g.*, 1 and 6 s following action execution) activate preferentially the striatum and the hippocampus, respectively, which both contribute with varying degrees to motor learning. Given the distinct functional role of these two areas, we hypothesized that reward timing could modulate how people learn and consolidate a new motor skill. In sixty healthy participants, we found that delaying reward delivery by a few seconds influenced motor learning. Indeed, training with a short reward delay (*i.e.*, 1 s) induced continuous improvement in performance across training, while a long reward delay (*i.e.*, 6 s) led to initially high learning rates that were followed by an early plateau in the learning curve and a lower performance at the end of training. Moreover, participants who successfully learned the skill with a short reward delay displayed overnight consolidation, while those who trained with a long reward delay exhibited an impairment in the consolidation of the motor memory. Overall, our data show that reward timing affects motor learning, potentially by modulating the engagement of different learning processes, a finding that could be exploited in future rehabilitation programs.

3.2. Introduction

When delivered following well-executed movements, reward can boost motor learning (Galea et al., 2015a; Chen et al., 2017; Dhawale et al., 2017; Vassiliadis et al., 2021) and the consolidation of motor memories (Abe et al., 2011). This observation has raised hope for rehabilitation, where reward is regarded as a promising means to magnify the positive effects of practice on motor control (Therrien et al., 2016, 2020; Quattrocchi et al., 2017; Vassiliadis et al., 2019; Vassiliadis and Derosiere, 2020). Yet, this branch of research is only burgeoning, and a current challenge in the field is to identify the features of reward feedback that may be critical for motor learning.

Recent studies have started to tackle this issue, showing that the magnitude (Vassiliadis et al., 2021), the valence (Galea et al., 2015a) and the stochasticity (Dayan et al., 2014a) of reward feedback bear all a decisive impact on motor learning. Another key feature of reward feedback that may directly affect motor learning is its timing – that is, the delay after which reward is delivered following movement execution. As such, previous studies have shown that reward prediction error signals, which are key for reward-based learning, are not only modulated by the value of the reward but also depend on the timing at which it is delivered (Fiorillo et al., 2008; Kobayashi and Schultz, 2008; Klein-Flügge et al., 2011). Moreover, converging lines of evidence from neuroimaging and electroencephalographic studies indicate that different brain structures exhibit activity changes in response to reward feedback depending on its timing. Indeed, in associative learning tasks, short reward delays (e.g., provided 1 s following action execution) activate a fronto-striatal network, while long reward delays (e.g., 6 s following execution) evoke changes in the activity of the hippocampus primarily (Foerde and Shohamy, 2011; Peterburs et al., 2016). Further, Parkinson's disease and ADHD patients, both known to exhibit striatal dysfunction (Mehler-Wex et al., 2006), are impaired in learning action-outcome associations based on short reward delays (Foerde and Shohamy, 2011; Foerde et al., 2012; Gabay et al., 2018; Weismüller et al., 2018), while amnesic patients with damage to the hippocampus are unable to learn associations with long reward delays (Foerde et al., 2013). Altogether, these findings indicate that the processing of reward preferentially engages striatum- or hippocampus-centred networks depending on the timing at which it is delivered.

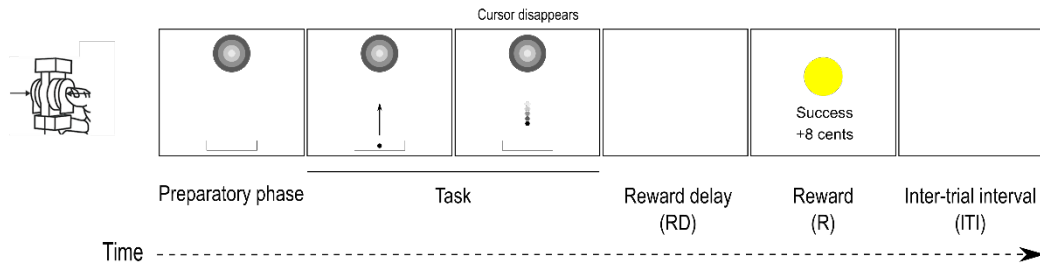
The striatum and the hippocampus show varying contributions during motor learning and consolidation (Schendan et al., 2003; Doyon and Benali, 2005; Fernández-Seara et al., 2009; Krakauer et al., 2019), which are thought to underlie the operation of distinct learning processes (Albouy et al., 2008, 2013a). Hence, it is sensible to assume that reward may boost different motor learning processes – potentially relying on the striatum or the hippocampus – depending on the timing at which it is delivered. Notably, previous studies on reward-based motor learning have only exploited short reward delays, impeding one to test this hypothesis directly. Here, we tested this idea by evaluating the performance of sixty healthy participants in a skill learning task (Vassiliadis et al., 2021), where reward was delivered either at a short or at a long delay following movement execution. We found that delaying reward delivery by a few seconds influenced the dynamics of learning. Indeed, training with a short reward delay induced continuous improvement in performance across training, while a long reward delay led to initially high learning rates that were followed by an early plateau in the learning curve and a lower endpoint performance. Moreover, participants who successfully learned the skill with a short reward delay displayed overnight consolidation, while those who trained with a long reward delay exhibited an impairment in the consolidation of the motor memory. Altogether, the present results provide evidence that reward timing can strongly influence motor learning, a finding that could be exploited in future rehabilitation protocols.

3.3. Results

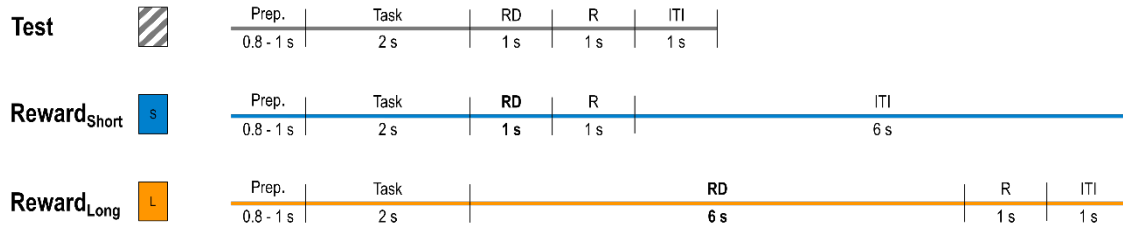
Sixty healthy participants practiced a pinch-grip force task over two consecutive days. Participants were required to hold a pinch grip transducer in their right hand and to squeeze it as quickly as possible in order to move a cursor displayed on a computer screen in front of them, from an initial position to a fixed target (**Figure 1A**; (Vassiliadis et al., 2021)). The force required to reach the target ($\text{Target}_{\text{Force}}$) corresponded to 10 % of the individual maximum voluntary contraction (MVC). In most of the trials (90 %), participants practiced the task with very limited sensory feedback: the cursor disappeared when the generated force reached half of the $\text{Target}_{\text{Force}}$ (see STAR Methods for more details on the task). To learn the task, subjects were provided with six Training blocks (T1 to T6; 40 trials each; *i.e.*, total of 240 training trials; **Figure 1B**) in which they received reinforcement feedback (*i.e.*, indicating Success or Failure) associated to a monetary reward. Success on the task was determined based on the Error, defined as the absolute force difference between the $\text{Target}_{\text{Force}}$ and the exerted force (Abe et al., 2011; Steel et al., 2016).

In different groups of participants, we varied the delay between the end of the movement period and the delivery of the reward during the Training blocks. As such, $\text{Group}_{\text{Short}}$ subjects trained with a short reward delay (*i.e.*, 1 s) while participants of the $\text{Group}_{\text{Long}}$ performed the task with a long reward delay (*i.e.*, 6 s). The total duration of the trials was kept constant by modulating the inter-trial interval (ITI; 6 s in $\text{Group}_{\text{Short}}$ and 1 s in $\text{Group}_{\text{Long}}$). Before, immediately and 24 hours after training, all participants performed Test blocks with no reward, a short reward delay (1 s) and a short ITI (1 s). Notably, the groups were comparable for a variety of features including Pre-training success rates, difficulty of the task, force required, sensitivity to reward and punishment, fatigue and final monetary gains (**Figure 1C**, **Table 1**). Altogether, this design allowed us to investigate the specific effect of reward timing on motor learning and consolidation.

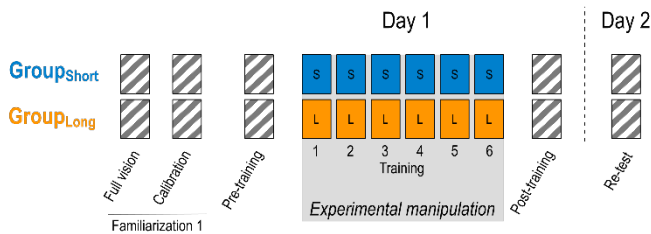
A. Trials



B. Block types



C. Training



D.

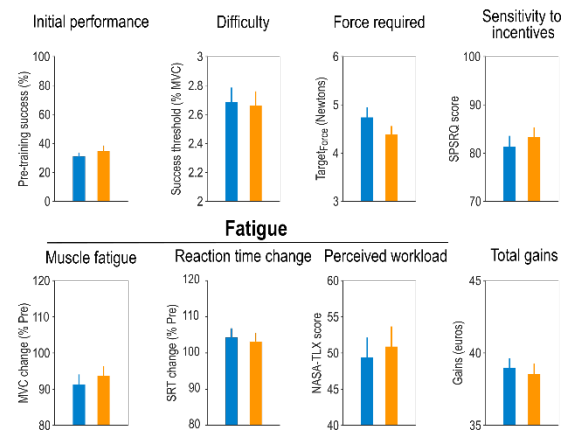


Figure 1. Motor skill learning task. **A. Time course of a trial in the motor skill learning task.** Each trial started with the appearance of a sidebar and a target. After a variable preparatory period (0.8-1s), a cursor appeared in the sidebar, playing the role of a “Go” signal. At this moment, participants were required to pinch the force transducer to bring the cursor into the target as quickly as possible and maintain it there until the end of the task (2 s). Notably, on most trials, the cursor disappeared halfway towards the target (as displayed here). Then, after a delay, a reward (R) appeared consisting in a reinforcement feedback and a monetary reward (a successful trial is shown here). Trials ended with an inter-trial interval (ITI). **B. Durations in the different block types.** Reward delays (RD) and ITIs were manipulated. Test blocks included a short reward delay (1 s), a short ITI (1 s) and no monetary reward (*i.e.*, only reinforcement). Reward_{Short} and Reward_{Long} blocks included monetary rewards and were performed with a short (1 s) and long (6 s) reward delay, respectively. The total duration of the trials was kept constant between Reward_{Short} and Reward_{Long} by varying the ITI. **C. Training procedure.** On Day 1, all participants performed two familiarization blocks in a Test blocks condition. The first one involved full vision of the cursor while the second one provided only partial vision and served to calibrate the difficulty of the task on an individual basis (See STAR Methods). Then, Pre- and Post-training Test blocks assessments were separated by 6 blocks of training in the condition corresponding to each individual group (Reward_{Short} for Group_{Short}, Reward_{Long} for Group_{Long}). Day 2 consisted in a short re-familiarization (5 trials with full vision, not represented) followed by a Re-test assessment (1 Test block). **D. Control analyses.** Group_{Short} and Group_{Long} were comparable for a variety of factors including initial performance, task difficulty, required force

to reach the target, sensitivity to reward and punishment (as assessed by the SPSRQ questionnaire), muscular and cognitive fatigue and final monetary gains (see also Table 1).

Training with long reward delays modifies the dynamics of motor learning

As a first step, we evaluated performance on the task by computing the average success rate per Training_{Block} (T1 to T6, **Figure S1**). To compare the learning process between the groups, we performed a Linear Mixed Model (LMM), with TRAINING_{BLOCK} and GROUP_{TYPE} (and their interaction) modeled as categorical fixed factors. Overall, participants of both groups significantly improved their success rates over training (main effect of TRAINING_{BLOCK}: $F_{(5, 290)} = 4.30$; $p < 0.001$; **Figure 2A**). Most importantly, the improvement in success rate over the blocks depended on the Group_{Type}, as revealed by a significant TRAINING_{BLOCK} \times GROUP_{TYPE} interaction ($F_{(5, 290)} = 2.69$; $p = 0.021$; **Figure 2A**). Interestingly, between-groups post-hoc comparisons further revealed that endpoint performance (*i.e.*, success rate at T6) was significantly lower in Group_{Long} than in Group_{Short} ($p = 0.045$; **Figure 2B**). Note though that this significant result would not survive multiple comparisons corrections, and therefore needs to be taken with caution. Conversely, success rates at all other Training_{Blocks} were comparable between the two groups (all $p > 0.22$; **Figure 2A**). This result suggests that reward timing influenced the dynamics of learning leading to a poorer endpoint performance in Group_{Long}.

In order to confirm these results, we ran another LMM on the single-trial Error data (**Figure 3C, Supplementary Table 1**) with the predictors TRAINING_{TRIAL} (continuous) and GROUP_{TYPE} (categorical). Focusing on the Error allowed us to evaluate the effect of reward timing on motor learning without having to bin the data in any way. This analysis confirmed that learning was influenced by the timing at which rewards were provided (**Figure 2D**; TRAINING_{TRIAL} \times GROUP_{TYPE} interaction: $F_{(1, 12114)} = 9.00$; $p = 0.0027$). This interaction reflected the fact that the slope of learning (*i.e.*, a proxy of the learning rate) was steeper in Group_{Short} than in Group_{Long} (**Figure 2D**). Importantly, comparison of the intercepts in both groups did not show any significant difference ($p = 0.60$), suggesting that the learning effect could not be explained by differences in initial performance. Put together, these two analyses show that training with long reward delays impairs the acquisition of a new motor skill.

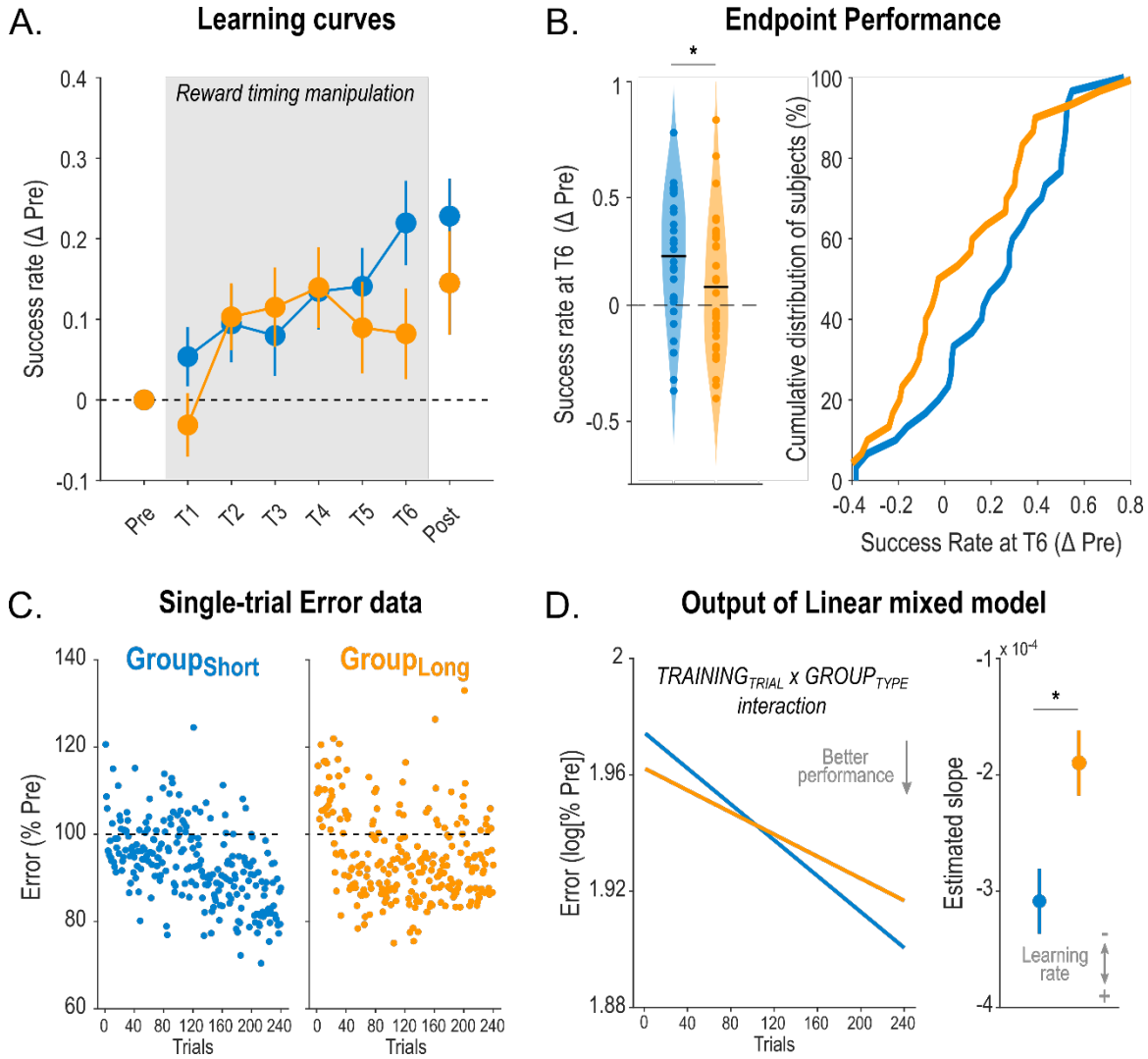


Figure 2. Effect of reward timing on motor skill learning. **A. Learning curves.** Proportion of successful trials (expressed as a difference with the individual Pre-training success rate) is represented across practice for the two experimental groups (blue: Group_{Short}, $n=30$, orange: Group_{Long}, $n=30$). The grey shaded area highlights the blocks concerned by the reward timing manipulation. The remaining blocks were Test blocks. **B. Endpoint performance.** Violin plot showing success rates at the end of the training period (i.e., at T6) for each participant (left panel) and the corresponding cumulative distributions of the data (right panel). **C. Single-trial Error data.** Normalized Error data obtained during training are averaged across groups and plotted for each single trial. Note that lower Errors were associated to better performance. **D. Output of LMM on the Error data.** Output of LMM run on the log-transformed Error data is plotted for each group (left panel). The Error data was log-transformed prior to respect key assumptions of LMMs (see STAR Methods section). The significant TRAINING_{TRIAL} X GROUP_{TYPE} interaction shows that the slope of learning was steeper in Group_{Short} than in Group_{Long} (right panel). Estimated intercepts were not different between groups ($p = 0.60$). Notably, more negative slopes reflect larger learning rates. *: significant difference ($p < 0.05$). Data are represented as mean \pm SE.

An important aspect of our experimental design is that we increased the duration of the ITI in Group_{Short} relative to Group_{Long} (6 s and 1 s, respectively; **Figure 1B**), in order to

match the total duration of the trials in both groups despite differences in reward timing. To evaluate how such manipulation may have impacted learning in our task, we added in the analysis another group of participants, who trained with a short reward delay (0.5 s) and an intermediate ITI (3 s; Group_{Short-PastStudy}, $n = 30$; from (Vassiliadis et al., 2021)). We reasoned that, if differences in learning dynamics were truly driven by differences in reward timing but not by differences in ITI duration, learning in Group_{Short-PastStudy} should be similar than in Group_{Short}, and therefore different than in Group_{Long}. As above, we ran a first LMM on the Success data with the factors TRAINING_{BLOCK} and GROUP_{TYPE}. Consistent with our hypothesis, we found a significant TRAINING_{BLOCK} \times GROUP_{TYPE} interaction ($F_{(10, 435)} = 2.84$; $p = 0.0020$), and post-hoc tests showed (1) no significant difference between Group_{Short-PastStudy} and Group_{Short}, at any Training_{Block} (all $p > 0.22$) and (2) a marginally significant difference in endpoint performance when comparing Group_{Short-PastStudy} and Group_{Long} (*i.e.*, $p = 0.048$ and 0.052 at T5 and T6, respectively; **Figure S2A, B**). In order to confirm these effects on non-binned, single-trial data, we ran the same LMM on the Error variable (*i.e.*, same analysis as in **Figure 2D**), but with the addition of the data from Group_{Short-PastStudy} (**Figure S2C**). Again, there was a TRAINING_{TRIAL} \times GROUP_{TYPE} interaction ($F_{(2, 18213)} = 14.99$; $p < 0.001$), that was driven by differences in the slopes of the learning curves between the groups (**Figure S2D**). As expected, post-hoc tests showed that the slopes were steeper in Group_{Short-PastStudy} than in Group_{Long} ($p < 0.001$). However, slopes were also steeper in Group_{Short-PastStudy} than in Group_{Short} ($p = 0.018$), suggesting that longer ITIs may also have some detrimental effect on the learning rates. Notably, no differences were found when comparing the intercepts (all $p > 0.59$). The apparent discrepancy between the LMM results obtained for Success vs. for Error data possibly arises from the fact that the former analysis was based on block-averaged performance, while the latter focused on the learning rates estimated based on single-trial data (reflected by the coefficient associated to TRAINING_{TRIAL} in the LMM). Notably, the difference in learning rate between Group_{Short} and Group_{Short-PastStudy} must be taken with caution as in addition to presenting a longer ITI duration, Group_{Short} also presented a slightly longer reward delay relative to Group_{Short-PastStudy} (1 s vs. 0.5 s, respectively), which might also have been detrimental for learning rates. Still, if anything, this analysis suggests that long ITIs had rather a negative impact on learning. Yet, participants of Group_{Short} exhibited better learning rates than participants of Group_{Long}, indicating that the positive effect of the shorter reward delays overcame the negative impact of the longer ITI

in this group. Overall, this analysis suggests that both reward delay and ITI duration influence motor skill learning but that reward delay plays a more prominent role in shaping learning. A direct corollary to this is that we may have underestimated the negative impact of long reward delays on learning when comparing Group_{Long} with Group_{Short}, given that this negative effect was partially counteracted by the longer ITI duration in the latter group.

To evaluate total learning, we computed success rates at Post-training, which was performed in a Test block setting in both groups. Importantly, whereas comparing performance at T6 informed us about the effect of our particular training features on learning within each training condition, Post-training performance provides information about total learning in the task, in identical Test block conditions. Overall, success rates at Post-training increased by 22.8 ± 4.69 % in Group_{Short}, and 14.5 ± 6.42 % in Group_{Long} with respect to Pre-training. Interestingly, success rates at Post-training were significantly different from 0 in Group_{Short} despite Bonferroni correction of the significance threshold (cutoff for significance: $p = 0.025$; $t_{(29)} = 4.87$, $p < 0.001$), indicative of a significant improvement in performance with respect to Pre-training. In contrast, success rates at Post-training were not significantly different from 0 in Group_{Long} after Bonferroni correction of the significance threshold ($t_{(29)} = 2.26$, $p = 0.031$). However, a t-test on these data did not show any significant difference between the Group_{TYPES} ($t_{(58)} = 1.05$; $p = 0.30$). Hence, reward timing only induced a subtle change in total learning that did not reach significance when comparing directly the groups.

Results of the first analysis showed that training with long reward delays was generally associated with lower learning rates (**Figure 2D**), leading to a reduced endpoint performance (**Figure 2B**). Inspection of the raw data (**Figure 2A, 2C**) also suggested that the learning dynamics could be different between the groups. To evaluate this, we ran three additional analyses. First, we asked for each group of participants whether the learning curves were best modeled as a linear or non-linear logarithmic function. Interestingly, we found that the data from Group_{Short} were better approximated by a linear function (linear fit: Adjusted $R^2 = 0.25$; logarithmic fit: Adjusted $R^2 = 0.21$; **Figure 3A**), whereas the Group_{Long} learning curve was better modeled with a logarithmic fit (linear fit: Adjusted $R^2 = 0.063$; logarithmic fit: Adjusted $R^2 = 0.18$; **Figure 3B**). This suggests that training with short reward delays was associated with generally stable learning rates while training with long reward delays was related to fast learning rates early on during practice that was quickly followed

by a plateau in performance. Indeed, simple linear regressions on the Success data showed that 76.7 % (23/30) of participants of Group_{Long} exhibited higher learning rates in the early than in the late phase of training, while this percentage was 46.7% (14/30) in Group_{Short} (Fisher's exact test on the proportions: $p = 0.033$; **Figure S3**). To further evaluate how learning rates varied across early and late phases of practice, we ran the same LMM on the Error data as described above (**Figure 2D**) with the addition of the factor TRAINING_{PHASE} which was modeled as a categorical fixed effect with two modalities (Training_{Early} vs. Training_{Late} for the first and last 120 trials of training, respectively; **Supplementary Table 2**). Interestingly, we found a triple TRAINING_{TRIAL} X GROUP_{TYPE} X TRAINING_{PHASE} interaction ($F_{(1,12110)} = 40.62$; $p < 0.001$), demonstrating that learning rates (reflected by the coefficients associated with the factor TRAINING_{TRIAL}) varied not only depending on the group but also based on the phase of practice. As illustrated on **Figure 3C**, this interaction was due to the fact that at Training_{Early}, the estimated learning rate was significantly higher in Group_{Long} than in Group_{Short}, whereas it was the opposite at Training_{Late} (both $p < 0.001$). Moreover, learning rates were significantly higher at Training_{Early} than at Training_{Late} in Group_{Long} ($p < 0.001$). In Group_{Short}, there was a trend for the opposite effect (*i.e.*, higher learning rates at Training_{Late} than at Training_{Early}; $p = 0.056$). Again, intercepts at Training_{Early} were not significantly different ($p = 0.36$), indicating comparable initial levels of performance in the two groups. Hence, this analysis confirms that reward timing impacts learning dynamics. More specifically, training with short reward delay appears to induce continuous gains in performance during training, while long reward timings favor non-linear dynamics with larger initial learning rates that then drop significantly, indicative of a plateau in learning.

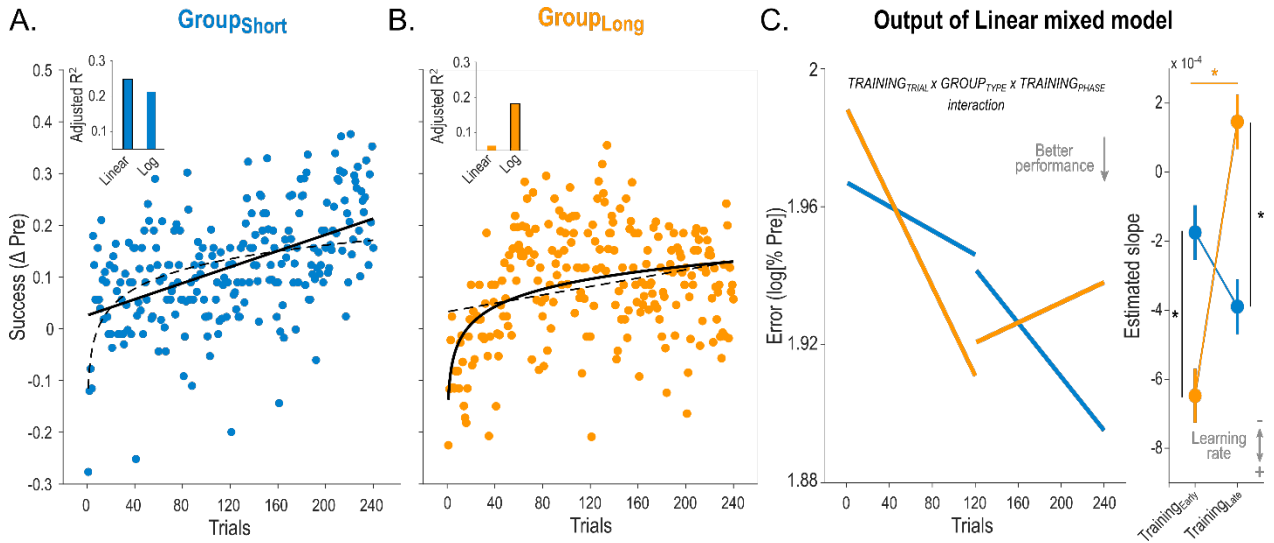


Figure 3. Effect of reward timing on the dynamics of learning. **A and B. Linear and Non-linear fits on Success learning data.** The group-averaged single-trial Success data in Group_{Short} (A, $n = 30$) and Group_{Long} (B, $n = 30$) were fitted with either a linear or a non-linear logarithmic function. Importantly, the best fit (i.e., represented by the solid trace) was linear for Group_{Short} and logarithmic for Group_{Long}, suggesting that the dynamics of learning were different in both groups. **C. Output of the LMM on the Error data including the factor TRAINING_{PHASE}.** Output of LMM run on the log-transformed Error data is plotted for each group (left panel). The significant TRAINING_{TRIAL} \times GROUP_{TYPE} \times TRAINING_{PHASE} interaction shows that, in the early phase of practice, the learning rates – reflected by the slope of learning – were steeper in Group_{Long} than in Group_{Short} ($p < 0.001$). This was the opposite in the late phase of practice ($p < 0.001$). Notably, there was also a significant reduction of the learning rates from Training_{Early} to Training_{Late} in Group_{Long} ($p < 0.001$; orange star), while there was a tendency for an increase in learning rates in Group_{Short} ($p = 0.056$). Note that lower Errors were associated to better performance and that more negative slopes reflect larger learning rates. *: significant difference ($p < 0.05$). Data are represented as mean \pm SE.

Training with long reward delay impairs overnight skill consolidation in learners

As a last step, we investigated the impact of the reward timing experienced during training on Day 1 on overnight consolidation of the skill (i.e., on Day 2). To evaluate consolidation, we ran a LMM on the normalized success rates obtained at Post-training of Day 1 and at Re-test of Day 2 (i.e., both performed in a Test block setting) with TEST_{BLOCK} and GROUP_{TYPE} as fixed effects. This analysis did not reveal any main effect of TEST_{BLOCK} ($F_{(1, 58)} = 0.75$; $p = 0.39$) and GROUP_{TYPE} ($F_{(1, 86.14)} = 1.18$; $p = 0.28$) nor any TEST_{BLOCK} \times GROUP_{TYPE} interaction ($F_{(1, 58)} = 0.48$; $p = 0.49$). The same results were obtained when running the LMM on the single-trial Error data. However, a potential caveat of these analyses is that they included participants who did not learn the task on Day 1 and even exhibited a deterioration of performance with practice. In these participants, a Re-test performance

similar to the Pre-training level would be considered as evidence for an offline gain in performance, when it would actually only reflect a return to the baseline level of performance. In a second step, we therefore focused on the learners – that is, participants who exhibited an improvement of performance with practice on Day 1 ($n = 22$ and 18 in Group_{Short} and Group_{Long}, respectively). This allowed us to compare offline consolidation in participants who actually responded to the training and who also happened to be very close in terms of Post-training success rates (**Figure 4A**), a crucial aspect in order to interpret any overnight change in performance. Interestingly, this analysis revealed a $\text{TEST}_{\text{PHASE}} \times \text{GROUP}_{\text{TYPE}}$ interaction ($F_{(1, 38)} = 5.77$; $p = 0.021$). In fact, as mentioned above, performance was strongly similar between learners of the two groups at Post-training on Day 1 ($p = 0.65$), but diverged between the groups on Day 2. Indeed, success rates were significantly reduced on Day 2 relative to Day 1 in Group_{Long} ($p = 0.0021$), but remained stable from one day to another in Group_{Short} ($p = 0.96$, **Figure 4B**). The difference in performance on Day 2 between the groups was only at the trend level ($p = 0.096$). Notably, this interaction was replicated when running the LMM on the single-trial Error data ($F_{(1, 2615.9)} = 7.25$; $p = 0.0071$). This indicates that delaying rewards on Day 1 impaired consolidation of the motor skill on Day 2 in learners. Overall, our results support the view that short or long reward delays support qualitatively different motor learning processes during training, leading to different consolidation of the skill.

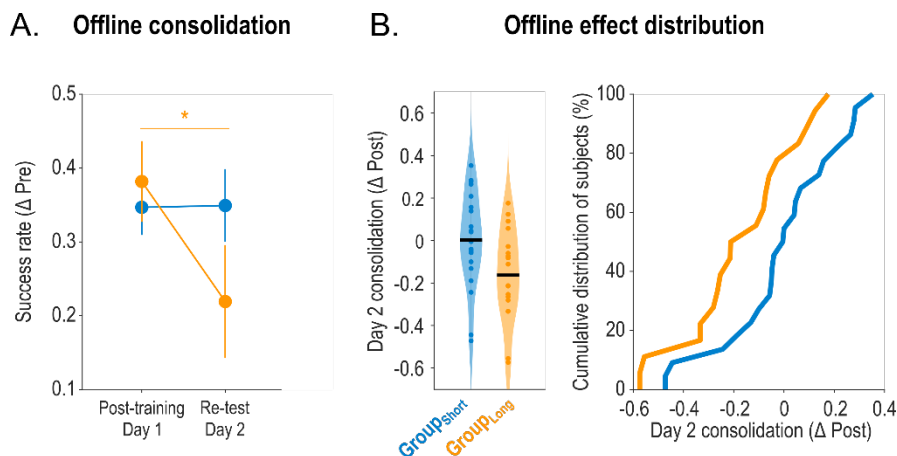


Figure 4. Effect of reward timing on overnight consolidation of the motor skill. A. Offline consolidation of the motor skill. Proportion of successful trials (expressed as a difference with the individual Pre-training success rate) at Post-training (Day 1) and Re-test (Day 2). Both assessments were Test blocks. This analysis only considered participants who demonstrated skill learning on Day 1 ($n = 22$ and 18 in Group_{Short} and Group_{Long}, respectively). Notably, a significant $\text{TEST}_{\text{PHASE}} \times$

GROUP_{TYPE} interaction showed that while Group_{Short} participants performed comparably well in Post-training and Re-test, demonstrating offline consolidation of the skill, this process was altered in Group_{Long}. **B. Offline effect distribution.** Violin plot showing the offline effect (Success rate in Day 2 – Success rate in Post-training) for each participant (left panel) and the corresponding cumulative distributions of the data (right panel).

3.4. Discussion

Previous studies have shown that reward timing can influence the response of brain structures involved in reward processing during associative learning (Fiorillo et al., 2008; Kobayashi and Schultz, 2008; Foerde and Shohamy, 2011; Klein-Flügge et al., 2011; Foerde et al., 2013). Inspired by these neurophysiological findings, we asked whether reward timing can also influence how people learn and consolidate a new motor skill. We found that delaying reward delivery by a few seconds influences motor learning dynamics: training with a short reward delay induced continuous gains in performance, while a long reward delay allowed high initial learning rates that were followed by an early plateau in the learning curve and a lower endpoint performance. Moreover, among participants who successfully learned the skill, those who trained with a short reward delay displayed overnight consolidation, while those who learned the task with a long reward delay exhibited an impairment in the consolidation of the motor memory. Overall, our findings show that reward timing can influence how the brain learns and consolidates new motor skills.

An important finding of our study is the overall impairment of learning when training with long compared to short reward delays, which was reflected by a reduction of global learning rates as well as endpoint performance during training. As such, efficient reward-based motor learning relies on the mapping between somatosensory sensations (e.g., elicited by the generated force in the present task) and the associated reward (Bernardi et al., 2015; Sidarta et al., 2016; Vassiliadis et al., 2021), and somatosensory working memory is known to decay quickly following movement execution, after only a few seconds (Harris et al., 2001; Sidarta et al., 2018). Hence, it is possible that delaying reward delivery blunted the reinforcement of somatosensory working memory (Sidarta et al., 2018), explaining the limited learning observed in the subjects of Group_{Long}. Another complementary interpretation is that reward delays affected the precision of dopaminergic reward prediction errors in the striatum (Fiorillo et al., 2008; Kobayashi and Schultz, 2008). In this case, the temporal uncertainty caused by increased reward delays would alter the association between the movement and the corresponding outcome due to imprecise learning signals in the reward system (Fiorillo et al., 2008). Overall, the present data indicate that the temporal contingency between movements and rewards is a decisive aspect of reward-based motor learning.

Despite clear effects of reward delay during the training phase, we did not find any between-group difference at Post-training (*i.e.*, performed in a Test block setting, with short reward delay and ITI). There are several ways to interpret this finding. First, it is possible that reward timing has dissociable effects on motor performance and learning (Schmidt and Bjork, 1992; Soderstrom and Bjork, 2015). As such, the introduction of reward delays during training may generally alter motor performance, but not the learning of the skill, as evaluated in the Post-training Test block. A second interpretation is that the reward timing manipulation affected the learning process but was not sufficient to evoke lasting behavioral differences. This would be in line with previous work on associative learning showing that reward delays modulate brain signatures of reward processing in healthy subjects but not behavioral learning in the Test phase (Foerde and Shohamy, 2011). Yet, the same researchers also found robust learning effects when testing populations of patients that presented specific dysfunctions of the striatum or the hippocampus (Foerde and Shohamy, 2011; Foerde et al., 2012, 2013). A possibility is therefore that our reward delay manipulation was not sufficient to modulate behavioral learning in young healthy individuals (potentially due to other compensatory learning mechanisms), but may still prove efficient when testing populations of patients exhibiting specific lesions of the networks involved in reward processing.

The differences in learning dynamics observed in subjects trained with short and long reward delays may indicate that reward boosted processes presenting different temporal dynamics. As such, a prevalent view in the field is that motor learning entails the operation of distinct processes, with either slow (*i.e.*, developing over a few trials) or fast (*i.e.*, developing over tens/hundreds of trials) temporal dynamics (Smith et al., 2006). The slow process is characterized by both a low learning rate and a sluggish forgetting of the acquired behaviour and is thought to reflect implicit learning (Trewartha et al., 2014; McDougale et al., 2015). In contrast, the fast process entails both a high learning rate and a quick forgetting of the new behavior and supports the explicit learning of new motor behaviors (Trewartha et al., 2014; McDougale et al., 2015). The nature of our task did not allow us to evaluate the relationship between reward timing and the relative contribution of implicit and explicit learning. Still, people who trained with a short reward delay exhibited learning dynamics that presented a low initial learning rate and a clear overnight consolidation – reminiscent of the

slow process, while those who trained with a long reward delay exhibited a high initial learning rate and an overnight forgetting of the motor memory – evocative of the fast process. Based on these results, one may suggest that short reward delays preferentially facilitate the slow (putatively more implicit) process, while long reward delays may favour the fast (potentially more explicit) learning process, accentuating their respective contribution to subjects' improvements. Interestingly, the striatum and hippocampus, which are involved in processing rewards offered after short and long delays, respectively (Foerde and Shohamy, 2011; Foerde et al., 2012, 2013), exhibit a pattern of activation during motor learning that is consistent with this interpretation. As such, the striatum displays slow, continuous changes in activity over the course of motor learning whereas the hippocampus usually exhibits a fast increase in activity in the early phase of learning that wanes later on (Schendan et al., 2003; Albouy et al., 2008, 2012, 2013a; Rieckmann et al., 2010; Doyon et al., 2018). Notably though, this parallel between our behavioral results and previous neurophysiological findings in motor learning needs to be taken with caution as the aforementioned studies mainly used motor sequence learning tasks that may engage partially different brain mechanisms than our motor skill learning task (Krakauer et al., 2019). Altogether, these elements suggest that the different learning dynamics observed in individuals training with short and long reward delays could result from the preferential engagement of distinct brain networks that exhibit different activation patterns during motor learning.

The impairment of motor consolidation observed in subjects who trained with a long reward delay also suggests that reward timing does not only affect the acquisition of the skill, but also the offline processing of the acquired motor memory. The reduction of overnight consolidation in learners of Group_{Long} may appear discordant with previous work showing improved episodic memory consolidation after training with long reward delays (Foerde and Shohamy, 2011). Notably though, the beneficial effect of long reward delays on episodic memory previously reported was not observed in Parkinson's disease patients nor in their age-matched controls (Foerde et al., 2012). Our results may also seem to differ from those of former motor learning studies showing consolidation improvements in hippocampal-related skills (Albouy et al., 2008, 2015). However, an important difference with respect to these studies is the nature of our task. As such, the hippocampus is known to be involved to various degrees in motor learning depending on the type of skill that is practiced

(McDoughle et al., 2022), contributing more to learning in settings requiring to build a spatial representation of the task (Albouy et al., 2015) or to learn a perceptual component (Rose et al., 2011). Hence, while the hippocampus is potentially involved in skill learning tasks involving the flexible selection of force parameters (i.e., as in the current study; (McDoughle et al., 2022)), its engagement may have been limited as learning did not involve a strong spatial or perceptual component. Another complementary interpretation is that rewards delivered after a long delay are temporally discounted and perceived as subjectively less valuable relative to when the delay is short (Shadmehr et al., 2010a, 2019), reducing their beneficial effect on offline consolidation mechanisms (Ambrose et al., 2016; Sterpenich et al., 2021).

Beyond reward timing, another feature that could have altered both the learning dynamics and consolidation in the present study is the post-reward delay – *i.e.*, the delay between reward delivery and the execution of the subsequent movement (referred to as ITI in the *Results* section, above). First, the comparison of Group_{Short} and Group_{Short-Replication} suggests that lengthening the post-reward delay had a rather negative impact on the learning dynamics, inducing a reduction in learning rates (**Figure S2**). Despite this detrimental impact, participants of Group_{Short} (ITI = 6 s) still exhibited better learning rates than participants of Group_{Long} (ITI = 1 s), suggesting that the positive effect of the shorter reward delays overcame the negative impact of the longer ITI in Group_{Short}. Overall, this analysis suggests that both reward delay and ITI duration influence motor learning but that reward delay plays a more prominent role in shaping learning. Second, the presence of resting periods of a few seconds during learning was recently shown to induce a rapid form of consolidation during motor sequence learning (Bönstrup et al., 2019, 2020; Jacobacci et al., 2020; Buch et al., 2021). We cannot rule out that the longer ITI experienced by Group_{Short} could have facilitated this form of consolidation. Notably though, this rapid form of consolidation was not correlated to overnight consolidation, suggesting different mechanisms for between-trials and between-days consolidation (Bönstrup et al., 2019). Hence, we believe it is unlikely that the longer ITIs in Group_{Short} drove the effect of reward timing on overnight consolidation.

In conclusion, our data indicate that the timing at which reward is delivered during motor training alters the dynamics of learning and the consolidation of the new motor

memory. Research is now required to gain further knowledge as to the brain networks involved in these time-dependent effects of reward on motor learning. Such knowledge would prove useful for the design of future reward-based rehabilitation programs, in which reward timing may be individualized depending on the brain networks and learning processes affected in specific populations of patients. For instance, short reward delays may be preferred during rehabilitation when brain lesions affect the medial temporal lobe (Foerde et al., 2013), whereas long reward delays may prove more efficient when patients suffer from dysfunction of the striatal network (Foerde and Shohamy, 2011; Foerde et al., 2012; Gabay et al., 2018; Weismüller et al., 2018). In addition, our study suggests that short reward delays and short ITIs should be generally preferred in motor rehabilitation when the motor deficit is not associated to any lesion of the reward circuitry, as occurs after spinal cord injury or lesions of the peripheral nervous system.

Methodological considerations and limitations of the study

Even if initial performance was not significantly different between the groups in any analysis, the fact that it was slightly lower in Group_{Long} may have caused an overestimation of early learning rates in this group. In this case, higher early learning rates in Group_{Long} (relative to Group_{Short}) would reflect a quick recovery from an initial perturbation caused by the introduction of long reward delays at T1. The present data do not allow us to rule out this interpretation completely. Notably though, while a decrement in initial performance in Group_{Long} may have contributed to bias our estimation of early learning rates, it cannot explain the between-group differences observed when considering the late phase of training, strongly suggestive of an effect of reward timing on learning dynamics.

Relatedly, the nature of our research question required us to employ different timings in the Training and in the Test blocks. As such, in Reward_{Short} blocks, reward delay (1 s) was identical to the Test blocks but the ITI (6 s) was different. Conversely, in Reward_{Long} blocks, ITI duration was identical to the Test blocks (1 s) but the reward delay was different (6 s). Therefore, strictly speaking, the overall similarity between the training and the Test blocks was identical in both groups. However, our results suggest that changes in reward delay has a stronger impact on motor performance than changes in ITIs, implying that performance in the Test blocks may be more affected in Group_{Long} due to the difference in the reward delay

experienced during training versus during the Test block. One may hypothesize that this could have subsequently altered performance on Day 2, which was reduced in learners of Group_{Long}. Even if we cannot definitely refute or confirm this hypothesis, we believe that it is unlikely. First, if this was true, Group_{Long} should be more disturbed than Group_{Short} when transitioning from the end of the training phase to the Post-training Test block. Importantly though, we observed the opposite pattern of results, with a tendency to improve performance from training to Post-training. Second, analysis of consolidation showed a reduction of performance on Re-test (on Day 2) compared to Post-training in learners of Group_{Long}, with both assessments being Test blocks. Any disturbance of Group_{Long} subjects due to the difference between the reward delay experienced during training and Test blocks should have affected similarly both Post-training and Re-test blocks. Overall, characterizing the impact of dynamic changes in reward delay on motor performance represents an interesting avenue for future research.

3.5. Materials and Methods

3.5.1. Resources table

RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Motor learning data ('All_Var_table.mat')	This paper	https://osf.io/4kqpe/
Subjects characteristics ('Subjects_characteristics_Timing.xlsx')	This paper	https://osf.io/4kqpe/
Software and Algorithms		
Matlab vR2007 7.5 and R2018a	Mathworks	www.mathworks.com/products/matlab.html
Statistica 10	StatSoft Inc.	https://www.statistica.com/en/
Psychophysics Toolbox	Psych-toolbox.org	http://psychtoolbox.org/

3.5.2. Participants

A total of sixty right-handed healthy volunteers participated in the present study (46 women, 23.7 ± 0.3 years old; mean \pm SE). Data from a previous group of thirty participants was also re-analyzed (20 women, 23.9 ± 0.43 years old; (Vassiliadis et al., 2021)). Handedness was determined via a shortened version of the Edinburgh Handedness inventory (Oldfield, 1971). None of the participants suffered from any neurological or psychiatric disorder, nor were they taking any centrally-acting medication. All participants gave their written informed consent in accordance with the Ethics Committee of the Université Catholique de Louvain (approval number: 2018/22MAI/219) and the principles of the Declaration of Helsinki. Subjects were financially compensated for their participation. Finally, all participants were asked to fill out a French adaptation of the Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ; (Torrubia et al., 2001; Lardi et al., 2008)) and a NASA Task Load Index questionnaire (NASA-TLX, (Hart and Staveland, 1988)).

3.5.3. Methods details

3.5.3.1. Motor skill learning task

Participants were seated approximately 60 cm in front of a cathode-ray tube screen (refresh rate: 100 Hz) with their right forearm positioned at a right angle on the table. The task was developed on Matlab 7.5 (the Mathworks, Natick, Massachusetts, USA) exploiting the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) and consisted in a previously described force modulation task (Vassiliadis et al., 2021). Briefly, the task required participants to squeeze a force transducer (Arsalis, Belgium) between the index and the thumb to control a cursor displayed on the screen. Increasing the force exerted resulted in the cursor moving vertically and upward. Each trial started with a preparatory period in which a sidebar appeared at the bottom of the screen and a target at the top (**Figure 1A**). After a variable time interval (0.8 to 1 s), a cursor popped up in the sidebar, indicating the start of the movement period. Participants had to pinch the transducer to move the cursor as quickly as possible from the sidebar to the target and maintain it there for the rest of the movement period, which lasted 2 s. The level of force required to reach the target ($\text{Target}_{\text{Force}}$) was individualized for each participant and set at 10 % of maximum voluntary contraction (MVC). Notably, squeezing the transducer before the appearance of the cursor was considered as an anticipation and therefore led to the interruption of the trial. Anticipation trials were discarded from further analyses. At the end of each trial, a binary reinforcement feedback was presented to the subject (yellow or blue circle for success or failure, respectively).

3.5.3.2. Sensory and reinforcement feedbacks

We provided only limited visual feedback to the participants in order to increase the impact of the reinforcement feedback on learning (Mawase et al., 2017). As such, on 90 % of the trials, the cursor disappeared shortly after the start of the movement period: it became invisible as soon as the generated force became larger than half of the $\text{Target}_{\text{Force}}$ (*i.e.*, 5 % of MVC). Conversely, the remaining trials (10 % of the trials) provided a continuous vision of the cursor (full vision trials). Full vision trials were not considered in the analyses.

As mentioned above, each trial ended with the presentation of a binary reinforcement feedback, indicating success or failure. Success on the task was determined based on the Error, defined as the absolute force difference between the $\text{Target}_{\text{Force}}$ and the exerted force

(Abe et al., 2011; Steel et al., 2016). The Error was first computed for each frame refresh from 0.15 s to the end of the trial (*i.e.*, providing 185 data points at 100 Hz), then averaged across the data points for each trial (Steel et al., 2016), and expressed in percentage of MVC. This indicator of performance allowed us to classify a trial as successful or not based on an individualized success threshold (see below). When the Error on a given trial was below the threshold, the trial was classified as successful, and when it was above the threshold, the trial was considered as failed. Hence, task success depended on the ability to approximate the Target_{Force} as quickly and as accurately as possible.

3.5.3.3. Reward timing manipulation

The protocol involved Training and Test blocks (see *Experimental protocol*, below). During Training blocks, reinforcement feedbacks were associated with a reward of 8 cents on successful trials, and failed trials led to 0 cent. Importantly, in two block types, we manipulated the timing at which the reinforcement feedback, and therefore the associated reward, was delivered after the movement period (**Figure 1A**). Indeed, the reward was displayed after either a short or a long delay – that is, 1 or 6 s following the movement period in Reward_{Short} and Reward_{Long} blocks, respectively (see (Foerde and Shohamy, 2011; Foerde et al., 2013) for the use of similar delays in decision-making tasks). In order to keep the total duration of the trial constant in these two block types, inter-trial intervals (ITI, which followed reward occurrence) were set to 6 and 1 s in the Reward_{Short} and the Reward_{Long} blocks, respectively. Finally, we re-analyzed data from a previous study (Vassiliadis et al., 2021), in which the training blocks involved a short reward delay timing (0.5 s) and an intermediate ITI (3 s; Reward_{Short-PastStudy} blocks). The latter analysis allowed us to test for the reproducibility of the effects of training obtained in the Reward_{Short} block.

In the Test blocks, reinforcement feedback occurred 1 s after the movement period, involved an ITI of 1 s, and was not associated with any reward.

3.5.3.4. Motor skill learning protocol

Subjects were tested on two consecutive days (Day 1 and Day 2; **Figure 1C**). On Day 1, we first measured the individual MVC to calculate the Target_{Force}. Notably, MVCs and simple reaction times (SRT) were measured before and after the training blocks to assess

potential fatigue related to the training (see Data and statistical analyses). Participants then performed 2 blocks of Familiarization, in a Test block setting. The first Familiarization block comprised 20 full vision trials. Subsequently, all blocks were composed of a mixture of partial vision trials (90 % of total trials) and full vision trials (10 % of total trials), as described above. The second Familiarization block involved 40 trials and allowed us to determine baseline performance to calibrate the difficulty of the task for the rest of the experiment (Calibration block; please see (Vassiliadis et al., 2021) for details on the Calibration procedure).

Following Familiarization, participants performed 320 trials divided in 8 blocks. All subjects started and ended the session with the realization of a Test block of 40 trials, allowing us to evaluate initial performance and total learning (*i.e.*, Pre- and Post-training blocks, respectively). In between, 6 Training blocks (T1 to T6) of 40 trials were performed by the participants (see **Figure 1B**). During the Training blocks, individuals were split into 2 separate groups depending on the type of training blocks they performed. As such, Group_{Short} and Group_{Long} trained with Reward_{Short} and Reward_{Long} blocks, respectively. The group trained with Reward_{Short-PastStudy} blocks was referred to as Group_{Short-PastStudy}. Comparing performance between the groups during the training period allowed us to test the effect of reward timing on the learning dynamics.

Day 2 was realized 24 hours later. Subjects performed the task again with the same Target_{Force} and success threshold. This assessment was composed of 5 full vision trials followed by a Test block of 40 trials (Re-test) and allowed us to assess the effect of reward timing on skill consolidation.

3.5.4. Data and statistical analyses

Statistical analyses were carried out with Matlab 2018a (the Mathworks, Natick, Massachusetts, USA) and Statistica 10 (StatSoft Inc., Tulsa, Oklahoma, USA). In the case of independent samples t-tests we verified the homogeneity of the variances systematically and non-parametric tests were used when variances were non-homogeneous. Linear mixed models (LMM) were fitted using the *fitlme* function in Matlab, with the restricted maximum likelihood fitting method. As random effects, we added intercepts for participants. Normality of residuals, skewness and homoscedasticity of the data were systematically tested and logarithmic transformations were applied when necessary. Significance of fixed effects was

tested by conducting ANOVAs on the models' coefficients (with Satterthwaite approximation of the degrees of freedom) with the function *anova* and post-hoc comparisons were conducted using the *coefTest* function. The significance level was set at $p \leq 0.05$, except in the case of correction for multiple comparisons (see below).

3.5.4.1. Motor skill learning

As a first step, we tested the impact of reward timing on motor performance during each block of Test and Training block. We quantified for each subject the percentage of successful trials (*i.e.*, the success rate) for each block and then normalized the data according to individuals' initial performance by subtracting the success rate values measured at Pre-training from the values obtained in every block. To evaluate the impact of reward timing on success rates across training, we performed a LMM with the categorical fixed effects $\text{GROUP}_{\text{TYPE}}$ ($\text{Group}_{\text{Short}}$ and $\text{Group}_{\text{Long}}$, $n=30$ each) and $\text{TRAINING}_{\text{BLOCK}}$ (T1 to T6). In order to confirm these results using single-trial data, we used the Error allowing us to obtain a continuous variable at each trial. Notably, for each participant, Errors measured during training were expressed in percentage of the average Pre-training level. In this case, we ran a LMM with the categorical fixed effect $\text{GROUP}_{\text{TYPE}}$ ($\text{Group}_{\text{Short}}$ and $\text{Group}_{\text{Long}}$) and the continuous fixed effect $\text{TRAINING}_{\text{TRIAL}}$ (trial 1 to 240). When the analysis revealed a significant interaction, we then compared the coefficient associated to $\text{TRAINING}_{\text{TRIAL}}$ to evaluate potential between-group differences in learning rates. Then, to characterize the effect of the ITI's duration on motor learning, we replicated these analyses with the inclusion of the $\text{Group}_{\text{Short-PastStudy}}$.

As a second step, we aimed at evaluating the effect of reward timing on the dynamics of the learning process. To do so, we ran the same LMM as described above with the addition of the fixed effect $\text{TRAINING}_{\text{PHASE}}$ which was modeled as a categorical fixed effect with two modalities ($\text{Training}_{\text{Early}}$ or $\text{Training}_{\text{Late}}$ for the first and last 120 trials or training, respectively). We were especially interested in a potential triple $\text{TRAINING}_{\text{TRIAL}} \times \text{GROUP}_{\text{TYPE}} \times \text{TRAINING}_{\text{PHASE}}$ interaction which would indicate that learning rates varied not only depending on the group but also depending on the phase of practice.

As a supplementary analysis to support our differences of learning dynamics between the groups, we also ran regression analysis for each subject on binned Success rates (presented in Figure S3). Specifically, we split the data into 24 non-overlapping bins of 10 trials, computed the success rate for each bin and normalized the data according to individuals' initial performance, as done in the first analysis. The bins were then separated into two equal parts (*i.e.*, of 12 bins each) depending on whether they belonged to the early or to the late phase of training (Training_{Early} and Training_{Late} phases, corresponding to T1-T3 and T4-T6, respectively). Finally, we performed linear regressions on these data and extracted the slope of the fits for the Training_{Early} and the Training_{Late} phases of the Group_{Short} and the Group_{Long} ($n=30$ each). The slope values – exploited here as a proxy of the learning rate – were compared using a two-way ANOVA with GROUP_{TYPE} (Group_{Short} and Group_{Long}) and TRAINING_{PHASE} (Training_{Early} and Training_{Late}) as between- and within-subjects factors, respectively.

Finally, we tested for any effect of reward timing on total learning, by comparing the success rates of Group_{Short} and Group_{Long} at Post-training, using an independent sample t-test. Further, in order to test the statistical significance of total learning within each group, we conducted two single sample t-tests on Post-training success rate, against a constant value of 0 (threshold for significance Bonferroni-corrected at $p \leq 0.025$).

3.5.4.2. Motor skill consolidation

A secondary goal of the study was to evaluate the effect of reward timing on skill consolidation. We first performed this analysis on the whole cohort ($n = 30$ per group). However, a potential caveat of these analyses is that they included participants who did not learn the task on Day 1 and even exhibited a deterioration of performance with practice on Day 1. In these participants, a Re-test performance (*i.e.*, on Day 2) similar to the Pre-training level would be considered as evidence for an offline stabilization or even gain in performance, when it would actually only reflect a return to the baseline level of performance. In a second step, we therefore focused only on participants who demonstrated skill learning on Day 1 ($\text{Success}_{\text{Post-training}} - \text{Success}_{\text{Pre-training}} > 0$). This allowed us to compare offline consolidation in participants who responded to the training on Day 1 and who also happened to

have very close Post-training success rates (**Figure 4A**), a crucial aspect in order to interpret any overnight change in performance. 40 participants were considered in this analysis (22 and 18 in Group_{Short} and Group_{Long}, respectively). Pre-training normalized Success rates (averaged per block) and Error (single-trial) data were analyzed by means of LMMs with GROUP_{TYPE} (Group_{Short} and Group_{Long}) and TEST_{BLOCK} (Post-training and Day 2) as categorical fixed effects.

3.5.4.3. Group features, initial performance and fatigue

As a control, we verified that the Group_{Short} and the Group_{Long} were comparable in terms of age, success threshold, Target_{Force}, sensitivity to reward and to punishment (*i.e.*, as assessed by the SPSRQ questionnaire), initial performance (*i.e.*, at Pre-training) and received monetary gains. As displayed in **Table 1**, independent sample two-tailed t-tests performed on these data did not reveal any significant differences between the groups (see also **Figure 1C**).

We also assessed if potential motor and cognitive fatigue generated by Day 1 training was different between the groups (Derosiere and Perrey, 2012; Derosière et al., 2014). To do so, we expressed MVCs, and SRTs obtained after training (MVC_{POST} and SRT_{POST}) in percentage of the values measured initially (MVC_{PRE} and SRT_{PRE}). We also assessed the perceived workload after training through the NASA-TLX questionnaire. Notably, these data did not differ between the groups (**Table 1**), suggesting that motor and cognitive fatigue were not responsible for the effect of reward timing on motor learning.

	Group _{Short} (n = 30)	Group _{Long} (n = 30)	t-value	p-value
Age (in years)	22.8 ± 0.58	23.0 ± 0.52	-0.26	0.80
Gender (number of females)	22	24	/	/
Success Threshold (% MVC)	2.7 ± 0.01	2.7 ± 0.01	0.18	0.86
Target _{Force} (Newtons)	4.74 ± 0.21	4.39 ± 0.17	1.32	0.19

Sensitivity to reward and punishment (score)	82.0 ± 2.32	83.3 ± 2.04	-0.43	0.67
Pre-training success rate	31.0 ± 2.61	34.9 ± 3.62	-0.86	0.39
Monetary Gains (euros)	39.0 ± 0.64	38.5 ± 0.73	0.45	0.66
Muscle fatigue (MVC _{POST} in % of MVC _{PRE})	91.3 ± 2.83	93.7 ± 2.67	-0.62	0.54
Simple Reaction Time change (SRT _{POST} in % of SRT _{PRE})	104.35 ± 2.51	103.21 ± 2.32	0.33	0.74
Perceived workload (NASA-TLX score)	49.4 ± 2.74	50.89 ± 2.79	-0.39	0.70

Table 1. Group features, initial performance and fatigue in the three experimental groups (mean ± SE). The 2 last columns provide the results of independent samples t-tests.

3.6. Supplementary materials

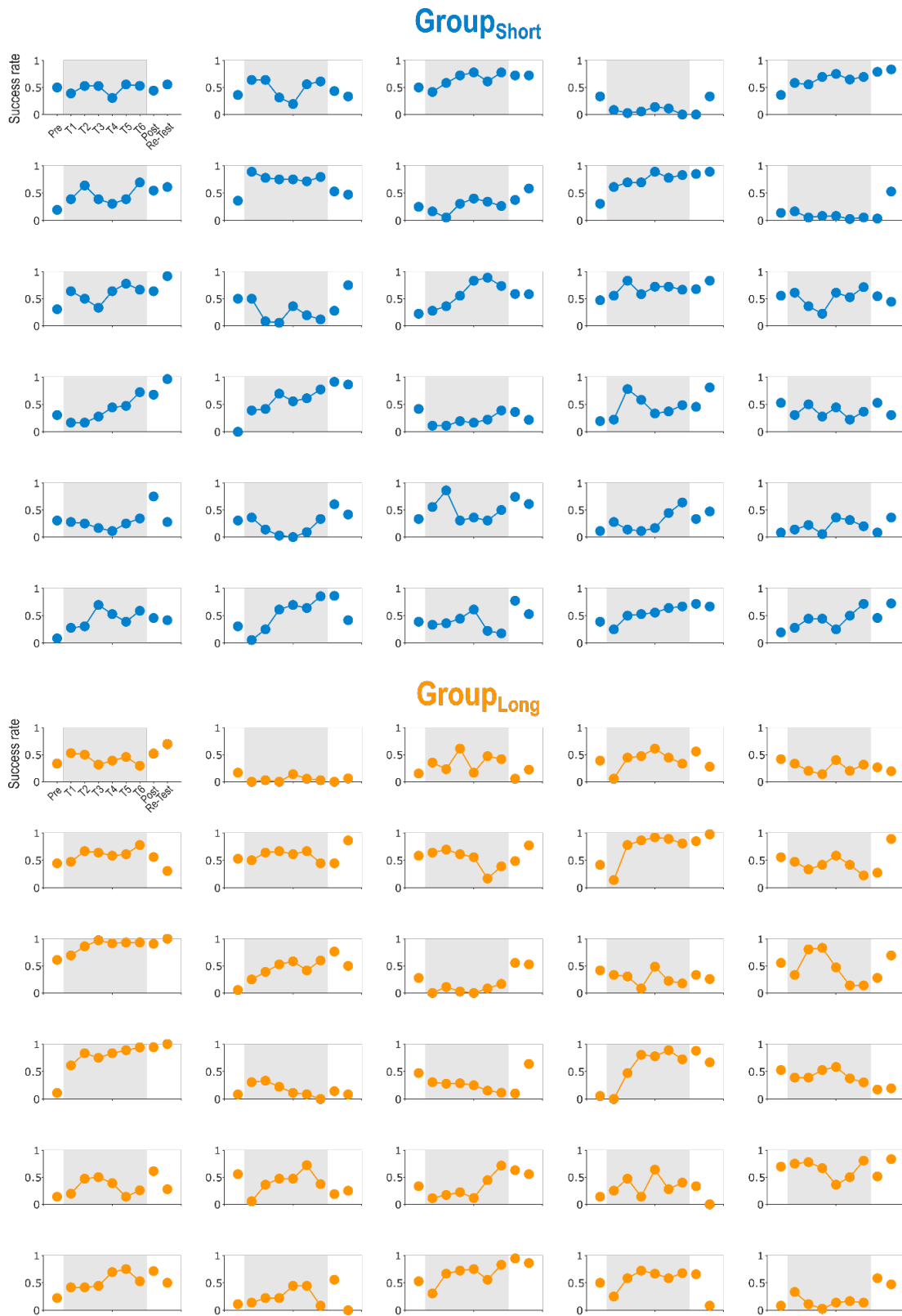


Figure S1. Individual learning curves. Success rates are displayed for each block of practice and each participant of Group_{Short} (n = 30; upper panel, blue traces) and Group_{Long} (n = 30; lower panel, orange traces).

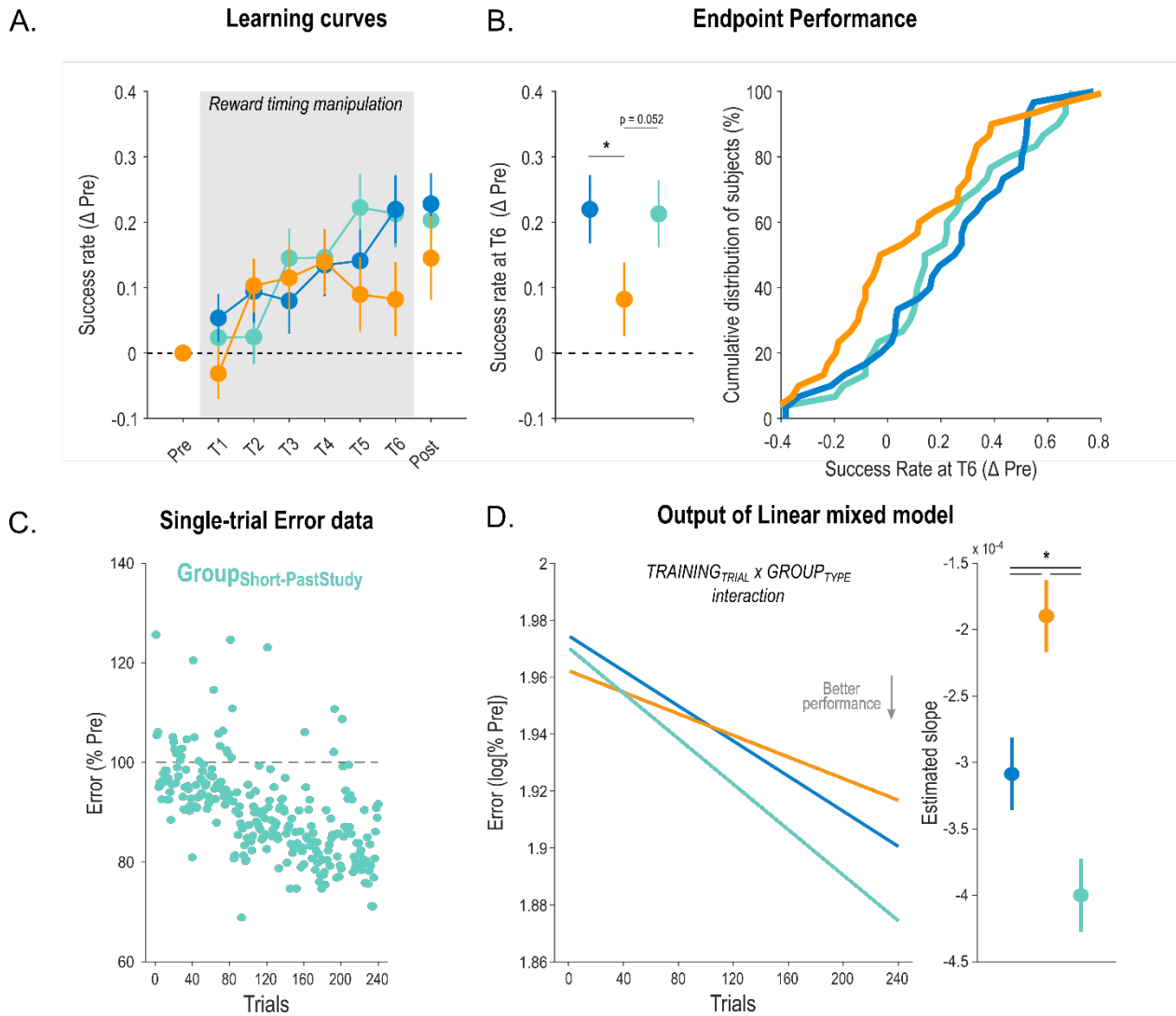


Figure S2. Effect of reward timing on motor skill learning with the inclusion of Group_{PastStudy} (n = 30). **A. Learning curves.** Proportion of successful trials (expressed as a difference with the individual Pre-training success rate) is represented across practice for the three experimental groups (blue: Group_{Short}; n = 30, orange: Group_{Long}; n = 30, turquoise: Group_{Short-PastStudy}; n = 30). The grey shaded area highlights the blocks concerned by the reward timing manipulation. The remaining blocks were Test blocks. Note the similarity in the learning curves of Group_{Short} and Group_{Short-PastStudy} despite the difference in ITI. **B. Endpoint performance.** Average success rates in the end of the training period (i.e., measured at T6) in the three groups (left panel) and the corresponding cumulative distributions of the data (right panel). **C. Single-trial Error data.** Normalized Error data obtained during training are averaged for Group_{Short-PastStudy} and plotted for each single trial. Note that lower Errors were associated with better performance. **D. Output of LMM on the Error data.** Output of LMM run on the log-transformed Error data is plotted for each group (left panel). The Error data was log-transformed prior to respect key assumptions of LMMs (see STAR Methods section). The significant TRAINING_{TRIAL} \times GROUP_{TYPE} interaction shows that the slope of learning was steeper in

Group_{Short-PastStudy} than in Group_{Long} ($p < 0.001$) and also than in Group_{Short} ($p = 0.018$; right panel). As in the main analysis, estimated intercepts were not different between groups ($p = 0.59$). Notably, more negative slopes reflect larger learning rates. *: significant difference ($p < 0.05$). Data are represented as mean \pm SE.

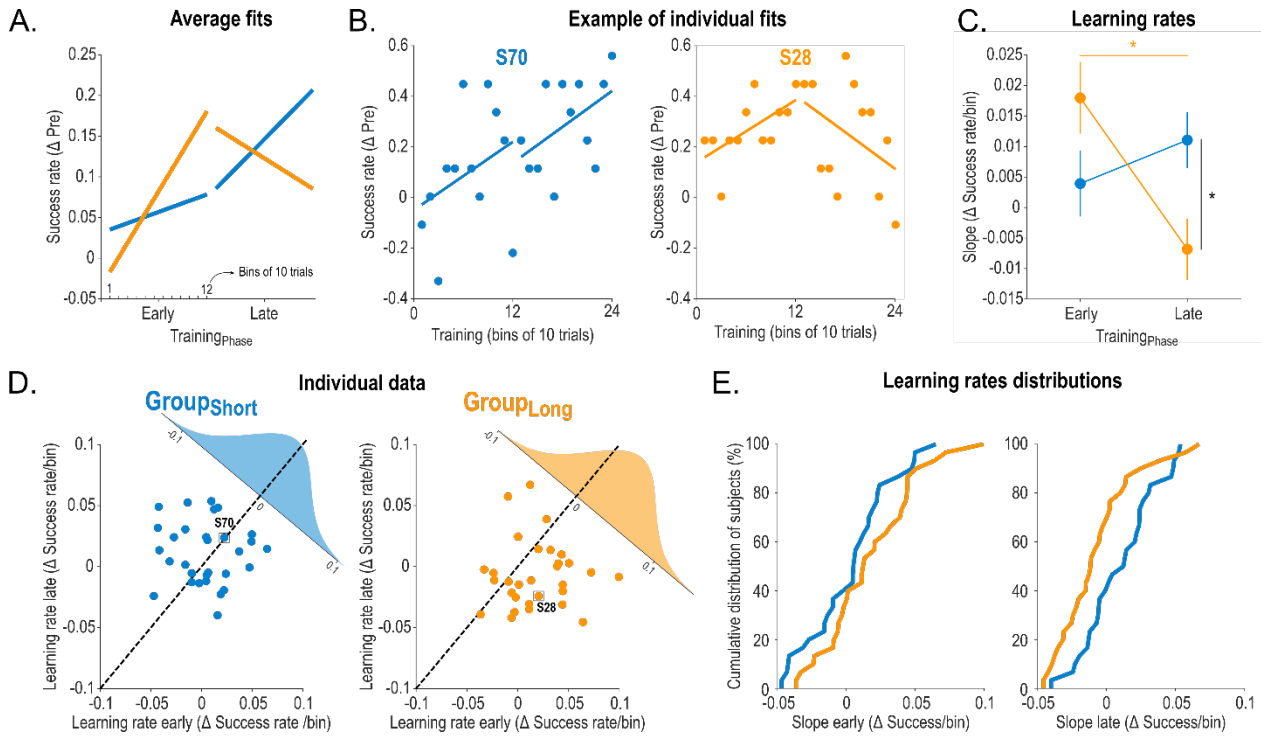


Figure S3. Additional analysis showing the effect of reward timing on learning rates for early and late phases of practice. **A. Average fits.** Average fits for GroupShort (blue) and GroupLong (orange) obtained through linear regression on each subject's binned success data (expressed as a difference from Pre-Training) at early (corresponding to T1, T2 and T3) and late phases of training (corresponding to T4, T5 and T6). Each bin was the average of 10 trials and there were 12 bins per TrainingPhase. **B. Example of individual fits.** Binned success rates and the corresponding linear fits are represented for an exemplar subject of GroupShort (left) and GroupLong (right). **C. Learning rates.** The slope of the individual fits – expressed as a delta of success rate per bin on the y-axis – was exploited as a proxy of the learning rate for early and late phases of training. A two-way ANOVA run on these data revealed a strong TRAININGPHASE \times GROUPTYPE interaction ($F_{(1, 58)} = 9.13$; $p = 0.0037$). This interaction was driven by the fact that learning rates were comparable across training phases in GroupShort ($p = 0.34$), while they were significantly lower at TrainingLate than at TrainingEarly in GroupLong ($p = 0.0016$; orange star). Consistently, learning rates were higher in GroupShort than in GroupLong at TrainingLate ($p = 0.017$; black star). This tended to be the opposite at TrainingEarly ($p = 0.061$). Notably, intercepts of the linear fits at TrainingEarly were not significantly different between the groups ($t_{(58)} = 1.07$, $p = 0.29$). **D. Individual learning rates.** Scatter plot representing each subject's learning rate for early (x-axis) vs late (y-axis) TrainingPhases. The group distributions of the change in learning rates (TrainingEarly – TrainingLate) are also represented (upper right corner of each plot). Note the shift in GroupLong below the identity line reflecting the higher learning rates at TrainingEarly than at TrainingLate. This shift was also demonstrated by the larger proportion of GroupLong participants (76.7 % [23/30]) exhibiting higher learning rates in the early than in the late phase of training, compared

to Group_{Short} (46.7% [14/30]; Fisher's exact test on the proportions: $p = 0.033$). **E. Learning rates distributions.** Cumulative distributions of the group data for early (left panel) and late (right panel) learning phases. *: significant difference ($p < 0.05$). Data are represented as mean \pm SE.

Log(Error) ~ 1 + TRAINING _{TRIAL} x GROUP _{TYPE} + (1 Subject)					
AIC = -11099 ; BIC = -11054					
Fixed effects	Estimate	95% CI Lower	95% CI Upper	t-value	p-value
Intercept	1.975	1.942	2.007	120.10	< 0.001
TRAINING _{TRIAL}	-0.0003086	-0.0003653	-0.0002537	-11.01	< 0.001
GROUP _{Long}	-0.01230	-0.05789	0.03329	-0.53	0.60
TRAINING _{TRIAL} x GROUP _{Long}	0.0001189	4.124x10 ⁻⁵	0.0001966	3.00	0.0027

Supplementary Table 1. Output of LMM run on the Error data with the fixed effects TRAINING_{TRIAL} and GROUP_{TYPE} (see also Figure 2D).

Log(Error) ~ 1 + TRAINING _{TRIAL} x GROUP _{TYPE} x TRAINING _{PHASE} + (1 Subject)					
AIC = -11099 ; BIC = -11025					
Fixed effects	Estimate	95% CI Lower	95% CI Upper	t-value	p-value
Intercept	1.967	1.934	2.000	116.44	< 0.001
TRAINING _{TRIAL}	-0.0001749	-0.0003300	-0.00001999	-2.21	0.027
GROUP _{Long}	0.02168	-0.02514	0.06851	0.91	0.36
TRAINING _{Late}	0.02160	-0.009096	0.05230	1.38	0.17
TRAINING _{TRIAL} x GROUP _{Long}	-0.0004730	-0.0006913	-0.0002548	-4.25	< 0.001
TRAINING _{TRIAL} x TRAINING _{Late}	-0.0002147	-0.0004348	5.3346 x10 ⁻⁶	-1.91	0.056
GROUP _{Long} x TRAINING _{Late}	-0.1072	-0.15048	-0.06399	-4.86	< 0.001
TRAINING _{TRIAL} x GROUP _{Long} x TRAINING _{Late}	0.001008	0.0006982	0.001318	6.37	< 0.001

Supplementary Table 2. Output of LMM run on the Error data with the fixed effects TRAINING_{TRIAL}, GROUP_{TYPE} and TRAINING_{PHASE} (see also Figure 3C).

4. Study 3: Non-invasive stimulation of the human striatum disrupts reinforcement learning of motor skills

Pierre Vassiliadis^{1,2,3}, Elena Beanato^{1,2}, Traian Popa^{1,2}, Fabienne Windel^{1,2}, Takuya Morishita^{1,2}, Esra Neufeld⁴, Julie Duque³, Gerard Derosiere³, Maximilian J. Wessel^{1,2,5} and Friedhelm C. Hummel^{1,2,6 *}

¹Defitech Chair of Clinical Neuroengineering, Neuro-X Institute (INX) and Brain Mind Institute (BMI), École Polytechnique Fédérale de Lausanne (EPFL), 1202 Geneva, Switzerland.

²Defitech Chair of Clinical Neuroengineering, INX and BMI, EPFL Valais, Clinique Romande de Réadaptation, 1951 Sion, Switzerland.

³Institute of Neuroscience, Université Catholique de Louvain, 1200, Brussels, Belgium

⁴Foundation for Research on Information Technologies in Society, Zurich, Switzerland

⁵Department of Neurology, University Hospital, Julius-Maximilians-University, Würzburg, Germany

⁶Clinical Neuroscience, University of Geneva Medical School, 1202 Geneva, Switzerland.

Reference: Vassiliadis P, Beanato E, Popa T, Windel F, Morishita T, Neufeld E, Duque J, Derosiere G, Wessel MJ, Hummel FC (2022). Non-invasive stimulation of the human striatum disrupts reinforcement learning of motor skills.

Personal contribution: Conceptualization, Methodology, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Funding acquisition.

* Correspondence

Friedhelm Hummel

Neuro-X Institute (INX)

École Polytechnique Fédérale de Lausanne (EPFL)

Campus Biotech, Room H4.3.132.084

Chemin des Mines 9, 1202 Geneva, Switzerland

and

EPFL Valais,

Clinique Romande de Réadaptation

Av. Grand-Champsec 90,

CH-1951 Sion

Email: friedhelm.hummel@epfl.ch

4.1. Abstract

Reinforcement feedback can improve motor learning, but the underlying brain mechanisms remain unexplored. Especially, the causal contribution of specific patterns of oscillatory activity within the human striatum is unknown. To address this question, we exploited an innovative, non-invasive deep brain stimulation technique called transcranial temporal interference stimulation (tTIS) during reinforcement motor learning with concurrent neuroimaging, in a randomised, sham-controlled, double-blind study. Striatal tTIS applied at 80Hz, but not at 20Hz, abolished the benefits of reinforcement on motor learning. This effect was related to a selective modulation of neural activity within the striatum. Moreover, 80Hz, but not 20Hz, tTIS increased the neuromodulatory influence of the striatum on frontal areas involved in reinforcement motor learning. These results show for the first time that tTIS can non-invasively and selectively modulate a striatal mechanism involved in reinforcement learning, opening new horizons for the study of causal relationships between deep brain structures and human behaviour.

Keywords:

Motor learning, reward, reinforcement learning, non-invasive brain stimulation, deep brain stimulation, temporal interference stimulation, striatum, neuroimaging

4.2. Introduction

The ability to learn from past outcomes, often referred to as reinforcement learning, is fundamental for biological systems (Neftci and Averbeck, 2019). Reinforcement learning has been classically studied in the context of decision making, when agents have to decide between a discrete number of potential options (Schultz, 2015). Importantly, there is an increasing recognition that reinforcement learning processes are also at play in other contexts including when one has to learn a new motor skill (Dhawale et al., 2017; Vassiliadis et al., 2021). For instance, the addition of reinforcement feedback during motor training can improve motor learning, presumably by boosting the retention of newly acquired motor memories (Huang et al., 2011; Galea et al., 2015a). Interestingly, reinforcement feedback also appears to be relevant for the rehabilitation of patients suffering from motor impairments (Therrien et al., 2016; Vassiliadis et al., 2019; Widmer et al., 2021). Yet, despite these promising results, there is currently a lack of understanding of the brain mechanisms that are critical to implement this behaviour.

A prominent hypothesis in the field is that the striatum, an area that is active both during reinforcement (Bartra et al., 2013) and motor learning (Hardwick et al., 2013), may be causally involved in the beneficial effects of reinforcement on motor learning. As such, the striatum shares dense connexions with dopaminergic structures of the midbrain as well as with pre-frontal and motor cortical regions (Haber, 2016), and is therefore well positioned to translate information about reinforcement into motor adjustments (Balleine et al., 2007; Piray et al., 2017; Hori et al., 2019). This idea is in line with neuroimaging studies showing reward-related activation of the striatum during motor learning (Wachter et al., 2009; Widmer et al., 2016). More specifically, within the striatum, oscillatory activity in specific frequency bands is suggested to be involved in aspects of reinforcement processing. As such, previous

rodent studies have shown that striatal high gamma oscillations (~ 80 Hz) are sensitive to reward and dopamine and are highly coherent with the frontal cortex, suggesting that they may be involved in reinforcement learning (Berke, 2009; van der Meer and Redish, 2009; van der Meer et al., 2010; Matsumoto et al., 2012; Dwiell et al., 2019). In particular, dynamic changes of high gamma activity in the striatum (Berke, 2009; Cohen et al., 2009; Kalenscher et al., 2010) and in other parts of the basal ganglia (Herrojo-Ruiz et al., 2014; Sepe-Forrest et al., 2021) seem to encode the outcome of previous movements (i.e., success or failure). Hence, this body of work suggests that the fine-tuning of striatal oscillatory activity, especially in the gamma range, may be crucial for reinforcement learning of motor skills. Conversely, striatal beta oscillations (~ 20 Hz) have been largely associated with sensorimotor functions (Jenkinson and Brown, 2011). For instance, beta oscillations in the striatum are exacerbated in Parkinson's disease and associated to the severity of motor symptoms (Brown, 2007; McCarthy et al., 2011; Kondabolu et al., 2016). Taken together, these elements suggest that striatal high gamma and beta activity may have different functional roles preferentially associated to reinforcement learning and sensorimotor functions, respectively.

The studies mentioned above provide associative evidence linking the presence of reinforcement with changes of neural activity within the striatum determined through neuroimaging (Wachter et al., 2009; Widmer et al., 2016), but do not allow to draw conclusions regarding its causal role in reinforcement motor learning in humans. The only causal evidence available to date comes from animal work showing modulation of reinforcement-based decision-making with striatal stimulation (Nakamura and Hikosaka, 2006; Williams and Eskandar, 2006). A reason for the current absence of investigations of the causal role of the striatum in human behaviour is related to its deep localization in the brain. As such, current non-invasive brain stimulation techniques, such as transcranial magnetic stimulation

(TMS) or classical transcranial electric stimulation (tES), do not allow to selectively target deep brain regions, because these techniques exhibit a steep depth-focality trade-off (Deng et al., 2013). Studies of patients with lesions of the striatum (Schmidt et al., 2008; Nickchen et al., 2017) or invasive deep brain stimulation of connected nuclei (Seymour et al., 2016; Atkinson-Clement et al., 2019) have provided insights into the role of the basal ganglia in reinforcement learning. However, their conclusions are partially limited by the fact that the studied patients also have altered network properties resulting from neurodegeneration, lesions or respective compensatory mechanisms and therefore do not allow to conclude comprehensively regarding the role of the striatum in the physiological state. Here, we address these challenges by exploiting transcranial electric Temporal Interference Stimulation (tTIS), a new non-invasive brain stimulation approach allowing to target deep brain regions in a frequency-specific and focal manner (Grossman et al., 2017; Song et al., 2021).

The concept of tTIS was initially proposed and validated on the hippocampus of rodents (Grossman et al., 2017) and was then further tested through computational modelling (Rampersad et al., 2019; Cao et al., 2020; Mirzakhilili et al., 2020; Esmaeilpour et al., 2021; von Conta et al., 2021). tTIS requires two pairs of electrodes to be placed on the head, each pair delivering a high frequency alternating current in phase opposition. One key element is that this frequency has to be high enough (i.e., in the kHz range) to avoid direct neuronal entrainment, based on the low filtering properties of neuronal membranes (Hutcheon and Yarom, 2000). The second key element consists in applying a small difference of frequency between the two alternating currents. The superposition of the electric fields creates an envelope oscillating at this low-frequency difference, which in turn is in a range able to influence neuronal activity. By optimizing the electrodes' placement and current intensity ratio across stimulation channels, it is possible to steer the maximal amplitude of the envelope

towards individual deep brain structures, whilst minimizing it in the surrounding and/or overlying areas (Acerbo et al., 2022). Another possibility offered by tTIS is to stimulate at a particular frequency of interest in order to preferentially interact with specific neuronal processes (Grossman et al., 2017; Song et al., 2021). Importantly, despite these exciting opportunities, current evidence for tTIS-related neuromodulation of deep brain structures, such as the striatum, is lacking in humans.

Here, we combine tTIS with electric field modelling for target localisation, behavioural data and functional magnetic resonance imaging (fMRI) to evaluate the causal role of specific striatal oscillations in reinforcement learning of motor skills. In particular, based on the studies mentioned above, we hypothesised that striatal tTIS at high gamma frequency (tTIS_{80Hz}) would disturb the fine-tuning of high gamma oscillatory activity in the striatum and thereby would perturb reinforcement motor learning in contrast to beta (tTIS_{20Hz}) or sham (tTIS_{Sham}) stimulation. In line with our prediction, we found that tTIS_{80Hz} disrupted motor learning compared to the controls, but only in the presence of reinforcement. These behavioural effects were associated to a specific modulation of BOLD activity in the putamen and caudate, supporting the ability of tTIS to selectively modulate striatal activity without affecting overlying cortical areas. Moreover, we found that tTIS_{80Hz} increased the neuromodulatory influence of the striatum on frontal cortical areas involved in reinforcement motor learning. Finally, we report that inter-individual variability in the neural effects of tTIS_{80Hz} can partially be explained by impulsivity levels, suggesting that responsiveness to tTIS_{80Hz} may depend on individual traits associated to reward processing. Overall, the present study shows for the first time that tTIS can non-invasively and selectively modulate a striatal mechanism involved in reinforcement learning. tTIS opens new horizons for the study of causal relationships between deep brain structures and human behaviour, and offers novel opportunities

for the treatment of neuro-psychiatric disorders associated to pathophysiological alterations of deep brain structures.

4.3. Results

24 healthy participants (15 women, 25.3 ± 0.1 years old; mean \pm SE) performed a force tracking task in the MRI with concurrent tTIS of the striatum. The task required participants to modulate the force applied on a hand grip force sensor in order to track a moving target with a cursor with the right, dominant hand (**Figure 1A**, Abe et al., 2011). At each block, participants had to learn a new pattern of motion of the target (**Figure S1**; see Methods). In Reinf_{ON} blocks, participants were provided with online reinforcement feedback during training, giving them real-time information about success or failure throughout the trial, indicated as a green or red target, respectively. The reinforcement feedback was delivered according to a closed-loop schedule, considering previous performance to update the success criterion for each sample (Therrien et al., 2016). In Reinf_{OFF} blocks, participants practiced with a visually matched random feedback (cyan/magenta). Importantly, in both types of blocks, training was performed with intermittent visual feedback of the cursor, a condition known to maximise the effect of reinforcement on motor learning (Izawa and Shadmehr, 2011; Mawase et al., 2017; Vassiliadis et al., 2021, 2022). Before and after training, participants performed Pre- and Post-training assessments with full visual feedback, no reinforcement and no tTIS, allowing us to evaluate motor learning. To evaluate the effect of tTIS on reinforcement-related benefits in motor learning and the associated neural changes, participants performed 6 blocks of trials in the MRI, with concurrent tTIS during training, delivered with a Δf of 20 Hz (tTIS_{20Hz}), 80 Hz (tTIS_{80Hz}) or as a sham (tTIS_{Sham}; 3 tTIS_{TYPE} \times 2 Reinf_{TYPE} conditions; **Figure 1B**, **1C**, see Methods for more details on the stimulation protocol).

Computational modelling of tTIS fields using realistic head models

To determine the best electrode montage to stimulate the human striatum (putamen caudate and nucleus accumbens [NAc]), computational modelling with realistic head models was conducted with Sim4Life (Iacono et al., 2015, see Methods for more details). Based on these simulations, the montage that exhibited the best exposure performance in the bilateral striatum was selected (F3-F4; TP7-TP8). Consistently, this montage generated a temporal interference electric field that was ~30-40% stronger in the striatum than in the overlying cortex, reaching magnitudes of 0.5 to 0.6 V/m (**Figure 1D, 1E**), which are compatible with entrainment of neuronal activity (Krause et al., 2019, 2022; Johnson et al., 2020).

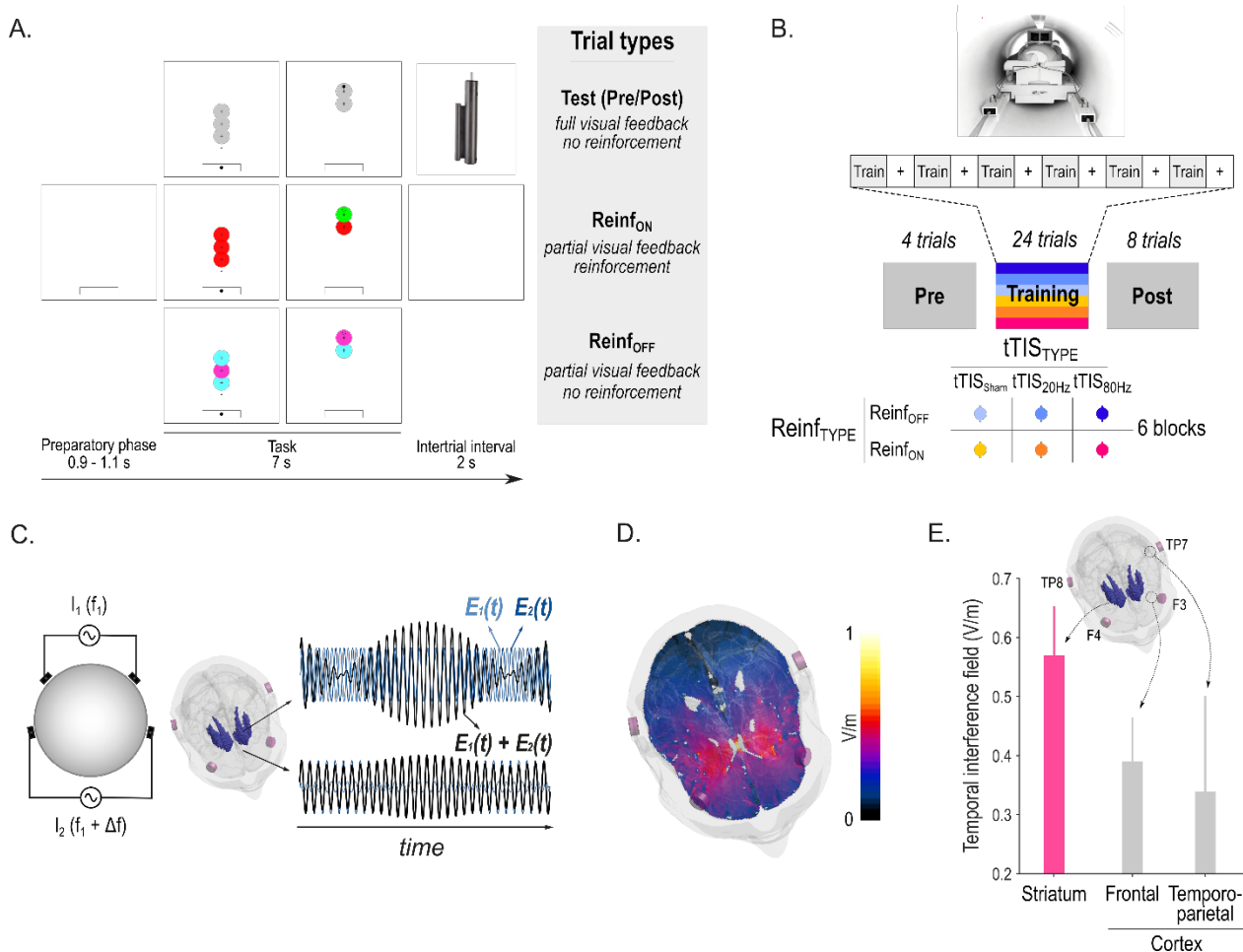


Figure 1. Striatal tTIS during reinforcement learning of motor skills in the MRI. A) Motor learning task. Participants were required to squeeze a hand grip force sensor (depicted in the upper right corner of the figure) in order to track a moving target (larger circle with a cross in the center) with a

cursor (black smaller circle; (Abe et al., 2011; Steel et al., 2016)). Pre- and Post-training assessments were performed with full visual feedback of the cursor and no reinforcement. In Reinf_{ON} and Reinf_{OFF} trials, participants practiced the task with or without reinforcement feedback, respectively. As such, in Reinf_{ON} trials, the color of the target varied in real-time as a function of the subjects' tracking performance. **B) Experimental procedure.** Participants performed the task in the MRI with concomitant TI stimulation. Blocks of training were composed of 36 trials (4 Pre-, 24 Training and 8 Post-training trials) interspersed with short resting periods (represented as + on the figure). The 6 training types resulted from the combination of 3 tTIS_{TYPE}S and 2 Reinf_{TYPE}S. **C) Concept of Temporal Interference stimulation.** On the left, two pairs of electrodes are shown on a head model and currents are applied with a frequency f_1 and $f_1 + \Delta f$. On the right, the interference of the two electric fields within the brain is plotted for two different locations with respectively high and low envelope modulation. tTIS was delivered either with a Δf of 20 or 80 Hz or as a sham (ramp-up and ramp-down of high frequency currents with flat envelope). **D) Electric field modelling with the striatal montage.** Temporal interference electric field magnitude. **E) Temporal interference electric field magnitude averaged in the striatum and in the overlying cortex.** Magnitude of the field in the cortex was extracted from the areas underneath the stimulation electrodes (F3-F4 and TP7-TP8).

tTIS_{80Hz} disrupts reinforcement-related benefits in motor learning.

Performance on the task was evaluated by means of the Error, which corresponded to the absolute difference between the applied and target force averaged across samples for each trial, as done previously ((Abe et al., 2011; Vassiliadis et al., 2021, 2022); **Figure 2A**). Across conditions, the Post-training Error was reduced compared to the Pre-training Error (single sample t-test on the normalised Post-training data: $t_{(24)} = -2.69$; $p = 0.013$; Cohen's $d = -0.55$), indicating significant motor learning during the task (**Figure 2B**). Such improvement was greater when participants had trained with reinforcement (Reinf_{TYPE} effect in the Linear Mixed Model (LMM): $F_{(1, 1062.2)} = 5.17$; $p = 0.023$; $d = -0.14$ for the contrast Reinf_{ON} – Reinf_{OFF}), confirming the beneficial effect of reinforcement on motor learning (Galea et al., 2015a; Mawase et al., 2017). Crucially though, this effect depended on the type of stimulation applied during the training (Reinf_{TYPE} x tTIS_{TYPE} interaction: $F_{(2, 1063.5)} = 2.11$; $p = 0.034$; **Figure 2C**). While reinforcement significantly improved learning when training was performed with tTIS_{Sham} ($p = 0.036$; $d = -0.22$) and tTIS_{20Hz} ($p = 0.0089$; $d = -0.27$), this was not

the case with tTIS_{80Hz} ($p = 0.43$; $d = 0.083$). Consistently, direct between-condition comparisons showed that in the Reinf_{ON} condition learning was reduced with tTIS_{80Hz} compared to tTIS_{20Hz} ($p = 0.039$; $d = 0.26$) and tTIS_{Sham} ($p < 0.001$; $d = 0.45$), but was not different between tTIS_{20Hz} and tTIS_{Sham} ($p = 0.15$; $d = 0.20$). This disruption of motor learning with tTIS_{80Hz} was not observed in the absence of reinforcement (tTIS_{80Hz} vs. tTIS_{20Hz}: $p = 0.59$; $d = -0.10$, tTIS_{80Hz} vs. tTIS_{Sham}: $p = 0.34$; $d = 0.15$). These results strongly point to the fact that high gamma striatal tTIS specifically disrupts the benefits of reinforcement on motor learning and not motor learning in general.

Although training with tTIS_{20Hz} did not alter the benefits of reinforcement on motor learning, we found that learning without reinforcement was significantly impaired in this condition (tTIS_{20Hz} vs. tTIS_{Sham}: $p = 0.046$; $d = 0.25$, **Figure 2C**). This suggests that tTIS_{20Hz} may disrupt a qualitatively different mechanism involved in learning from sensory feedback (Areshenkoff et al., 2022), in line with the role of striatal beta oscillations in sensorimotor function (Jenkinson and Brown, 2011).

As a next step, we evaluated the effect of tTIS on motor performance during training itself. As shown in Figure 2A, the Error was generally higher during Training than in Test trials due to the presence of visual uncertainty during this phase (see also Methods). The extent of this disruption was reduced in the presence of reinforcement (Reinf_{TYPE}: $F_{(1, 3262.4)} = 339.89$; $p < 0.001$; $d = -0.64$ for the contrast Reinf_{ON} – Reinf_{OFF}), demonstrating the ability of subjects to exploit real-time reinforcement information to improve tracking (**Figure 2D**). Notably, this effect was not modulated by tTIS_{TYPE} (Reinf_{TYPE} x tTIS_{TYPE}: $F_{(2, 3265.8)} = 0.91$; $p = 0.40$), indicating that tTIS did not directly influence reinforcement gains during tracking. Interestingly though, striatal stimulation did impact on general tracking performance independently of reinforcement as indicated by a significant tTIS_{TYPE} effect (tTIS_{TYPE}: $F_{(2, 3262.4)} =$

42.85; $p < 0.001$). This effect was due to an increase in the Error when tTIS_{20Hz} was applied ($p < 0.001$; $d = 0.28$ when compared to tTIS_{Sham}), which was even stronger during tTIS_{80Hz} ($p < 0.001$; $d = 0.38$ and $p = 0.031$; $d = 0.11$ when compared to tTIS_{Sham} and tTIS_{20Hz}, respectively). These results suggest that striatal tTIS altered motor performance in a frequency-dependent manner, but did not influence the ability to rapidly adjust motor commands based on reinforcement feedback during training. Hence, tTIS_{80Hz} may not disrupt real-time processing of reinforcement feedback, but may rather impair the beneficial effect of reinforcements on the retention of motor memories (Huang et al., 2011; Galea et al., 2015a).

In line with our hypothesis, the data suggest that striatal tTIS_{80Hz} disrupts a striatal mechanism involved in reinforcement learning of motor skills. Notably, this effect could not be explained by potential differences in initial performance between conditions (Reinf_{TYPE} \times tTIS_{TYPE}: $F_{(2, 519.99)} = 1.08$; $p = 0.34$), nor by changes in the flashing properties of the reinforcement feedback (i.e., the frequency of color change during tracking; Reinf_{TYPE} \times tTIS_{TYPE}: $F_{(2, 3283)} = 0.19$; $p = 0.82$), or by differences in success rate in the Reinf_{ON} blocks (i.e., the proportion of success feedback during tracking; tTIS_{TYPE}: $F_{(2, 1702)} = 0.17$; $p = 0.84$; see **Supplementary materials** for more details on these analyses).

Finally, these results can also not be a consequence of an inefficient blinding. As such, when debriefing after the experiment, only 6/24 participants were able to successfully identify the order of the stimulation applied (e.g., real – real – placebo; chance level: 4/24; Fisher exact test on proportions: $p = 0.74$). Consistently, the magnitude (**Figure S2A**) and type (**Figure S2B**) of tTIS-evoked sensations evaluated before the experiment were qualitatively similar across conditions and tTIS was generally well tolerated in all participants (no

adverse events reported). This suggests that blinding was successful and is unlikely to explain our findings. More generally, this is a first indication that tTIS evokes very limited sensations (e.g., only 2/24 and 1/24 subjects rated sensations evoked at 2 mA as “strong” for tTIS_{20Hz} and tTIS_{80Hz}, respectively; **Figure S2A**) that are compatible with efficient blinding.

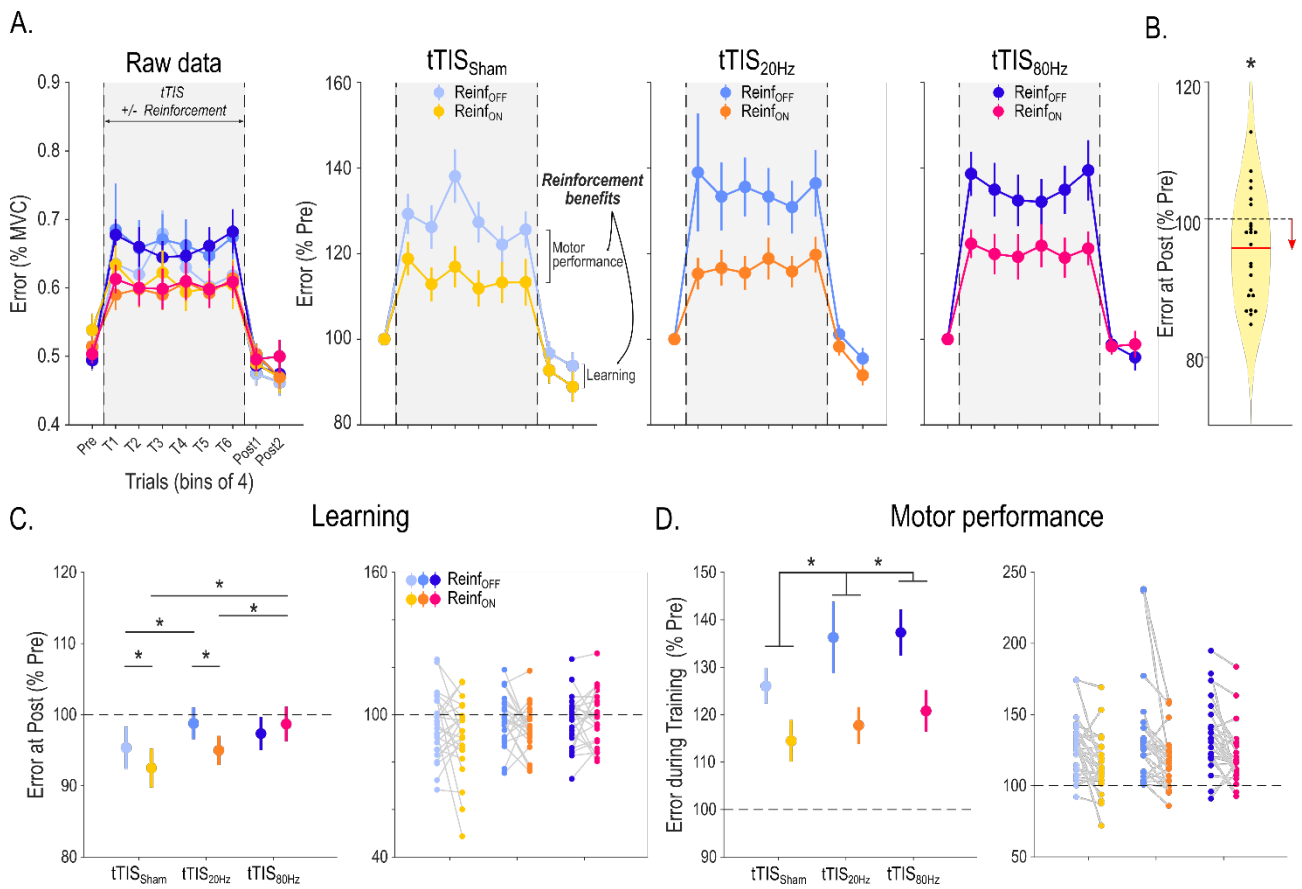


Figure 2. Behavioural results. A) Motor performance across training. Raw Error data (expressed in % of Maximum Voluntary Contraction [MVC]) are presented on the left panel for the different experimental conditions in bins of 4 trials. The increase in Error during Training is related to the visual uncertainty (i.e., intermittent disappearance of the cursor) that was applied to enhance reinforcement effects. On the right, the three plots represent the Pre-training normalised Error in the tTIS_{Sham}, tTIS_{20Hz} and tTIS_{80Hz} blocks. Reinforcement-related benefits represent the improvement in the Error measured in the Reinf_{ON} and Reinf_{OFF} blocks, during Training (reflecting benefits in motor performance) or at Post-training (reflecting benefits in learning). **B) Averaged learning across conditions.** Violin plot showing the Error distribution at Post-training (expressed in % of Pre-training) averaged across conditions, as well as individual subject data. A single-sample t-test showed that the Post-training Error was reduced compared to the Pre-training level, indicating significant learning in the task. **C) Motor learning.** Averaged Error at Post-training (normalised to Pre-training) and the

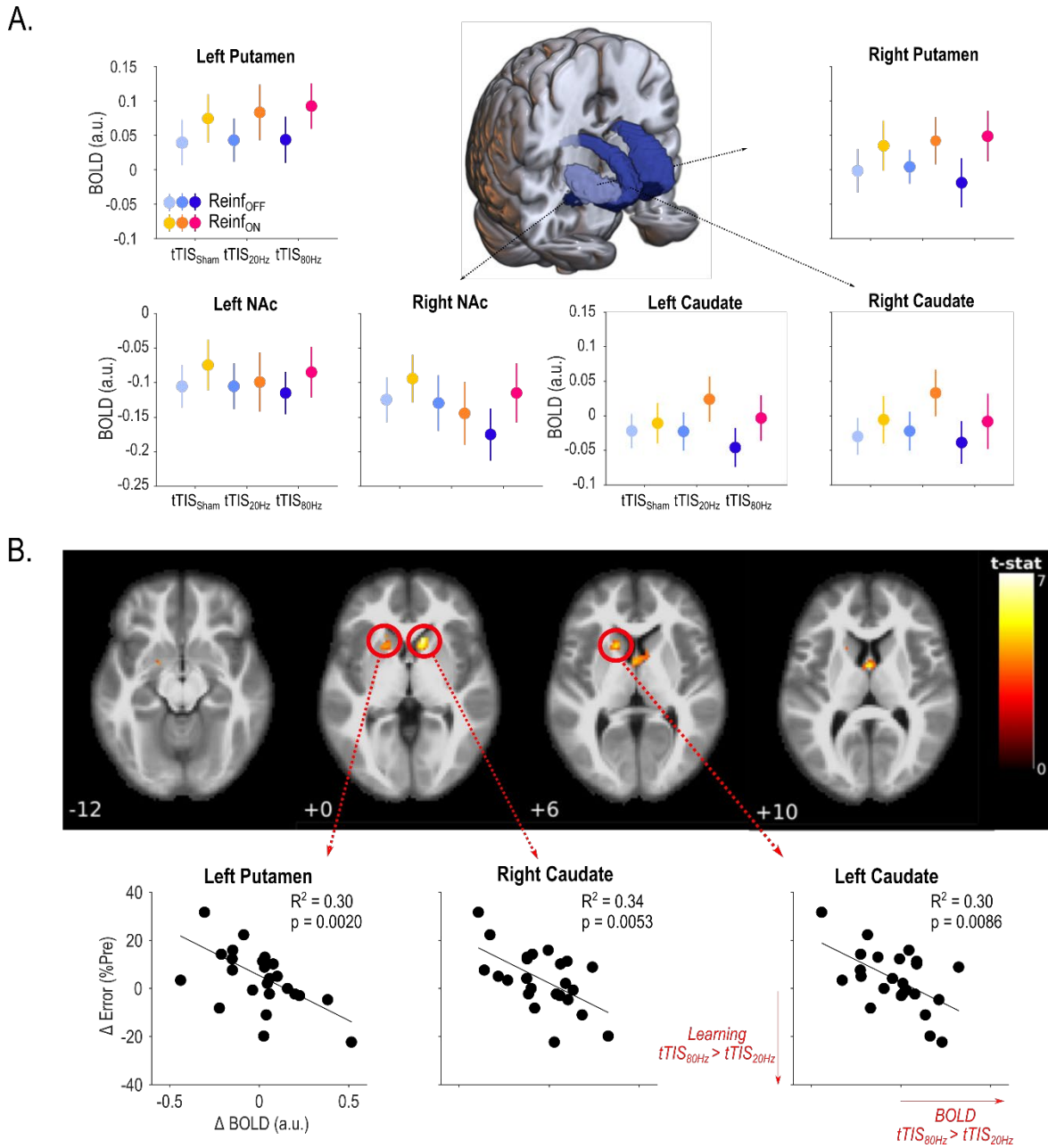
corresponding individual data points in the different experimental conditions are shown on the left and right panels, respectively. Reduction of Error at Post-training reflects true improvement at tracking the target in Test conditions (in the absence of reinforcement, visual uncertainty or tTIS). The LMM ran on these data revealed a specific effect of tTIS_{80Hz} on reinforcement-related benefits in learning. **D) Motor performance.** Averaged Error during Training (normalised to Pre-training) and the corresponding individual data points in the different experimental conditions are shown on the left and right panels, respectively. Individual data points are shown on the right panel. Error change during Training reflect the joint contribution of the experimental manipulations (visual uncertainty, potential reinforcement and tTIS) on motor performance. The LMM ran on these data showed a frequency-dependent effect of tTIS on motor performance, irrespective of reinforcement. *: $p < 0.05$. Data are represented as mean \pm SE.

The effect of tTIS_{80Hz} on reinforcement motor learning is related to modulation of neural activity in the striatum

As mentioned above, task-based fMRI was acquired during Training with concomitant tTIS. This allowed us to evaluate the neural effects of tTIS and their potential relationship to the behavioural effects reported above. As a first qualitative evaluation of the data, we performed a whole-brain analysis in the tTIS_{Sham} condition to assess the network activated during reinforcement motor learning (Reinf_{ON} condition). Consistent with previous neuroimaging studies employing similar tasks (Floyer-Lea and Matthews, 2004, 2005), we found prominent BOLD activations in a motor network including the putamen, thalamus, cerebellum and sensorimotor cortex, particularly on the left hemisphere, contralateral to the trained hand (**Figure S3, Table S2**). Notably though, contrasting Reinf_{ON} and Reinf_{OFF} conditions did not reveal any significant cluster at the whole-brain level.

As a second step, we evaluated the effect of tTIS on striatal activity, as a function of the type of reinforcement feedback and focusing on the very same regions of interest (ROI) that were used to optimise tTIS exposure in the modelling (see above). Based on this, we extracted averaged BOLD activity within the bilateral putamen, caudate and NAc based on the Brainnetome atlas (BNA, (Jiang, 2013)), in the different experimental conditions and

considered these six striatal ROIs (ROI_{STR}) as fixed effects in the LMM. This model revealed a strong enhancement of striatal activity with Reinf_{ON} with respect to Reinf_{OFF} ($F_{(1, 800.01)} = 13.23$; $p < 0.001$; $d = 0.25$ for the contrast Reinf_{ON} – Reinf_{OFF}) consistent with previous literature (Bartra et al., 2013), but no tTIS_{TYPE} effect ($F_{(2, 800.01)} = 0.46$; $p = 0.63$) and no interaction effect (all $p > 0.65$; **Figure 3A**). Despite the absence of effects of tTIS on averaged striatal activity, we then asked whether the behavioural effects of tTIS_{80Hz} on reinforcement motor learning (i.e., tTIS_{80Hz} vs. tTIS_{20Hz} and tTIS_{Sham} with Reinf_{ON}) could be linked to modulation of activity in core brain regions. To do so, we ran a whole-brain analysis focusing on the main behavioural effects mentioned above. Results revealed that the effect of tTIS_{80Hz} (with respect to tTIS_{20Hz}) on motor learning in the Reinf_{ON} condition was specifically related to modulation of activity in two clusters encompassing the left putamen and bilateral caudate (**Figure 3B, Table S3**; uncorrected voxel-wise FWE: $p=0.001$, and corrected cluster-based FDR: $p = 0.05$). No significant clusters were found for the tTIS_{80Hz} – tTIS_{Sham} contrast, neither for the control tTIS_{20Hz} - tTIS_{Sham} contrast. Overall, these results provide evidence that the detrimental effect of tTIS_{80Hz} on reinforcement learning of motor skills is related to a specific modulation of oscillatory activity in the striatum, supporting the idea that high gamma striatal oscillations are causally involved in reinforcement learning. More generally, this analysis shows for the first time that tTIS stimulation can indeed modulate the activity of deep brain regions in humans in a frequency-dependent and behaviourally-relevant way.



tTIS_{80Hz} enhances effective connectivity between the striatum and frontal cortex.

Interactions between the striatum and frontal cortex are crucial for a variety of behaviours including motor and reinforcement learning (Haber, 2016). In particular, reinforcement motor learning requires to use information about task success to guide future motor commands (Vassiliadis et al., 2021), a process for which the striatum may play an integrative role at the interface between fronto-striatal loops involved in reward processing and motor control (Graybiel and Grafton, 2015; Haber, 2016). In a subsequent analysis, we asked whether striatal tTIS modulates striatum to frontal cortex communication during reinforcement motor learning. More specifically, we computed effective connectivity (using the generalized psychophysiological interactions method; (McLaren et al., 2012)) between striatal and frontal regions classically associated with motor and reward-related functions, and thought to be involved in reinforcement motor learning (Sidarta et al., 2016; Codol et al., 2020). For the motor network, we evaluated effective connectivity between motor parts of the striatum (i.e., dorso-lateral putamen (dIPu) and dorsal caudate (dCa)) and two regions strongly implicated in motor learning: the medial part of the supplementary motor area (SMA) and the part of the primary motor cortex (M1) associated to upper limb functions as defined in the BNA atlas (**Figure 4A**; (Hardwick et al., 2013)). For the reward network, we assessed connectivity between parts of the striatum classically associated to limbic functions (i.e., the NAc and the ventro-medial putamen (vmPu) and two frontal areas involved in reward processing: the anterior cingulate cortex (ACC) and the ventro-medial prefrontal cortex (vmPFC; **Figure 4B**; Bartra et al., 2013). The LMM ran with the fixed effects $\text{Reinf}_{\text{TYPE}}$, $\text{tTIS}_{\text{TYPE}}$ and $\text{Network}_{\text{TYPE}}$ showed a significant effect of $\text{tTIS}_{\text{TYPE}}$ ($F_{(2, 2264.0)} = 5.42$; $p = 0.0045$), that was due to higher connectivity in the tTIS_{80Hz} condition with respect to tTIS_{Sham} ($p =$

0.0038; $d = 0.16$) and $tTIS_{20Hz}$ (at the trend level, $p = 0.069$; $d = 0.11$). There was no difference in connectivity between $tTIS_{20Hz}$ and $tTIS_{Sham}$ ($p = 0.58$; $d = 0.051$). Hence, $tTIS_{80Hz}$, but not $tTIS_{20Hz}$, enhanced effective connectivity between the striatum and frontal cortex during motor training.

The LMM did not reveal any effect of $Reinf_{TYPE}$ ($F_{(1, 2264.0)} = 0.010$; $p = 0.92$), $Network_{TYPE}$ ($F_{(1, 2264.0)} = 3.16$; $p = 0.076$) and no double interaction (note the trend for a $Reinf_{TYPE} \times Network_{TYPE}$ effect though: $F_{(1, 2264.0)} = 3.52$; $p = 0.061$). Yet, we did find a significant $Reinf_{TYPE} \times tTIS_{TYPE} \times Network_{TYPE}$ interaction ($F_{(2, 2264.0)} = 4.87$; $p = 0.0078$). Such triple interaction was related to the fact that $tTIS_{80Hz}$ increased connectivity in the $Reinf_{ON}$ condition in the motor network ($Reinf_{ON}$ vs. $Reinf_{OFF}$: $p = 0.0012$; $d = 0.33$; Figure 4A), while it tended to have the opposite effect in the reward network ($p = 0.063$; $d = -0.19$; Figure 4B). This increase was not present in any of the two networks when either $tTIS_{Sham}$ or $tTIS_{20Hz}$ were applied (all $p > 0.40$). Moreover, in the motor network, connectivity in the $Reinf_{ON}$ condition was higher with $tTIS_{80Hz}$ than with $tTIS_{Sham}$ ($p < 0.001$; $d = 0.42$) and $tTIS_{20Hz}$ (at the trend level; $p = 0.059$; $d = 0.23$, Figure 4A). These data suggest that $tTIS_{80Hz}$ enhanced the neuromodulatory influence of the striatum on motor cortex during task performance, but only in the presence of reinforcement. In the reward network, post-hocs revealed that connectivity in the $Reinf_{OFF}$ condition was significantly higher with $tTIS_{80Hz}$ compared to $tTIS_{20Hz}$ ($p = 0.045$; $d = 0.25$; Figure 4B), in line with the general effect of $tTIS_{TYPE}$ on connectivity reported above. This pattern of results suggests that the increase of connectivity from striatum to frontal cortex observed with $tTIS_{80Hz}$ depends on the presence of reinforcement, in particular in the motor network. Such reinforcement-dependent increase of connectivity may reflect the preferential entrainment of striatal gamma oscillations with $tTIS_{80Hz}$ (Krause et al., 2019) in a situation

where these oscillations are already boosted by the presence of reinforcement ((Berke, 2009); see Discussion).

Notably, contrary to the BOLD results presented above, we did not find any correlations between the effects of tTIS_{80Hz} on connectivity and motor learning, neither in the motor (robust linear regression: tTIS_{80Hz} – tTIS_{Sham}: $R^2 = 0.019$; $p = 0.48$; tTIS_{80Hz} – tTIS_{20Hz}: $R^2 = 0.034$; $p = 0.54$) nor in the reward (tTIS_{80Hz} – tTIS_{Sham}: $R^2 = 0.037$; $p = 0.46$; tTIS_{80Hz} – tTIS_{20Hz}: $R^2 < 0.001$; $p = 0.75$) network, suggesting some degree of independence between the effect of tTIS_{80Hz} on reinforcement motor learning and on effective connectivity.

As a control, we verified that the effects of tTIS_{TYPE} on connectivity could not be observed in a control network associated to language (as defined by Shirer et al. 2012), which was unlikely to be involved in the present task and did not include the striatum (see Methods). As expected, effective connectivity within the language network was not modulated by Reinf_{TYPE} ($F_{(1, 547)} = 0.81$; $p = 0.37$), nor by tTIS_{TYPE} ($F_{(2, 547)} = 0.58$; $p = 0.56$), or by Reinf_{TYPE} x tTIS_{TYPE} ($F_{(2, 547)} = 0.45$; $p = 0.64$).

Overall, these results highlight the ability of tTIS_{80Hz}, but not tTIS_{20Hz}, to modulate striatum to frontal cortex connectivity. Moreover, they suggest that a potential mechanism of action of tTIS_{80Hz} is the induction of a state of hyper-connectivity between striatum and motor cortex that may be detrimental for reinforcement motor learning.

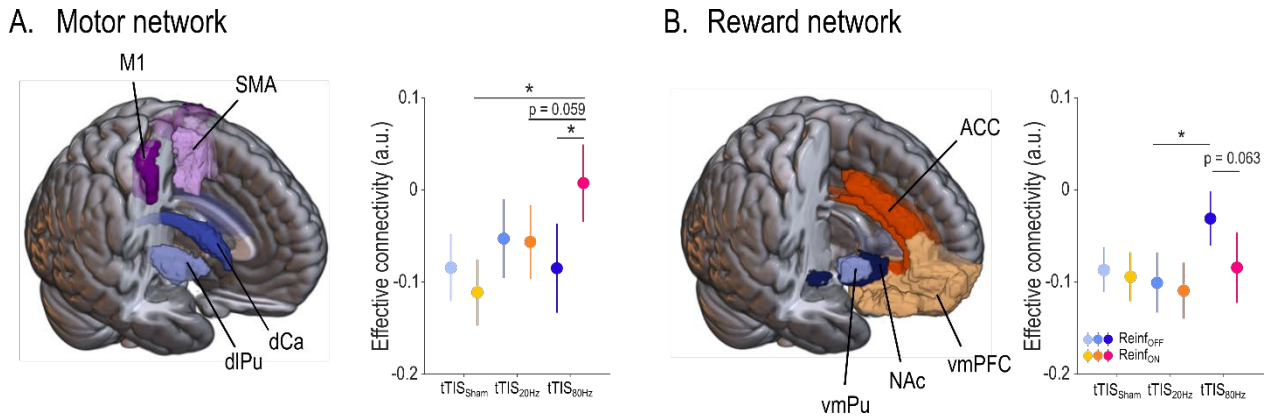


Figure 4. Striatum to frontal cortex effective connectivity. A) Motor network. On the left, 3D reconstruction of the masks used for the motor network (i.e., dorso-lateral putamen, dorsal caudate, M1, SMA). On the right, plot showing effective connectivity from motor striatum to motor cortex in the different experimental conditions. Note the increase of connectivity with tTIS_{80Hz} in the presence of reinforcement. **B) Reward network.** On the left, 3D reconstruction of the masks used for the reward network (i.e., ventro-medial putamen, NAc, vmPFC, ACC). On the right, plot showing effective connectivity from motor striatum to motor cortex in the different experimental conditions. *: $p < 0.05$. Data are represented as mean \pm SE.

Neural effects of tTIS_{80Hz} depend on impulsivity

Hyper-connectivity in fronto-striatal circuits has been described as a pathophysiological mechanism in multiple neuro-psychiatric disorders involving impulsivity (e.g., (Ma et al., 2016; Wang et al., 2016; Hampton et al., 2017; Mosley et al., 2019)). Thus, in a subsequent exploratory analysis, we asked if increased striatum to motor cortex connectivity with tTIS_{80Hz} was similarly associated to individual levels of impulsivity, as evaluated by a well-established independent delay-discounting questionnaire performed at the beginning of the experiment (Kirby et al., 1999; Mitchell et al., 2005). As a first step, we asked if striatum to motor cortex connectivity was related to impulsivity during reinforcement motor learning in the absence of stimulation (i.e., in the tTIS_{Sham} condition). Indeed, we found a significant positive relationship between impulsivity and striatum to motor cortex connectivity (robust linear regression: $R^2 = 0.10$; $p = 0.0038$), in line with previous results (Ma et al., 2016; Wang et al., 2016; Hampton et al., 2017; Mosley et al., 2019). Then, we evaluated whether the

increase of connectivity observed with tTIS_{80Hz} in the Reinf_{ON} condition (Figure 4A) could be related to impulsivity. Indeed, we found that the effect of tTIS_{80Hz} on connectivity was negatively correlated to impulsivity both when contrasting tTIS_{80Hz} with tTIS_{Sham} ($R^2 = 0.19$; $p = 0.043$, **Figure 5A, left**) and with tTIS_{20Hz} ($R^2 = 0.28$; $p = 0.021$, **Figure 5A, middle**): participants with the largest increase in connectivity with tTIS_{80Hz} in the Reinf_{ON} condition were also the least impulsive ones. Such correlation was absent when contrasting tTIS_{20Hz} and tTIS_{Sham} ($R^2 = 0.0031$; $p = 0.31$, **Figure 5A, right**), but also when considering the same contrasts in the reward instead of the motor network ($p = 0.93$ and $p = 0.86$ for the tTIS_{80Hz}-tTIS_{Sham} and tTIS_{80Hz}-tTIS_{20Hz} contrasts, respectively). Hence, striatum to motor cortex effective connectivity during the task was positively correlated to impulsivity, but the change in connectivity induced by tTIS_{80Hz} was rather negatively associated with impulsivity. This may be due to a ceiling effect in the most impulsive participants: exhibiting initially high levels of connectivity may leave less room for further modulation by tTIS_{80Hz}. This result suggests that inter-individual variability in impulsivity might influence neural responses to striatal tTIS_{80Hz}.

To further evaluate this idea, we ran the same analysis on the BOLD data presented above. More specifically, we asked if inter-individual variability in the neural effects of tTIS_{80Hz} during reinforcement motor learning (i.e., in the Reinf_{ON} condition) was related to impulsivity at the whole-brain level. Strikingly, this analysis revealed that impulsivity was associated to the effect of tTIS_{80Hz} (with respect to tTIS_{20Hz}) specifically in the left caudate nucleus (**Figure 5B, Table S4**; uncorrected voxel-wise FWE: $p=0.001$, and corrected cluster-based FDR: $p = 0.05$). No other clusters were found. As such, the most impulsive participants exhibited an increase of left caudate activity with tTIS_{80Hz} (compared to tTIS_{20Hz}) while the least impulsive ones rather presented a decrease of BOLD signal, consistent with the

idea that impulsivity modulates the neuronal responsiveness to tTIS ($R^2 = 0.47$; $p < 0.001$; **Figure 5C**). No significant clusters of correlation were found for the tTIS_{80Hz} – tTIS_{Sham} contrast, neither for the control tTIS_{20Hz} - tTIS_{Sham} contrast. Hence, this analysis suggests that the effect of tTIS_{80Hz} on caudate activity depends on participants' impulsivity.

As a last step, we verified if impulsivity was also predictive of the behavioural effects of tTIS_{80Hz} on reinforcement motor learning. We did not find any significant correlation between impulsivity and the effect of tTIS_{80Hz} on motor learning (tTIS_{80Hz} – tTIS_{Sham}: $R^2 = 0.098$; $p = 0.17$; tTIS_{80Hz} – tTIS_{20Hz}: $R^2 = 0.11$; $p = 0.21$). Hence, impulsivity was associated to the neural, but not the behavioural effects of tTIS_{80Hz}.

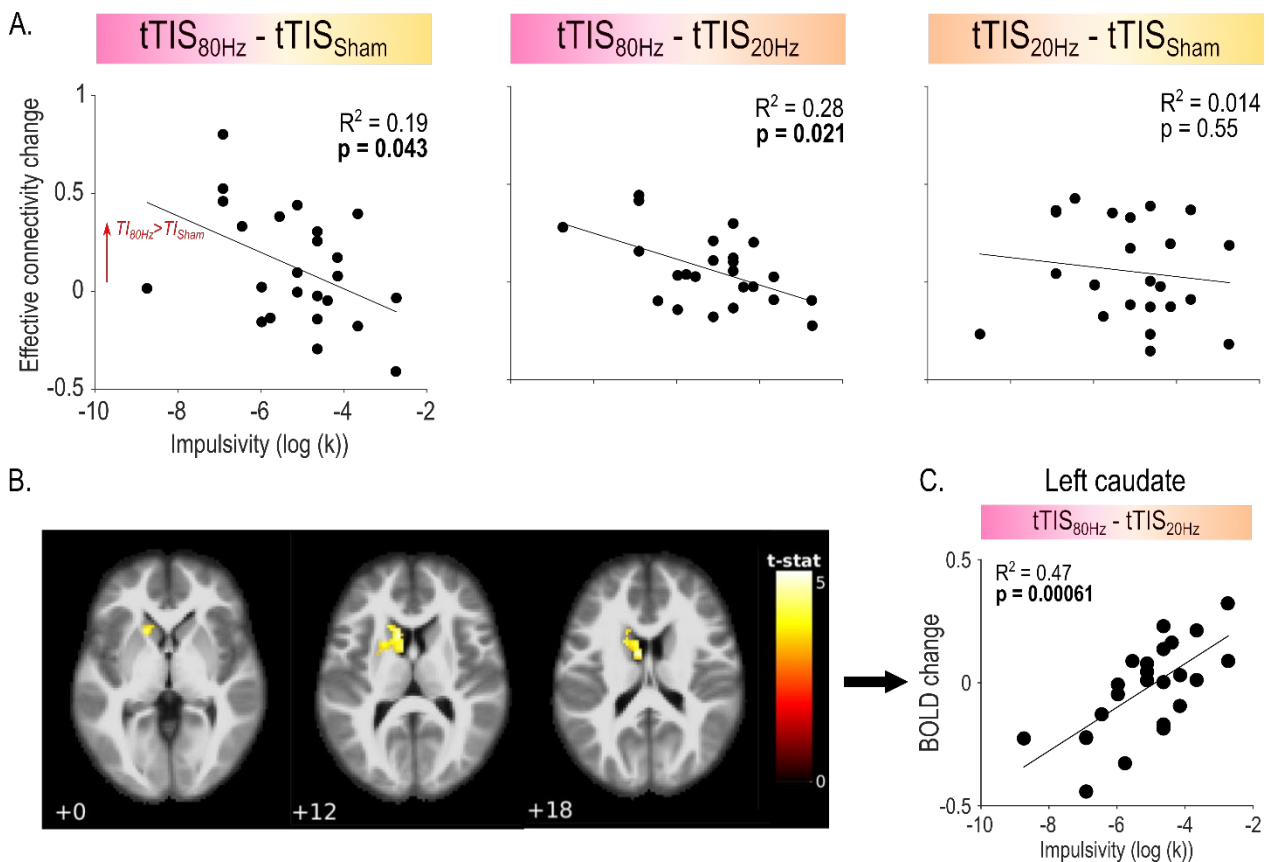


Figure 5. Relationship between impulsivity and the neural effects of tTIS_{80Hz}. **A)** Correlations between impulsivity and tTIS-related modulation of effective connectivity. Impulsivity

was associated to the neural effects of tTIS_{80Hz} both when contrasting to tTIS_{Sham} (left) and tTIS_{20Hz} (middle), but was not correlated to the effect of tTIS_{20Hz} (right). **B) Whole-brain correlation between the neural effects of tTIS_{80Hz} (with respect to tTIS_{20Hz}) and impulsivity.** Correlation between tTIS-related modulation of striatal activity (tTIS_{80Hz} – tTIS_{20Hz}) during reinforcement motor learning (Rein_{ON}) and individual impulsivity levels. A single significant cluster of correlation was found in left caudate (uncorrected voxel-wise FWE: $p=0.001$, and corrected cluster-based FDR: $p = 0.05$). **C) Correlation between left caudate activity and impulsivity.** A positive correlation was found showing that participants with higher levels of impulsivity exhibited stronger activation of the left caudate in the tTIS_{80Hz} (with respect to tTIS_{20Hz}).

4.4. Discussion

In this study, we combined striatal tTIS with electric field modelling, behavioural and fMRI analyses to evaluate the causal role of the striatum in reinforcement learning of motor skills in healthy humans. tTIS_{80Hz}, but not tTIS_{20Hz}, disrupted the ability to learn from reinforcement feedback. This behavioural effect was associated to modulation of neural activity specifically in the striatum. As a second step, we show that tTIS_{80Hz}, but not tTIS_{20Hz}, increased the neuromodulatory influence of the striatum on connected frontal cortical areas involved in reinforcement motor learning. Finally, inter-individual variability in the neural effects of tTIS_{80Hz} could be partially explained by impulsivity, suggesting that this trait may constitute a determinant of responsiveness to high gamma striatal tTIS. Overall, the present study shows for the first time that striatal tTIS can non-invasively modulate a striatal mechanism involved in reinforcement learning, opening new horizons for the study of causal relationships between deep brain structures and human behaviour.

In this work we investigated the causal role of the human striatum in reinforcement learning of motor skills in healthy humans; a question that cannot be addressed with conventional non-invasive brain stimulation techniques. In particular, by stimulating at different frequencies, we aimed at dissociating striatal mechanisms involved in reinforcement and sensorimotor learning. In line with our main hypothesis, we found that striatal tTIS_{80Hz} altered reinforcement learning of a motor skill. Such disruption was frequency- and reinforcement-specific: learning was not altered with striatal tTIS_{20Hz} in the presence of reinforcement, or when striatal tTIS_{80Hz} was delivered in the absence of reinforcement. The rationale to stimulate at high gamma frequency was based on animal work showing reinforcement-related

modulation of gamma oscillations in the striatum (Berke, 2009; van der Meer and Redish, 2009; Kalenscher et al., 2010; van der Meer et al., 2010; Donnelly et al., 2014; Catanese et al., 2016) and in the frontal cortex (Rothé et al., 2011; Catanese et al., 2016; Del Arco et al., 2017; Yoshimoto et al., 2022). Hence, a potential explanation of the present result is that striatal tTIS_{80Hz} interfered with the natural high gamma oscillations in the striatum (Esmaeilpour et al., 2021) and thereby reduced the ability of participants to learn from reinforcement information. In contrast, we found that striatal tTIS_{20Hz} reduced learning, but only in the absence of reinforcement, consistent with the literature linking striatal beta oscillations to sensorimotor functions (Courtemanche et al., 2003; Costa et al., 2006; Brown, 2007; Jenkinson and Brown, 2011; Kondabolu et al., 2016). Hence, these elements suggest that different oscillations within the striatum support qualitatively different motor learning mechanisms with beta activity contributing mostly to sensory-based learning and high gamma activity being particularly important for reinforcement learning. More generally, these results add to the growing body of evidence showing that sensory- and reinforcement-based motor learning rely on partially different neural mechanisms (Sidarta et al., 2016; Therrien et al., 2016; Mathis et al., 2017; Uehara et al., 2018; Vassiliadis et al., 2019; Areshenkoff et al., 2022).

Interestingly, striatal tTIS (especially tTIS_{80Hz}) also impaired tracking performance during training, irrespective of the presence of reinforcement. This frequency-dependent reduction of motor performance may be due to altered neuronal processing in the sensorimotor striatum that may lead to less fine-tuned motor control abilities (Brücke et al., 2012). Importantly though, tTIS did not modulate the ability of participants to benefit from real-time

reinforcement feedback during motor performance. This suggests that striatal tTIS_{80Hz} altered the beneficial effects of reinforcement on learning (as evaluated in Test conditions at Post-training), but not on motor performance (as evaluated during Training). Such dissociation between the effects of striatal tTIS_{80Hz} on reinforcement-related gains in motor performance and learning may be explained by the fact that these two phases of the protocol probe different processes (Soderstrom and Bjork, 2015; Vassiliadis et al., 2022). While improvement of motor performance with reinforcement feedback relies on rapid adjustments of motor variability based on recent outcomes (Dhawale et al., 2019), reinforcement gains in learning (i.e., probed in Test conditions without reinforcement) may rather reflect the beneficial effect of performance feedback on the retention of motor memories (Abe et al., 2011; Galea et al., 2015a). Hence, a potential explanation for the present results is that striatal tTIS_{80Hz} did not disrupt real-time processing of reinforcement feedback, but may rather alter the strengthening of the memory trace based on reinforcements (Huang et al., 2011; Galea et al., 2015a). Overall, these results are consistent with the view that specific patterns of oscillatory activity in the striatum are causally involved in motor control and learning processes (Kondabolu et al., 2016), and can be modulated with electrical stimulation (Krause et al., 2019; Johnson et al., 2020; Beliaeva et al., 2021).

To better understand the neural effects and frequency-specificity of tTIS, we coupled striatal tTIS and task performance with simultaneous fMRI acquisition. The imaging results support the view that the effect of tTIS_{80Hz} on reinforcement learning of motor skills was indeed related to neuromodulation of the striatum. As such, when considering averaged BOLD activity, we found a general increase of striatal activity when reinforcement was provided (Bartra et al., 2013), but no effect of tTIS. Crucially though, the detrimental effect of

tTIS_{80Hz} on reinforcement learning was related to a specific modulation of activity in the caudate and putamen, providing evidence that the present behavioural effects were indeed driven by focal neuromodulation of the striatum (Figure 3). Interestingly, participants with stronger disruption of reinforcement learning at the behavioural level were also the ones exhibiting stronger suppression of striatal activity with tTIS_{80Hz} (compared to tTIS_{20Hz}), suggesting that tTIS-induced reduction of striatal activity is detrimental for reinforcement motor learning. Further analyses showed that tTIS_{80Hz}, but not tTIS_{20Hz}, increased the neuromodulatory influence of the striatum on frontal areas known to be important for motor learning and reinforcement processing (Krakauer et al., 2019; Averbeck and O'Doherty, 2022). Interestingly, this effect depended on the type of network considered (reward vs. motor) and on the presence of reinforcement. Striatal tTIS_{80Hz} coupled with reinforcement increased connectivity between the motor striatum and the motor cortex while it tended to have the opposite effect when considering the connectivity between limbic parts of the striatum and prefrontal areas involved in reward processing (Figure 4). This result may reflect the differential influence of striatal tTIS on distinct subparts of the striatum, depending on their pattern of activity during the task (Wessel et al., 2021). As such, a recent study in non-human primates showed that tACS can have opposite effects on neuronal activity based on the initial entrainment of neurons to the target frequency (Krause et al., 2019). Hence, the present differential effects of tTIS_{80Hz} on motor and reward striato-frontal pathways may be due to different initial patterns of activity in these networks in the presence of reinforcement. Electrophysiological recordings with higher temporal resolution than fMRI are required to confirm or refute this hypothesis. Overall, the present neuroimaging results support the idea that the behavioural effects of striatal tTIS_{80Hz} on reinforcement learning are associated to a selective modulation of striatal activity that influence striato-frontal communication.

Interestingly, the state of increased connectivity observed with striatal tTIS_{80Hz}, in particular in the motor network and in the presence of reinforcement, parallels previous findings showing hyper-connectivity in fronto-striatal networks in neuro-psychiatric disorders associated to impulsivity (Ma et al., 2016; Wang et al., 2016; Hampton et al., 2017; Mosley et al., 2019). Consistently, in a subsequent exploratory analysis, we found that individual impulsivity levels of the participants were positively associated to striatum to motor cortex connectivity in the absence of stimulation. Most importantly, we found that impulsivity was also predictive of the effect of tTIS_{80Hz} (but not tTIS_{20Hz}) on connectivity with the subjects with lower impulsivity exhibiting stronger increases of connectivity with tTIS_{80Hz}. Strikingly, impulsivity was also associated to tTIS_{80Hz}-related BOLD changes specifically in the left caudate. This link between impulsivity and neural responsiveness to tTIS is in line with the idea that impulsivity is associated to differences in striatal gamma oscillatory activity (Donnelly et al., 2014), which may determine the neural effects of striatal tTIS (Esmailpour et al., 2021). Hence, impulsivity could constitute a behavioural factor allowing to determine responsiveness to striatal tTIS_{80Hz}. Conversely, an interesting avenue for future research could aim at determining whether impulsivity can be modulated by striatal tTIS_{80Hz}.

Striatal tTIS had dissociable effects on motor learning depending on the frequency of stimulation and the presence of reinforcement. Such frequency- and state-dependent effects of non-invasive brain stimulation have been also shown at the cortical level with transcranial alternating current stimulation (tACS; e.g., (Ozen et al., 2010; Krause et al., 2019, 2022; Johnson et al., 2020; Beliaeva et al., 2021)). In particular, tACS studies on non-human primates have shown that oscillating electric fields act by entraining particular populations of

neurons to specific frequencies (Krause et al., 2019; Johnson et al., 2020) and that such neuromodulatory influence depends on the initial activity of these neurons (Ozen et al., 2010; Polanía et al., 2018; Krause et al., 2022). Similarly, responsiveness of a particular region to tTIS may depend on intrinsic properties of neurons (Esmaeilpour et al., 2021) and their pattern of activity during the task (Wessel et al., 2021). Taken together, these elements suggest that focality of tTIS depends not only on the distribution of the low-frequency electric field in the brain due to modelling-based electrode arrangement, but also on the stimulation frequency and the concurrent engagement of the target region in a particular task.

The present results also add experimental support to the idea that the effects of tTIS are related to amplitude modulation of electric fields deep in the brain and not to the high frequency fields themselves, in line with recent animal work (Grossman et al., 2017; Song et al., 2021). As such, the different behavioural and neural effects of striatal tTIS_{80Hz} and tTIS_{20Hz} despite comparable carrier frequencies (centered on 2kHz) indicate that temporal interference was indeed the driving force of the present effects. Moreover, disruption of reinforcement motor learning with tTIS_{80Hz} (relative to tTIS_{20Hz}) was specifically related to neuromodulation of the striatum, where the amplitude of the tTIS field was highest according to our simulations and in line with previous experiments in cadavers (Acerbo et al., 2022; Violante et al., 2022). Hence, we believe that the frequency- and state-dependent tTIS effects reported here cannot be explained by direct modulation of neural activity by the high frequency fields. Yet, disentangling the neural effects of the low-frequency envelope and the high frequency component appears as an important next step to better characterise the mechanisms underlying tTIS (Mirzakhali et al., 2020).

The present findings show for the first time the ability of non-invasive striatal transcranial temporal interference stimulation to interfere with reinforcement learning in humans through a selective modulation of striatal activity. Such deep brain stimulation was well tolerated and compatible with efficient blinding, suggesting that tTIS provides the exciting option to circumvent the steep depth-focality trade-off of current non-invasive brain stimulation approaches in a safe and effective way. Overall, tTIS opens new possibilities for the study of causal brain-behaviour relationships and for the treatment of neuro-psychiatric disorders associated to alterations of deep brain structures.

4.5. Methods

4.5.1. Participants

A total of 24 right-handed healthy volunteers participated in the present study (15 women, 25.3 ± 0.1 years old; mean \pm SE). Handedness was determined via a shortened version of the Edinburgh Handedness inventory ((Oldfield, 1971); laterality index = $89.3 \pm 2.14\%$). None of the participants suffered from any neurological or psychiatric disorder, nor taking any centrally-acting medication (see Supplementary Materials for a complete list of exclusion criteria). All participants gave their written informed consent in accordance with the Declaration of Helsinki and the Cantonal Ethics Committee Vaud, Switzerland (project number 2020-00127). Finally, all participants were asked to fill out a delay-discounting monetary choice questionnaire (Kaplan et al., 2016), which evaluates the propensity of subjects to choose smaller sooner rewards over larger later rewards, a preference commonly associated to choice impulsivity (Kirby et al., 1999; Mitchell and Potenza, 2014).

4.5.2. Experimental procedures

The study employed a randomised, double-blind, sham-controlled design. Following screening and inclusion, participants were invited to a single experimental session including performance of a motor learning task with concurrent transcranial electric Temporal Interference stimulation (tTIS) of the striatum and functional magnetic resonance imaging (fMRI). Overall, participants practiced 6 blocks of trials, that resulted from the combination of two reinforcement feedback conditions (Reinf_{TYPE}: Reinf_{ON} or Reinf_{OFF}) with three types of striatal stimulation (tTIS_{TYPE}: tTIS_{Sham}, tTIS_{20Hz} or tTIS_{80Hz}).

4.5.2.1. Motor learning task

4.5.2.1.1. General aspects

Participants practiced an adaptation of a widely used force-tracking motor task (Abe et al., 2011, Steel et al., 2016) with a fMRI-compatible fiber optic grip force sensor (Current designs, Inc., Philadelphia, PA, USA) positioned in their right hand. The task was developed on Matlab 2018 (the Mathworks, Natick, Massachusetts, USA) exploiting the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) and was displayed on a computer screen with a refresh rate of 60 Hz. The task required participants to squeeze the force sensor to control a cursor displayed on the screen. Increasing the exerted force resulted in the cursor moving vertically and upward in a linear way. Each trial started with a preparatory period in which a sidebar appeared at the bottom of the screen (**Figure 1A**). After a variable time interval (0.9 to 1.1 s), a cursor (black circle) popped up in the sidebar and simultaneously a target (grey larger circle with a cross in the middle) appeared, indicating the start of the movement period. Subjects were asked to modulate the force applied on the transducer to keep the cursor as close as possible to the center of the target. The target moved in a sequential way along a single vertical axis for 7 s. The maximum force required (i.e., the force required to reach the target when it was in the most upper part of the screen; $\text{MaxTarget}_{\text{Force}}$) was set at 4% of maximum voluntary contraction (MVC) evaluated in the beginning of the experiment. This low force level was chosen based on pilot experiments to limit muscular fatigue. Finally, each trial ended with a blank screen displayed for 2 s before the beginning of the next trial.

4.5.2.1.2. Trial types and reinforcement manipulation

During the experiment, participants were exposed to different types of trials (**Figure 1A, Supplementary Video**). In Test trials, the cursor remained on the screen and the target was consistently displayed in grey for the whole duration of the trial. These trials served to evaluate Pre- and Post-training performance for each block. In Reinf_{ON} and Reinf_{OFF} trials (used during Training only), we provided only limited visual feedback to the participants in order to increase the impact of reinforcement on learning (Izawa and Shadmehr, 2011; Mawase et al., 2017; Vassiliadis et al., 2021, 2022). As such, the cursor was only intermittently displayed during the trial: it was always displayed in the first second of the trial, and then disappeared for a total of 4.5 s randomly split on the remaining time by bits of 0.5 s. The cursor was therefore displayed 35.7% of the time during these trials (2.5 s over the 7 s trial). Importantly, contrary to the cursor, the target always remained on the screen for the whole trial and participants were instructed to continue to track the target even when the cursor was away.

In addition to this visual manipulation, in Reinf_{ON} trials, participants also trained with reinforcement feedback indicating success or failure of the tracking in real time. As such, participants were informed that, during these trials, the color of the target would vary as a function of their performance: the target was displayed in green when tracking was considered as successful and in red when it was considered as failure. Online success on the task was determined based on the Error, defined as the absolute force difference between the force required to be in the center of the target and the exerted force (Abe et al., 2011; Steel et al., 2016; Vassiliadis et al., 2021, 2022). The Error, expressed in percentage of MVC, was

computed for each frame refresh and allowed to classify a sample as successful or not based on a closed-loop reinforcement schedule (Therrien et al., 2016). More specifically, for each training trial, a force sample was considered as successful if the Error was below the median Error over the 4 previous trials at this specific sample. Put differently, to be successful, participants had to constantly beat their previous performance. This closed-loop reinforcement schedule allowed us to deliver consistent reinforcement feedback across individuals and conditions, while maximizing uncertainty on the presence of reinforcement, an aspect that is crucial for efficient reinforcement motor learning (Dayan et al., 2014a). Notably, in addition to this closed-loop design, samples were also considered as successful if the cursor was very close to the center of the target (i.e., within one radius around the center, corresponding to an Error below 0.2% of MVC). This was done to prevent any conflict between visual information (provided by the position of the cursor relative to the target) and reinforcement feedback (provided by the color of the target), which could occur in situations of extremely good performance (when the closed-loop Error cut-off was below 0.2% of MVC).

As a control, Reinf_{OFF} trials were similar to Reinf_{ON} trials with the only difference that the displayed colors were either cyan or magenta, and were generated randomly. Participants were explicitly told that, in this condition, colors were displayed randomly and could be ignored. The visual properties of the target in the Reinf_{OFF} condition were designed to match the Reinf_{ON} condition in terms of relative luminance (cyan: RGB = [127.5 242.1 255] matched to green: [127.5 255 127.5] and magenta: [211.7 127.5 255] to red: [255 127.5 127.5]) and average frequency of change in colors (i.e., the average number of changes in colors divided by the total duration of a trial, see Supplementary materials).

4.5.2.1.3. Motor learning protocol

After receiving standardised instructions about the force-tracking task, participants practiced 5 blocks of familiarization (total of 75 trials) without tTIS. The first block of familiarization included 20 trials with the target moving in a regular fashion (0.5 Hz sinuoid). Then, in a second block of familiarization, participants performed 35 trials of practice with an irregular pattern, with the same properties as the training patterns (see below). Finally, we introduced the reinforcement manipulation and let participants perform 2 short blocks (8 trials each) including Reinf_{ON} and Reinf_{OFF} trials. These four first blocks of familiarization were performed outside the MRI environment. A last familiarization block (4 trials) was performed after installation in the scanner, to allow participants to get used to performing the task in the MRI. This long familiarization allowed participants to get acquainted with the use of the force sensor, before the beginning of the experiment.

During the main part of the experiment, participants performed 6 blocks of trials in the MRI with concurrent striatal tTIS (**Figure 1B**). Each block was composed of 4 Pre-training trials followed by 24 Training and 8 Post-training trials. Pre- and Post-training trials were performed in Test conditions, without tTIS and were used to evaluate motor learning. Training trials were performed with or without reinforcement feedback and with concomitant striatal tTIS and were used as a proxy of motor performance. During Training, trials were interspersed with 25 s resting periods every 4 trials (i.e., used for fMRI contrasts, see below). The order of the 6 experimental conditions was pseudo-randomised across participants: the 6 blocks were divided into 3 pairs of blocks with the same tTIS condition and each pair was then composed of one Reinf_{ON} and one Reinf_{OFF} block. Within this structure, the order of the Tl_{TYPE} and Reinf_{TYPE} conditions were evenly balanced among the 24 participants.

As mentioned above, the protocol involved multiple evaluations of motor learning within the same experimental session. In order to limit carry-over effects from one block to the following, each experimental block was associated to a different pattern of movement of the target (**Figure S1**). Put differently, in each block, participants had to generate a new pattern of force to successfully track the target. To balance the patterns' difficulty, they all consisted in the summation of 5 sinusoids of variable frequency (range: 0.1-1.5 Hz) that presented the following properties: a) Average force comprised between 45 and 55% of the $\text{MaxTarget}_{\text{Force}}$; b) Absolute average derivative of the $\text{MaxTarget}_{\text{Force}}$ comprised between 54 and 66 % / s; c) Number of peaks = 14 (defined as an absolute change of force of at least 1% of $\text{MaxTarget}_{\text{Force}}$). These parameters were determined based on pilot experiments to obtain a relevant level of difficulty for young healthy adults and consistent learning across the different patterns.

4.5.2.2. Transcranial Electric Temporal Interference Stimulation (tTIS) applied to the striatum

4.5.2.2.1. General concept

Transcranial temporal interference stimulation (tTIS) is an innovative non-invasive brain stimulation approach, in which two or more independent stimulation channels deliver high-frequency currents in the kHz range (oscillating at f_1 and $f_1 + \Delta f$; **Figure 1C**). These high-frequency currents have been proposed to be too high to modulate neuronal activity (Grossman et al., 2017; 2018). Still, by applying a small shift in frequency, they result in a

modulated electric field with the envelope oscillating at the low-frequency Δf (target frequency) between the two sources. The peak of the envelope amplitude can be steered towards specific areas located deep in the brain, by tuning the electrodes' position and current ratio across stimulation channels (**Figure 1C, 1D**; (Grossman et al., 2017)). Based on these properties, tTIS has been shown to focally target activity of deep structures in rodents, without engaging overlying tissues (Grossman et al., 2017). Here, we applied temporal interference stimulation transcranially via surface electrodes applying a low-intensity, sub-threshold protocol respecting the currently accepted cut-offs and safety guidelines for low intensity transcranial electric stimulation in humans (Antal et al., 2017).

4.5.2.2.2. Stimulators

The currents for tTIS were generated by two independent DS5 isolated bipolar constant current stimulators (*Digitimer Ltd, Welwyn Garden City, UK*). The stimulation patterns were created using custom-based Matlab graphical user interface and transmitted to the current sources using a standard digital-analog converter (*DAQ USB-6216, National Instruments, Austin, TX, USA*). Finally, an audio transformer was added between stimulators and subjects, in order to avoid possible direct current accumulation.

4.5.2.2.3. Stimulation protocols

During the 6 Training blocks, we applied three different types of striatal tTIS (2 blocks each): a stimulation with a tTIS envelope modulated at 20Hz (tTIS_{20Hz}), a stimulation with a

tTIS envelope modulated at 80Hz (tTIS_{80Hz}) and a sham stimulation (tTIS_{Sham}). For tTIS_{20Hz}, the posterior stimulation channel (TP7-TP8, see below) delivered a 1.99 kHz stimulation while the anterior one delivered a 2.01 kHz ($\Delta f = 20$ Hz). For tTIS_{80Hz}, the posterior and anterior channels delivered 1.96 kHz and 2.04 kHz, respectively ($\Delta f = 80$ Hz). Hence in both conditions, the high frequency component was comparable and the only difference was Δf . During each block, tTIS was applied for 5 minutes (6 x 50 s) during Training. Each stimulation period started and ended with currents ramping-up and -down, respectively, for 5 s. tTIS was applied only while participants were performing the motor task and not during resting periods or Pre- and Post-training assessments. Finally, tTIS_{Sham} consisted in a ramping-up (5 s) immediately followed by a ramping-down (5 s) of 2 kHz currents delivered without any shift in frequency. This condition allowed us to mimic the sensations experienced during the active conditions tTIS_{20Hz} and tTIS_{80Hz}, while delivering minimal brain stimulation (Figure S2). A trigger was sent 5 seconds before the beginning of each trial in order to align the beginning of the task and the beginning of the frequency shift after the ramp-up. Other TI stimulation parameters were set as follows: current intensity per stimulation channel = 2 mA, electrode type: round, conductive rubber with conductive cream/paste, electrode size = 3 cm² (see ContES checklist in Supplementary materials for more details).

The stimulation was applied within the MRI environment (Siemens 3T MAGNETOM Prisma; Siemens Healthcare, Erlangen, Germany) employing a standard RF filter module and MRI-compatible cables (*neuroConn GmbH, Ilmenau, Germany*). The technological, safety and noise tests, and methodological factors can be found in Supplementary materials (Table S1) on the basis of the ContES Checklist (Ekhtiari et al., 2022).

4.5.2.2.4. Modelling

Electromagnetic simulations were performed to identify optimised electrode placement and current steering parameters. Simulations were performed using the MIDA head model (Iacono et al., 2015), a detailed anatomical head model featuring 117 distinguished tissues and regions that was derived from multi-modal image data of a healthy female volunteer. Importantly, for brain stimulation modelling, the model distinguishes different scalp layers, skull layers, grey and white matter, CSF, and the dura. Circular electrodes (radius = 0.7 cm) were placed on the skin according to the 10-10 system and the electromagnetic exposure was determined using the ohmic-current-dominated electro-quasistatic solver from Sim4Life version 5.0 (ZMT Zurich MedTech AG, Switzerland), which is suitable due to the dominance of ohmic currents over displacement currents and the long wavelength compared to the simulation domain. Dielectric properties were assigned according to the IT'IS Tissue Properties Database v4.0 (Hasgall et al., 2022). Rectilinear discretisation was used, and grid convergence as well as solver convergence analyses were performed to ensure negligible numerical uncertainty resulting in a grid, which contained over 54M voxels selected for the simulations. Dirichlet voltage boundary conditions were applied, followed by current normalization, and the electrode-head interface contact was treated as ideal. tTIS exposure was quantified according to the maximum modulation envelope magnitude formula from Grossman et al., 2017. Subsequently, a sweep over 960 permutations of the four electrode locations was performed, considering symmetric and asymmetric montages with parallel (sagittal = 729 configurations; coronal = 231 configurations) or crossing current paths, while quantifying bilateral striatum (putamen, caudate and nucleus accumbens) exposure performance according to three metrics: 1) target exposure strength, 2) focality ratio (the volume ratio of target tissue above threshold and overall brain tissue above threshold;

a measure for stimulation selectivity), and (3) activation ratio (percentage of target volume above threshold; a measure for target coverage). From the resulting Pareto-optimal front, two configurations stood out particularly: one that maximized focality and activation (Pair 1: AF3 and AF4, Pair 2: TP7 and TP8 montage; focality = 30.3%, activation = 28.2%, threshold = 0.19 V/m) and a second one that accepts a reduction of these two metrics by a quarter, while increasing the target exposure strength by more than 50% (Pair 1: F3 and F4, Pair 2: TP7 and TP8; focality = 23.9%, activation = 22.1%, threshold = 0.31 V/m). The latter montage was selected, while the predicted tTIS field had larger stimulation intensity and to ensure that the target could actually be stimulated (Figure 1C, 1D).

4.5.2.2.5. Electrode positioning and evaluation of stimulation-associated sensations

Based on the modelling approach described above, we defined the stimulation electrode positions in the framework of the EEG 10-20 system (Seeck et al., 2017). The optimal montage leading to the best stimulation of the target structure, i.e. the bilateral striatum, was composed of the following electrodes: F3, F4, TP7 and TP8. Their locations were marked with a pen on the scalp and, after skin preparation (cleaned with alcohol), round conductive rubber electrodes of 3 cm² were placed adding a conductive paste (*Ten20, Weaver and Company, Aurora, CO, USA* or *Abralyt HiCl, Easycap GmbH, Woerthsee-Ettersschlag, Germany*) as interface with the skin. Electrodes were held in position with tape and oriented towards the top in order to allow good positioning inside the scanner. Impedances were checked and optimised until they were lower than 20 kΩ. Once good contact was obtained,

we tested different intensities of stimulation for each stimulation protocol in order to familiarise the participants with the perceived sensations and to systematically document them. tTIS_{Sham}, tTIS_{20Hz} and tTIS_{80Hz} were applied for 20 seconds with the following increasing current amplitudes per channel: 0.5 mA, 1 mA, 1.5 mA and 2 mA. Participants were asked to report any kind of sensation and, if a sensation was felt, they were asked to grade the intensity from 1 to 3 (light to strong) as well as give at least one adjective to describe it (Figure S2). After this step, cables were removed to be then replaced by MRI-compatible cables and a bandage was added to apply pressure on the electrodes and keep them in place. An impedance check was repeated in the MRI right before the training with concomitant stimulation and after the intervention.

4.5.2.3. MRI data acquisition

Structural and functional images were acquired using a 3T MAGNETOM PRISMA scanner (*Siemens, Erlangen, Germany*). T1-weighted images were acquired via the 3D MPRAGE sequence with the following parameters: TR = 2.3 s; TE = 2.96 ms; flip angle = 9°; slices = 192; voxel size = 1 × 1 × 1 mm, FOV = 256 mm. Anatomical T2 images were also acquired with the following parameters: TR = 3 s; TE = 409 ms; flip angle = 120°; slices = 208; voxel size = 0.8 × 0.8 × 0.8 mm, FOV = 320 mm. Finally, functional images were recorded using Echo-Planar Imaging (EPI) sequences with the following parameters: TR = 1.25 s; TE = 32 ms; flip angle = 58°; slices = 75; voxel size = 2 × 2 × 2 mm; FOV = 112 mm.

4.5.3. Data and statistical analyses

Data and statistical analyses were carried out with Matlab 2018a (the Mathworks, Natick, Massachusetts, USA) and the R software environment for statistical computing and graphics (R Core Team 2021, Vienna, Austria). Robust linear regressions were fitted with the Matlab function `robustfit`. Linear mixed models (LMM) were fitted using the `lmer` function of the `lme4` package in R (Bates et al., 2015). As random effects, we added intercepts for participants and block. Normality of residuals, skewness and homoscedasticity of the data were systematically tested, and logarithmic transformations were applied when necessary. To mitigate the impact of isolated influential data points on the outcome of the final model, we employed tools of the `influence.ME` package to detect and remove influential cases based on the following criterion: $\text{distance} > 4 * \text{mean distance}$ (Nieuwenhuis et al., 2012). Statistical significance was determined using the `anova` function with Satterthwaite's approximations of the `lmerTest` package (Luke, 2017). For specific post-hoc comparisons we conducted pairwise comparisons by computing estimated marginal means with the `emmeans` package with Tukey adjustment of p-values (Searle et al., 1980). Effect size measures were obtained using the `effectsize` package (Ben-Shachar et al., 2020). The level of significance was set at $p < 0.05$.

4.5.3.1. Behavioural data

The main goal of the present study was to evaluate the influence of striatal tTIS on reinforcement motor learning. To do so, we first removed trials, in which participants did not react within 1 s after the appearance of the cursor and target, considering that these extremely long preparation times may reflect significant fluctuations in attention (Derosière et al., 2015). This occurred extremely rarely (0.52 % of the whole data set). For each subject

and each trial, we then quantified the tracking Error as the absolute force difference between the applied and required force as done previously (Abe et al., 2011; Vassiliadis et al., 2021, 2022). Tracking performance during Training and Post-training trials were then normalised according to subjects' initial level by expressing the Error data in percentage of the average Pre-training Error for each block. In order to test our main hypothesis predicting specific effects of striatal tTIS on reinforcement motor learning, we performed a LMM on the Post-training data with $tTIS_{TYPE}$ and $Reinf_{TYPE}$ as fixed effects. We then also ran the same analysis on the Training data, to evaluate if striatal tTIS also impacted on motor performance, while stimulation was being delivered.

As a control, we checked that initial performance at Pre-training was not different between conditions with a LMM on the Error data obtained at Pre-training. Again, $tTIS_{TYPE}$ and $Reinf_{TYPE}$ were considered as fixed effects. Finally, another LMM was fitted with the fixed effect $tTIS_{TYPE}$ to verify that the amount of positive reinforcement (as indicated by a green target) in the $Reinf_{ON}$ blocks was similar across $tTIS_{TYPES}$.

4.5.3.2. fMRI data

4.5.3.2.1. Imaging Preprocessing

We analyzed functional imaging data using Statistical Parametric Mapping 12 (SPM12; *The Wellcome Department of Cognitive Neurology, London, UK*) implemented in MATLAB R2018a (*Mathworks, Sherborn, MA*). All functional images underwent a common preprocessing including the following steps: slice time correction, spatial realignment to the first image, normalization to the standard MNI space and smoothing with a 6 mm full-width

half-maximal Gaussian kernel. T1 anatomical images were then co-registered to the mean functional image and segmented. This allowed to obtain bias-corrected gray and white matter images, by normalizing the functional images via the forward deformation field. To select subjects with acceptable level of head movement, framewise displacement was calculated for each run. A visual check of both non-normalised and normalised images was performed in order to ensure good preprocessing quality. Finally, possible tTIS-related artifacts were investigated based on signal to noise ratio maps (see below).

4.5.3.2.2. Signal to Noise Ratio

Total signal to noise ratio (tSNR) maps were computed to check the presence of possible artifacts induced by the electrical stimulation. The values were calculated as mean over standard deviation of each voxel time series. Spherical regions of interest were then defined both underneath the tTIS electrodes and at 4 different locations, distant from the electrodes as a control. The center of each spherical ROI was obtained by projecting the standard MNI coordinates on the scalp (Okamoto et al., 2004) toward the center of the brain. After visual inspection of the ROIs, average tSNR maps were extracted and a LMM was used to compare signal to noise ratio underneath the electrodes and in the control regions (Figure S4).

4.5.3.2.3. Task-based BOLD activity analysis

A general linear model was implemented at the single-subject level in order to estimate signal amplitude. Eight regressors were included in the model: 6 head motion parameters (displacement and rotation) and normalised time series within the white matter and the corticospinal fluid. Linear contrasts were then computed to estimate specific activity during the motor task with respect to resting periods. Functional activation was also extracted within specific ROIs individually defined based on structural images. More specifically, the Free-surfer recon-all function was run based on the structural T1w and T2w images (<https://surfer.nmr.mgh.harvard.edu/>). The BNA parcellation was derived on the individual subject space and the selected ROIs were then co-registered to the functional images and normalised to the MNI space. BOLD activity within the individual striatal masks was averaged and compared between different striatal nuclei namely the putamen, caudate and NAc. Multiple comparison correction was applied at a cluster level by controlling for the False Discovery Rate (FDR).

4.5.3.2.4. Effective connectivity analyses

As an additional investigation, we computed task-modulated effective functional connectivity by means of the CONN toolbox 2021a (www.nitrc.org/projects/conn, RRID:SCR_009550) running in Matlab R2018a (*Mathworks, Sherborn, MA*). An additional denoising step was added by applying a band-pass filtering from 0.01 to 0.1 Hz and by regressing potential confounders (white matter, CSF and realignment parameters). After that, generalized Psycho-Physiological Interactions (gPPI) connectivity was extracted within specific pre-defined customised sub-networks: a reward and a motor network. The reward network was defined as following: two regions within the striatum, namely the NAc (BNA regions 223 and 224) and the ventro-medial putamen (BNA regions 225 and 226, left and

right respectively), and two frontal areas, namely the anterior cingulate (BNA regions 177, 179, 183 and 178, 180, 184, left and right respectively) and the orbitofrontal cortex within the vmPFC (BNA regions 41, 45, 47, 49, 187 and 42, 46, 48, 50, 188 for left and right respectively). The motor network included the following areas: the medial part of the SMA (BNA regions 9 and 10, left and right respectively) and the part of the M1 associated to upper limb function (BNA regions 57 and 58, left and right respectively). Notably, we considered connectivity in the left and right motor and reward networks regardless of laterality. Finally, gPPI was also extracted within a control language network, defined based on the functional atlas described by Shirer et al. (Shirer et al., 2012).

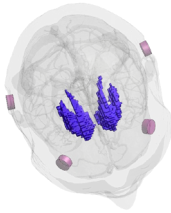

4.6. Supplementary material

Exclusion criteria

- Unable to consent
- Severe neuropsychiatric (e.g., major depression, severe dementia) or unstable systemic diseases (e.g., severe progressive and unstable cancer, life threatening infectious diseases)
- Severe sensory or cognitive impairment or musculoskeletal dysfunctions prohibiting to understand instructions or to perform the experimental tasks
- Color blindness
- Inability to follow or non-compliance with the procedures of the study
- Contraindications for NIBS or MRI:
 - Electronic or ferromagnetic medical implants/device, non-MRI compatible metal implant
 - History of seizures
 - Medication that significantly interacts with NIBS being benzodiazepines, tricyclic antidepressant and antipsychotics
- Regular use of narcotic drugs
- Left-handedness
- Pregnancy
- Request of not being informed in case of incidental findings
- Concomitant participation in another trial involving probing of neuronal plasticity.

ContES Checklist

Technological factors	
Manufacturer of MR Conditional Stimulator	DS5 Isolated Bipolar Constant Current Stimulator (Digitimer)
MR Conditional Electrode Details	Round, 3 cm ² conductive rubber electrodes
Electrode Positioning	<p>F3 → F4</p> <p>TP7 → TP8</p> <p>A bandage is warped around the head to apply pressure and keep the electrodes in place</p> <p>Electrodes are oriented in order to have vertical cables entering parallel to the MRI coil</p> <p>Head was fixed with pillows to avoid movements</p>
MR Conditional Skin-Electrode Interface	<p>10-20 gel</p> <p>One or two drops of saline were added when impedances were too high</p>
Amount of Contact Medium (Paste/Gel/Electrolyte)	Around 1mm of paste was manually placed on the electrodes
Electrode Placement Visualization	Pictures

	
RF Filter	NeuroConn DC-STIMULATOR MR RF filter module with MRI-compatible cables and electrodes
Wire Routing Pattern	<p>10 m ethernet cables between inner and outer box pass through a conduit along the wall of the MRI room until reaching the back of the MRI. Cables are then fixed with straps on the ground and on the wall of the MRI machine in order to avoid loops until reaching the interior of the coil.</p> <p>Cables between the head and the inner boxes were also fixed with straps and they were oriented in order to exit the magnetic field direction as soon as possible as indicated by the red arrows of the image below.</p> 
tES-fMRI Machine Synchronization/Communication	Stimulation was triggered by the stimulus delivery PC via parallel port to BNC cable. The parallel port of the stimulus delivery PC was connected to the DAQ controlling the stimulators.

	Stimulus delivery PC, in turn, was also receiving the scanner trigger from the scanner via USB port.
Safety and noise tests	
MR Conditionality Specifics for tES Setting	Please refer to Section “Methods-Imaging acquisition”
tES-fMRI Setting Test - Safety Testing	<p>Impedances were checked before and after the stimulation.</p> <p>No temperature tests were performed during the experiment.</p> <p>Intensity titration was performed prior to entering the MRI, testing increasing currents (0.5, 1, 1.5 and 2 mA) and asking the subject to report any type of sensation.</p> <p>A sensation questionnaire was also performed at the end of the experiment.</p>
tES-fMRI Setting Test - Subjective Intolerance Reporting	No intolerances were reported by any subject
tES-fMRI Setting Test - Noise/Artifact	Signal to Noise Ratio (SNR) analysis was performed on the fMRI images, please refer to Section “Methods-Signal to Noise Ratio”
Impedance Testing	<p>Impedances were checked right after electrodes positioning outside the scanner, before and after the stimulation inside.</p> <p>One or two drops of saline solution were</p>

	added if impedances were higher than 20k Ω
Methodological factors	
Concurrent tES-fMRI Timing	<p>For timings, please refer to the “Methods-Stimulation protocols” section</p> <p>To mitigate the impact of potential carry-over effects on our experimental results we used the following strategy:</p> <ol style="list-style-type: none"> 1) We stimulated for short periods in each condition (5 minutes interspersed with resting periods without stimulation; see “Methods-Stimulation protocols”); 2) We imposed breaks (~7-8 minutes) between each stimulation protocol; 3) We randomised the order of the Stimulation conditions
Imaging Session Timing	All sequences were performed with T1 stimulation electrodes placed on the subjects’ head.
tES Experience Report	Please refer to “Results” section and to Figure S2.

Table S1. ContES checklist as recommended in Ekhtiari et al., 2022 for concurrent tES-fMRI studies.

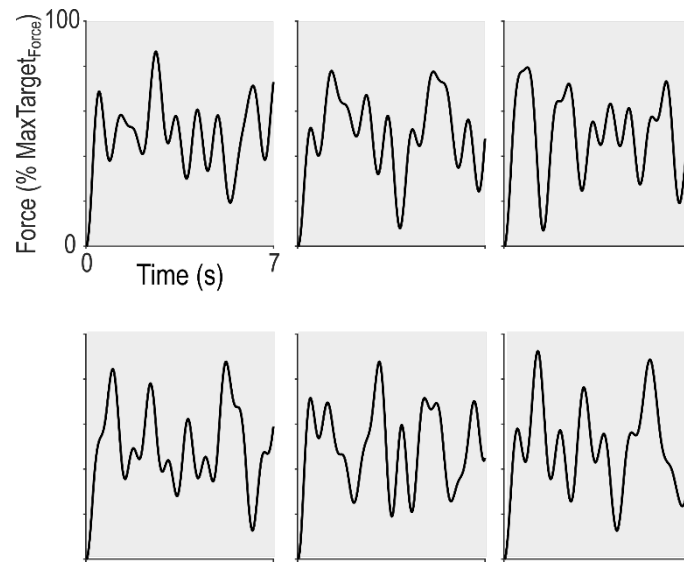


Figure S1. Patterns of motion of the target. For each block of training, participants had to learn a new pattern of motion of the target. The patterns had similar mathematical properties and their relationship to a condition was randomised (see Methods for more details).

Control analyses of behavioural data

Pre-training performance

In order to verify that our main behavioural results were not influenced by potential differences in initial performance between conditions despite randomization, we analysed the Error at Pre-training between conditions. We did not find any $tTIS_{TYPE}$ ($F_{(2,519.15)} = 1.64$; $p = 0.20$) or $tTIS_{TYPE} \times Reinf_{TYPE}$ effect ($F_{(2,519.99)} = 1.08$; $p = 0.34$), suggesting that the main behavioural results could not be accounted for by differences in initial performance between conditions. However, the LMM did reveal a $Reinf_{TYPE}$ effect ($F_{(1,519.15)} = 12.47$; $p < 0.001$), that was due to the fact that Pre-training performance was generally better in $Reinf_{OFF}$ blocks. This effect, which was opposite to our learning results (generally better learning with $Reinf_{ON}$), may be related to an expectancy effect stemming from the repetitive structure of the reinforcement conditions (see Methods). However, the absence of interaction with $tTIS_{TYPE}$ is strongly suggestive that this effect did not drive any of the main findings. Put together,

these data provide confidence that the differential effects of striatal tTIS on motor learning depending on the presence of reinforcement were not the result of different initial performance between conditions.

Success rate

Overall, the amount of positive reinforcement (i.e., when the target was green) averaged $52.78 \pm 0.42\%$ and was comparable across tTIS_{TYPE} ($F_{(2,1702)} = 0.17$; $p = 0.84$), suggesting that the closed-loop reinforcement schedule was successful at providing similar reinforcement feedback despite differences in performance between conditions. Hence, different success rates during training cannot explain the effect of the different striatal tTIS conditions on motor learning.

Frequency of flashing

Analysis of the frequency of flashing in the different conditions did not reveal any effect of tTIS_{TYPE} ($F_{(2,3283)} = 0.85$; $p = 0.43$) nor any Reinf_{TYPE} \times tTIS_{TYPE} interaction ($F_{(2,3283)} = 0.19$; $p = 0.82$), suggesting that the behavioural effects of tTIS could not be explained by a visual confound. However, this analysis did reveal a Reinf_{TYPE} effect ($F_{(1,3283)} = 33.62$; $p < 0.001$) which was due to the fact that the average frequency in the Reinf_{OFF} condition (4.28 ± 0.097 Hz) was slightly but significantly higher than with Reinf_{ON} (4.08 ± 0.098 Hz; $F_{(1,3283)} = 33.62$; $p < 0.001$). Notably, in absolute terms, this difference represented only a difference of 1.4 change of color over the whole 7 s trial, which we think is unlikely to explain the improvement of performance in the Reinf_{ON} condition.

Blinding integrity and tTIS-evoked sensations

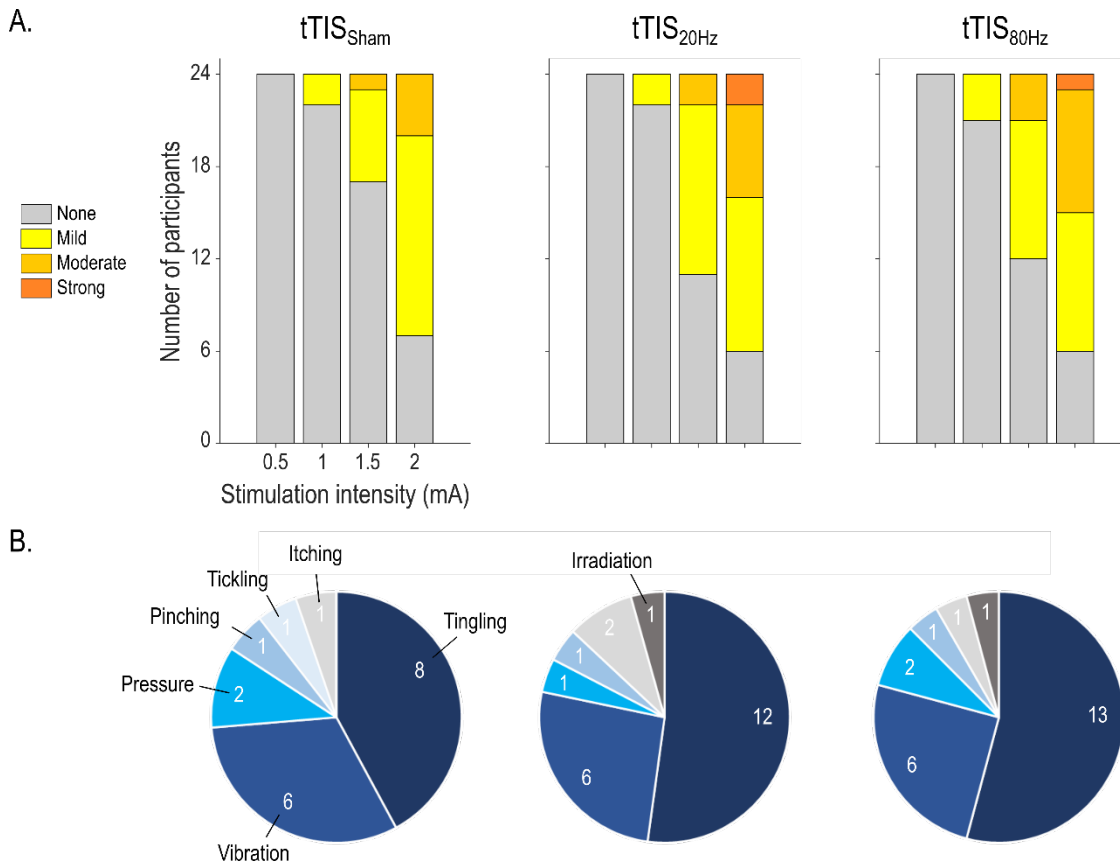


Figure S2. tTIS-related sensations. A) Magnitude of tTIS-related sensations. Magnitude of sensations reported before the experiment for current amplitudes ranging from 0.5 to 2 mA for each tTIS_{TYPE}. The current amplitude used in the present experiment was 2 mA. **B) Types of tTIS-related sensations.** Type of sensations as described by the participants, at 2 mA. Note that subjects were allowed to describe their sensations with up to two different words.

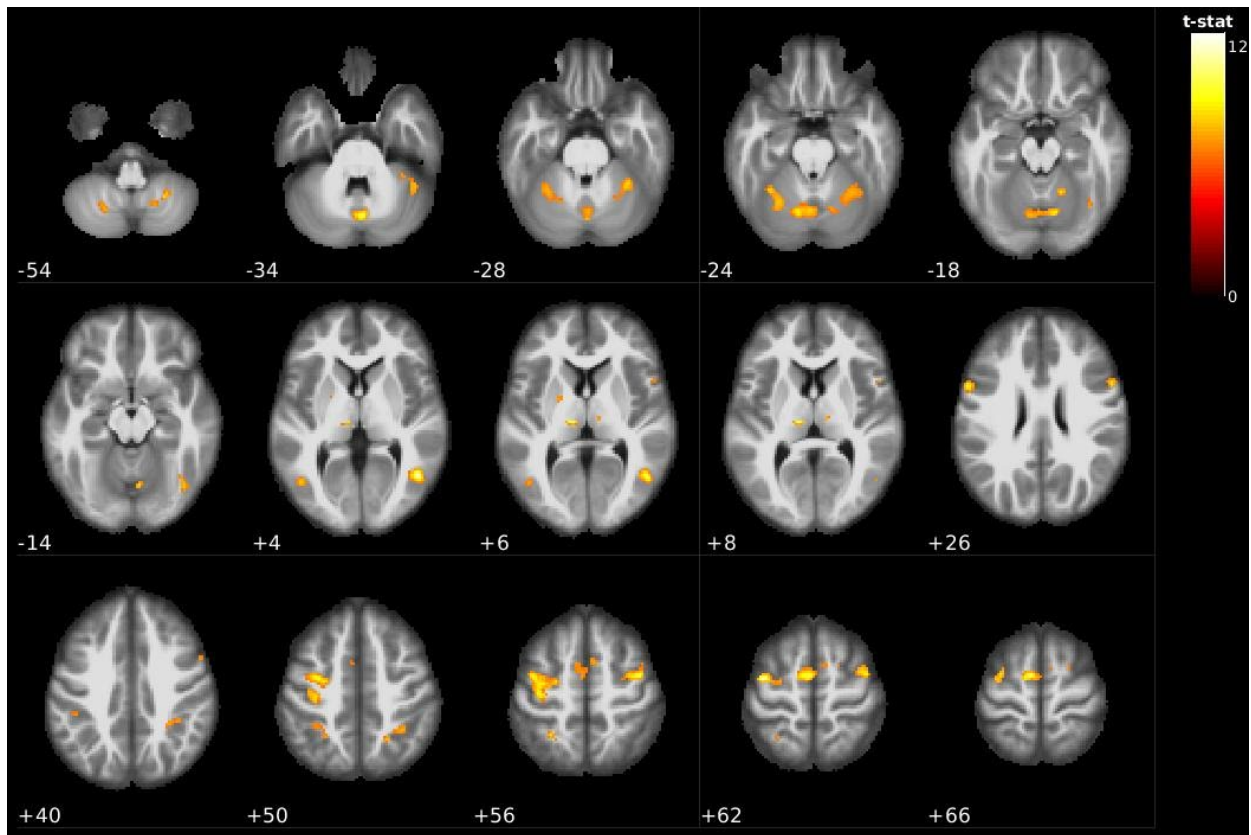
Brain activity during reinforcement motor learning

Figure S3. Whole-brain activity during reinforcement motor learning. Activation maps for the contrast task>rest in the tTIS_{Sham}, Reinf_{ON} condition showing activation of key areas of the reinforcement motor learning network including the putamen, thalamus, cerebellum and sensorimotor network, especially on the left side. Significant clusters are shown for uncorrected voxel-wise family wise error (FWE), $p=0.001$, and corrected cluster-based false discovery rate (FDR), $p = 0.05$.

Cluster-level				Peak-level					x	y	z	Region
pFWE-corr	qFDR-corr	kE	P _{uncorr}	pFWE-corr	qFDR-corr	T	(ZE)	P _{uncorr}				
<0.001	<0.001	135	<0.001	<0.001	0.005	12.63	6.84	<0.001	46	-62	4	Temporal_Mid_R
<0.001	<0.001	523	<0.001	<0.001	0.005	12.32	6.77	<0.001	-40	-8	62	Precentral_L
				<0.001	0.021	10.62	6.33	<0.001	-34	-6	52	Postcentral_L
				<0.001	0.021	10.43	6.28	<0.001	-36	-20	54	Precentral_L
<0.001	<0.001	335	<0.001	<0.001	0.018	11.08	6.46	<0.001	-8	-6	64	Supp_Motor_Area_L
				0.003	0.145	8.21	5.56	<0.001	6	6	58	Supp_Motor_Area_R
				0.003	0.145	8.20	5.55	<0.001	-4	-2	54	Supp_Motor_Area_L
<0.001	<0.001	44	<0.001	<0.001	0.021	10.65	6.34	<0.001	-10	-20	6	Thal_IL_L
<0.001	<0.001	162	<0.001	<0.001	0.021	10.36	6.26	<0.001	42	-6	56	Frontal_Mid_2_R
				<0.001	0.042	9.48	5.99	<0.001	34	-4	58	Frontal_Sup_2_R
<0.001	<0.001	175	<0.001	<0.001	0.021	10.27	6.23	<0.001	-58	10	28	Precentral_L
				<0.001	0.037	9.60	6.03	<0.001	-56	8	20	Frontal_Inf_Oper_L
				0.019	0.490	7.32	5.21	<0.001	-48	2	16	Rolandic_Oper_L
<0.001	<0.001	601	<0.001	<0.001	0.024	10.06	6.17	<0.001	2	-74	-34	Vermis_7
				<0.001	0.025	9.99	6.15	<0.001	-12	-70	-22	Cerebellum_6_L
				<0.001	0.027	9.88	6.12	<0.001	12	-70	-20	Cerebellum_6_R
<0.001	<0.001	82	<0.001	<0.001	0.070	9.14	5.88	<0.001	56	10	26	Frontal_Inf_Oper_R
				0.006	0.234	7.86	5.42	<0.001	56	10	38	Precentral_R
<0.001	<0.001	141	<0.001	0.001	0.092	8.89	5.80	<0.001	-34	-52	-24	Cerebellum_6_L
				0.002	0.117	8.47	5.65	<0.001	-28	-62	-24	Cerebellum_6_L
<0.001	<0.001	76	<0.001	0.001	0.092	8.87	5.79	<0.001	-28	-52	56	Parietal_Sup_L
				0.011	0.341	7.57	5.31	<0.001	-30	-44	48	Parietal_Inf_L
<0.001	<0.001	200	<0.001	0.001	0.092	8.77	5.76	<0.001	32	-48	-28	Cerebellum_6_R
				0.013	0.382	7.49	5.28	<0.001	34	-40	-34	Cerebellum_6_R
<0.001	<0.001	36	<0.001	0.001	0.092	8.73	5.74	<0.001	16	-54	-18	Cerebellum_4_5_R
<0.001	<0.001	28	<0.001	0.001	0.101	8.63	5.71	<0.001	26	-58	-54	Cerebellum_8_R
<0.001	<0.001	62	<0.001	0.001	0.113	8.51	5.67	<0.001	38	-62	-16	Fusiform_R
				0.002	0.117	8.45	5.64	<0.001	42	-72	-12	Occipital_Inf_R
<0.001	<0.001	21	<0.001	0.002	0.117	8.41	5.63	<0.001	-46	-68	4	Occipital_Mid_L
<0.001	<0.001	141	<0.001	0.002	0.130	8.33	5.60	<0.001	22	-56	50	Location not in atlas
				0.002	0.130	8.30	5.59	<0.001	30	-48	48	Parietal_Sup_R
				0.007	0.266	7.76	5.39	<0.001	36	-40	42	SupraMarginal_R
<0.001	<0.001	29	<0.001	0.004	0.170	8.09	5.51	<0.001	44	-50	-34	Cerebellum_Crus1_R
<0.001	<0.001	59	<0.001	0.004	0.178	8.04	5.49	<0.001	-22	-66	-52	Cerebellum_8_L
<0.001	0.006	12	0.003	0.004	0.190	7.99	5.47	<0.001	10	-16	8	Thal_MDI_R
0.001	0.043	6	0.028	0.009	0.319	7.63	5.33	<0.001	-22	-2	6	Putamen_L
<0.001	<0.001	34	<0.001	0.009	0.319	7.63	5.33	<0.001	18	-64	-54	Cerebellum_8_R
0.001	0.300	7	0.019	0.023	0.545	7.23	5.17	<0.001	20	2	62	Frontal_Sup_2_R
0.001	0.030	7	0.019	0.024	0.560	7.21	5.16	<0.001	52	12	8	Frontal_Inf_Oper_R
0.001	0.030	7	0.019	0.025	0.568	7.19	5.16	<0.001	-44	-36	40	Parietal_Inf_L

Table S2: Significant clusters and the respective local maxima in the tTIS_{Sham}, Reinf_{ON} condition. Regions were identified with the Automated Anatomical Labelling atlas 3 (AAL3, Rolls et al., 2020). Significant clusters were selected for uncorrected voxel-wise family wise error (FWE), $p=0.001$, and corrected cluster-based false discovery rate (FDR), $p = 0.05$.

Correlation between effect of tTIS_{80Hz} on reinforcement motor learning and modulation of whole-brain activity

Cluster-level				Peak-level					x	y	z	Region
pFWE-corr	qFDR-corr	kE	P _{uncorr}	pFWE-corr	qFDR-corr	T	(Z _E)	P _{uncorr}				
0.003	0.005	157	<0.001	0.027	0.065	7.29	5.14	<0.001	10	18	0	Caudate_R
				0.639	0.678	5.38	4.25	<0.001	0	0	10	Location not in atlas
				0.921	0.757	4.89	3.98	<0.001	6	6	2	Location not in atlas
0.007	0.005	138	<0.001	0.693	0.678	5.30	4.21	<0.001	-16	14	6	Location not in atlas
				0.923	0.757	4.88	3.98	<0.001	-22	14	-2	Putamen_L
				1.000	0.810	4.26	3.60	<0.001	-18	8	-6	Putamen_L

Table S3. Significant clusters for the correlation between the behavioural and neural effects of tTIS_{80Hz} (vs. tTIS_{20Hz}). Two significant clusters were found with several local maxima. Notably, the left cluster also encompassed a portion of the left caudate (related to Figure 3). Regions were identified with the Automated Anatomical Labelling atlas 3 (AAL3, Rolls et al., 2020). Significant clusters were selected for uncorrected voxel-wise family wise error (FWE), $p=0.001$, and corrected cluster-based false discovery rate (FDR), $p = 0.05$.

Correlation between impulsivity and modulation of whole-brain activity with tTIS_{80Hz}

Cluster-level				Peak-level					x	y	z	Region
pFWE-corr	qFDR-corr	kE	P _{uncorr}	pFWE-corr	qFDR-corr	T	(Z _E)	P _{uncorr}				
<0.001	<0.001	254	<0.001	0.707	0.524	5.29	4.20	<0.001	-8	0	18	Location not in atlas
				0.719	0.524	5.27	4.19	<0.001	-14	16	16	Caudate_L
				0.971	0.620	4.72	3.88	<0.001	-16	16	0	Location not in atlas

Table S4. Significant clusters for the correlation between impulsivity and neural effects of tTIS_{80Hz} (vs. tTIS_{20Hz}). One significant cluster encompassing the left caudate nucleus was found (related to Figure 5). Regions were identified with the Automated Anatomical Labelling atlas 3 (AAL3, Rolls et al., 2020).

Imaging quality control

A threshold of 0.5 was chosen to discard subjects showing more than 40% of voxels with framewise displacement FD higher than this threshold. In the current study cohort, no subject exceeded the limit value, thus the whole dataset could be used. Furthermore, successful cleaning of the data was ensured by visual checking the preprocessing results. In particular, good registration between anatomical and functional images and normalization to standard space were checked.

Signal to noise ratio analysis showed significantly higher tSNR values underneath the stimulating electrodes ($F_{(1,1122)} = 249.25$, $p < 0.001$; **Figure S4**). This result suggests that the

stimulation did not introduce additional noise to the MR images. In summary, all controls confirmed the good quality of the imaging data.

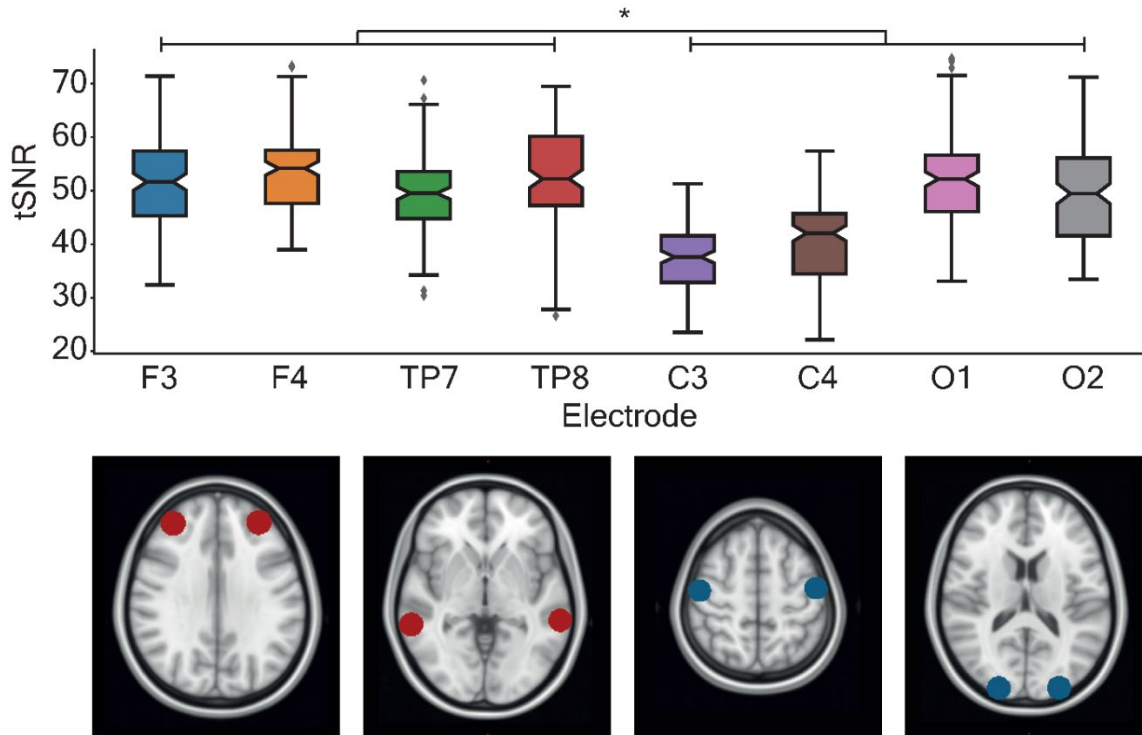


Figure S4. Total signal to noise ratio (tSNR). Total signal to noise ratio investigation. On the top panel, the average tSNR is shown within spheres of 10mm radius underneath the 4 stimulation electrodes (F3, F4, TP7 and TP8) and underneath other 4 locations more distal from the electrodes (C3, C4, O1 and O2). A significant higher tSNR was found underneath the electrodes with respect to the distal locations ($F_{(1,1122)} = 249.25$, $p < 0.001$). This indicates that there was no reduction of the tSNR due to the presence of electrical current. On the bottom panel, the location of the spheres from where the average tSNRs were extracted: F3 and F4 in red in the first image from the left, TP7 and TP8 in red on the second image from the left, C3 and C4 in blue on the third image from the left, O1 and O2 in blue on the forth image from the left.

5. Discussion

5.1. Main conclusions

In this PhD, I have investigated some of the mechanisms at play during reinforcement learning of motor skills, with the ultimate goal to favor incorporation of reinforcement feedback in future neuro-rehabilitation strategies.

In **Study 1**, I found that beyond reinforcement feedback, motivation by reward strongly impacts motor learning and leads to gains in performance that are maintained after training, in the absence of reward and even 24h later. Such effects were accompanied by a persistent up-regulation of motor variability based on previous reinforcements. Hence, when motivated by reward during training, participants relied more strongly on reinforcement information and this persisted 24h later in the absence of reward. This result suggests that associating a training with a particular motivational context is a powerful tool to durably enhance motor learning.

In **Study 2**, I asked if the timing at which rewards are provided with respect to movements, which is a crucial modulator of reinforcement learning in decision-making tasks, could also influence motor learning. I found that reward timing impacted on the dynamics and consolidation of learning. Training with short reward delays led to continuous, linear gains in performance while long reward delays led to fast initial learning followed by an early plateau in performance and to an overall reduced performance in the end of training. Moreover, skill consolidation was altered following training with long reward delay in learners. Overall, this study shows that reward timing matters in motor learning, potentially by modulating the engagement of different neural systems involved in learning (Foerde and Shohamy, 2011).

Finally, in **Study 3**, I used transcranial temporal interference stimulation, an innovative brain stimulation approach allowing to target deep brain structures non-invasively in humans. I exploited this technique to question the causal role of the striatum in reinforcement-related improvements in motor learning and found a specific disruption of reinforcement motor learning when stimulating in the high gamma range, a frequency band previously associated to reward processing in the striatum (Berke, 2009). These behavioral effects

were related to a specific modulation of activity in the caudate nucleus and putamen, pointing towards the idea that the disruption of reinforcement motor learning was indeed caused by neuromodulation of the striatum. In that study, I also show that high gamma tTIS enhances effective connectivity between striatum and frontal cortex, a pathway known to be crucial for reward processing. Such hyper-connectivity was particularly pronounced between the striatum and motor cortex and in the presence of reinforcement. Finally, the neural effects of the stimulation were consistently correlated to impulsivity levels of participants, suggesting that impulsivity could constitute a biomarker of the neural sensitivity to high gamma tTIS.

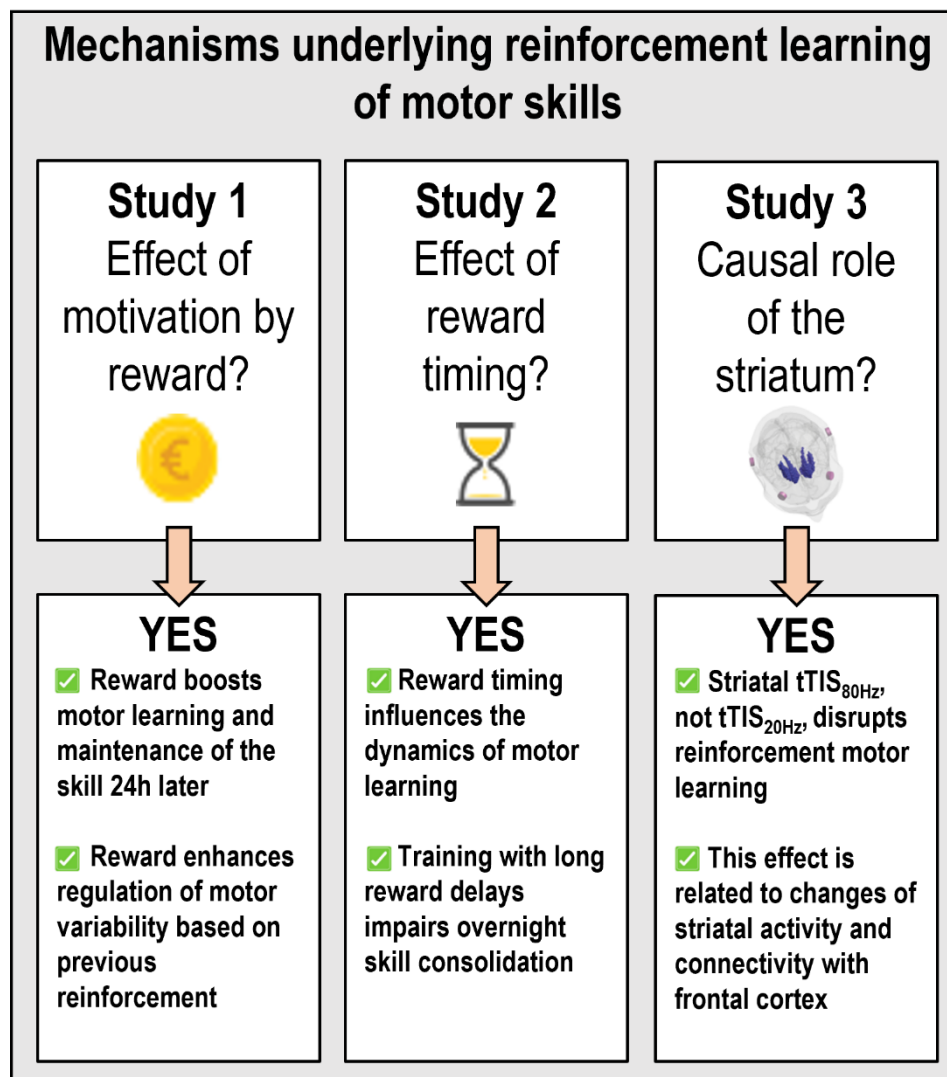


Figure 1. Main conclusions of the studies presented in this PhD.

5.2. General discussion

In this section, I discuss the main conclusions of the present work in light of the current literature and how they improve our mechanistic understanding of reinforcement learning of motor skills and may guide the development of future rehabilitation protocols.

5.2.1. Motivational interventions during motor learning

A central question of this work was to better understand whether motivation could impact on motor learning in addition to reinforcement feedback and the respective underlying mechanisms. In this regard, the results of Study 1 showed that there is a clear advantage in combining reinforcement feedback with motivational interventions (Vassiliadis et al., 2021). We have used monetary rewards in Study 1, but other motivational interventions have been proposed for motor learning including social praise (Sugawara et al., 2012), boosting sense of agency during practice (Lewthwaite and Wulf, 2017) or gamification (Doumas et al., 2021). Our design allowed us to isolate the net effect of motivation on motor learning and identify potential mechanisms. One of those mechanisms is the specific boosting of motor adjustments based on reinforcement feedback. Put differently, participants tended to increase more their motor variability following failure and reduce it more following success, indicating an improved use of exploration and exploitation, respectively, based on previous outcomes (Dhawale et al., 2017). Such increased reliance on reinforcement feedback, which was associated to improved learning at the single-trial level, even persisted 24h later, in the absence of training. This strongly points towards the idea that the long-term effect of motivational interventions (Abe et al., 2011) may be related to a durable sensitization to reinforcement feedback. This idea seems promising for rehabilitation, because it indicates that combining rehabilitation with short motivational interventions (e.g., during sessions performed at the clinic) could have positive effects on future training performed in neutral contexts (e.g., at home). Moreover, the data also show that this regulation of motor variability did not concern all motor components, but rather the most important component for task success in the task (i.e., the amplitude of the force), potentially reflecting the operation of a credit assignment mechanism (McDougle et al., 2016, 2018), allowing to adjust motor components based on their estimated weight in the outcome. Hence, an interesting

next step could be to evaluate whether specific motor components can be preferentially improved by modulating their weight in the computation of the reinforcement feedback. Such finding could pave the way to interventions individualizing reinforcement feedback based on the specific deficit experienced by a patient.

An important aspect of Study 1 is that we studied the effect of reward in a situation where it was coupled to reinforcement feedback. Hence, our results on their own do not allow to conclude definitively on whether reward needs to be coupled to reinforcement to be beneficial in motor learning, because we did not test any group with random reward (i.e., not related to performance). However, our single-trial results showing that reward boosted specifically the regulation of motor variability based on reinforcement information suggests that this coupling is indeed a crucial factor. Consistently, a recent study which compared performance-based and random reward found that the association between correct performance and reward is necessary to observe reward-related gains in motor learning (Sporn et al., 2022). Hence, providing rewards that depend on performance appears as a promising way to improve motor learning, and potentially recovery of motor function after injury of the motor system.

5.2.2. Reward timing: what's next?

In Study 2, we found that the delay between movement and reward plays a prominent role in motor skill acquisition and consolidation. This implies that the rapid mapping between a movement (and its sensory consequences) and the ensuing reward is a crucial aspect of reinforcement motor learning, potentially because it facilitates the coupling between a somatosensory sensation and the ensuing reward (Sidarta et al., 2018). Another potential explanation for the alteration of learning when training with long reward delays could be a reduction in the precision of striatal RPEs (Fiorillo et al., 2008; Kobayashi and Schultz, 2008). Study 2 also showed very different learning dynamics depending on reward timing: while training with short reward delay led to slow, continuous improvement of performance, training with long reward delays allowed fast initial learning that quickly plateaued. Hence, short reward delays may be generally preferred for rehabilitation, except in the case of very short sessions in which long reward delays may provide an added value. Characterizing the brain regions involved in learning from short and long reward delays appears as a crucial

next step to be able to individualize reward delays based on the brain networks affected by a specific pathology. As such, previous associative learning studies have shown dissociable roles of the striatum and hippocampus in learning from short and long reward delays, respectively (Foerde and Shohamy, 2011; Foerde et al., 2013). More specifically, this body of work has revealed that Parkinson's disease patients who suffer from striatal dysfunction were unable to learn from short reward delays (Foerde and Shohamy, 2011; Foerde et al., 2012) whereas amnesic patients suffering from hippocampal lesions could not learn from long reward delays (Foerde et al., 2013). Whether such dissociation also holds in the context of motor learning is not known, and could constitute an interesting future question. Overall, individualization of reward timing based on a patient's lesion may be an important step to optimize reinforcement motor learning.

Another exciting future direction of this work would be to evaluate the effect of the variability in reward timing on motor learning and the associated neural correlates. As such, previous research has shown that RPEs in the VTA are modulated by the predictability of reward timing with stronger RPEs being associated to less predictable timings (Klein-Flügge et al., 2011). Hence, reducing predictability of reward timing may improve motor learning by boosting dopaminergic RPEs, a hypothesis that could be tested in future work.

5.2.3. Causal role of the striatum in reinforcement learning of motor skills

Following Study 1 and Study 2, we have hypothesized that the striatum may be a crucial hub for translation of (short-delayed) reinforcement information into motor adjustments (Piray et al., 2017; Hori et al., 2019). Especially, the striatum is densely connected to multiple frontal areas involved in reward processing and motor learning that could contribute to the beneficial effect of reinforcement on motor learning. To evaluate the causal role of the striatum in this process, I used tTIS, a new non-invasive approach allowing to stimulate deep brain structures. Interestingly, our results highlight that a specific striatal mechanism, likely relying on high gamma oscillations, is involved in the beneficial effect of reinforcement on motor learning. Interestingly, the data also suggest that another striatal mechanism, relying on beta oscillations, is preferentially involved in sensory-based motor learning. This dissociation between the mechanisms involved in sensory- and reinforcement-based motor learn-

ing confirms that these two types of feedback preferentially recruit different neural mechanisms (Sidarta et al., 2016; Therrien et al., 2016; Mathis et al., 2017; Uehara et al., 2018) and further suggests that these neural mechanisms are associated to differential oscillatory activity within the striatum. Thanks to concurrent fMRI acquired during training and stimulation, we were able to evaluate the neurophysiological effects of tTIS. We found that the detrimental effect of high gamma tTIS on reinforcement motor learning was related to modulation of activity in the caudate nucleus and putamen, and also to a specific increase in effective connectivity between motor striatum and motor cortex. To my knowledge, this constitutes the first evidence that striatal activity can be modulated non-invasively by means of brain stimulation in humans with relevant behavioral and neurophysiological effects. More specifically, this study shows that reinforcement motor learning causally relies on functional integrity of the striatum and related fronto-striatal circuits, as previously suggested through associative approaches (Wachter et al., 2009; Widmer et al., 2016). As I argue below, such causal link between striatal integrity and reinforcement motor learning could be useful to stratify stroke patients based on the localization of their lesion. Overall, the present work shows that reinforcement motor learning is highly sensitive to motivation experienced during training, to the timing of the reinforcement feedback and relies on functional integrity of the striatum.

5.3. Perspectives

In this section, I combine the present work with previous literature to propose a framework for the integration of reinforcement feedback into clinical rehabilitation. In particular, I present preliminary results of three studies initiated during my PhD that support the necessity to incorporate reinforcement in a personalized manner. Finally, I also discuss some ideas to improve our understanding of the mechanisms underlying tTIS and propose perspectives to move this emerging field forward.

5.3.1. Towards incorporation of reinforcement into motor rehabilitation: optimization of training features and personalization

Contingency, stochasticity and valence of reinforcement feedback

Along with previous studies, this work aimed at delineating key training parameters that can guide the development and individualization of reinforcement-based motor rehabilitation procedures. As mentioned above, our results support the use of motor component-specific motivational interventions and short-delay rewards during motor rehabilitation. Additionally, previous studies have characterized other features of the reinforcement feedback that may impact the learning process. First, it was shown that it is important to provide reinforcement that is contingent on the learner's performance, rather than random feedback (Sporn et al., 2022). This is consistent with the idea that reinforcements are useful to inform a learning process that uses outcomes to adjust future movements (Dhawale et al., 2017). Moreover, the uncertainty on the presence of reinforcement (sometimes referred to as the stochasticity of reinforcement) can substantially improve motor learning (Dayan et al., 2014a). An efficient way to combine individualized performance-based feedback and stochasticity is through the use of closed-loop reinforcement schedules (Therrien et al., 2016, 2020). This method constantly adapts the motor criterion to consider a trial as successful by taking into account recent performance. As shown in Study 3, this procedure gave us success rates of around 52%, ensuring almost maximum uncertainty on the outcome (in line with the stochasticity principle), while providing personalized veridical feedback. An additional advantage of this method, also observed in Study 3, is that it allows to match success rates between different experimental conditions even in the case of differences in motor performance, an aspect that will be crucial for future clinical trials involving reinforcement feedback.

Another factor that also influences reinforcement motor learning is valence (i.e., whether a correct trial yields a reward or an absence of punishment and vice versa for incorrect trials). There is evidence that even though the amount of information provided by the reinforcement feedback is exactly the same in the reward and punishment versions of the task, people learn differently (Abe et al., 2011; Galea et al., 2015; Steel et al., 2016; Quattrocchi et al., 2017) and the brain regions involved in learning differ (Steel et al., 2019). More specifically, punishment has been suggested to accelerate learning while reward

would mainly impact on long-term retention (Galea et al., 2015). Importantly though, such effects have been shown to depend on the motor learning task performed (Steel et al., 2016). Even if we did not test any condition involving punishment, we found that reward could enhance skill acquisition in our force modulation task and that such gains in performance were maintained 24h later, in line with the aforementioned positive effects of reward on retention. Moreover, a previous motor adaptation study did not show any added value of using punishment compared to reward in stroke patients (Quattrocchi et al., 2017). Hence, the benefits of using punishment for motor rehabilitation are not clear yet, especially given the potential negative consequences of punishment-based training on patients' mood and levels of anxiety.

The role of sensory uncertainty in reinforcement motor learning

The three studies presented in this PhD evaluated the effect of reinforcement feedback, but always in a context of high visual feedback uncertainty about the cursor being controlled. As explained above, this manipulation was used based on previous studies suggesting that participants used more reinforcement information when sensory feedback was uncertain (Izawa and Shadmehr, 2011; Cashaback et al., 2017). Notably, such situation may be relevant in clinical contexts since patients suffering from motor impairments often also present sensory deficits (Connell, 2008; Hepworth et al., 2016). Moreover, many daily-life activities involve the precise control of force without direct visual feedback about the pressure being applied (e.g., such as when adjusting the force applied on the accelerator pedal when driving, or controlling the force applied when carrying a fragile object, Clemente et al., 2016). However, there is still a need to understand exactly how sensory and reinforcement feedback interact, and if people can learn more from reinforcement in high uncertainty contexts. Preliminary analyses on another dataset acquired in the context of my PhD (in collaboration with the group of Prof. Micera from EPFL) suggests that this might be the case (**Figure 2**, Vassiliadis et al., 2022 (in prep)). As such, when testing young healthy participants on the same task as in Study 3, with varying levels of visual uncertainty and with or without reinforcement (**Figure 2A, B, C**), we found that reinforcement feedback elicited larger benefits both in tracking performance (during training, **Figure 2D, E**) and in learning (at Post-training, **Figure 2F**) in a context of high sensory uncertainty. This provides direct evidence for the idea that the benefits of reinforcement feedback on motor learning directly depend

on the quality of the sensory feedback experienced during training. In subsequent experiments, we aim at evaluating whether these results can generalize when uncertainty does not only come from vision, but also from somatosensory information, a situation experienced by patients suffering from somatosensory deficits such as in the context of peripheral neuropathy or limb amputation (Raspopovic et al., 2014).

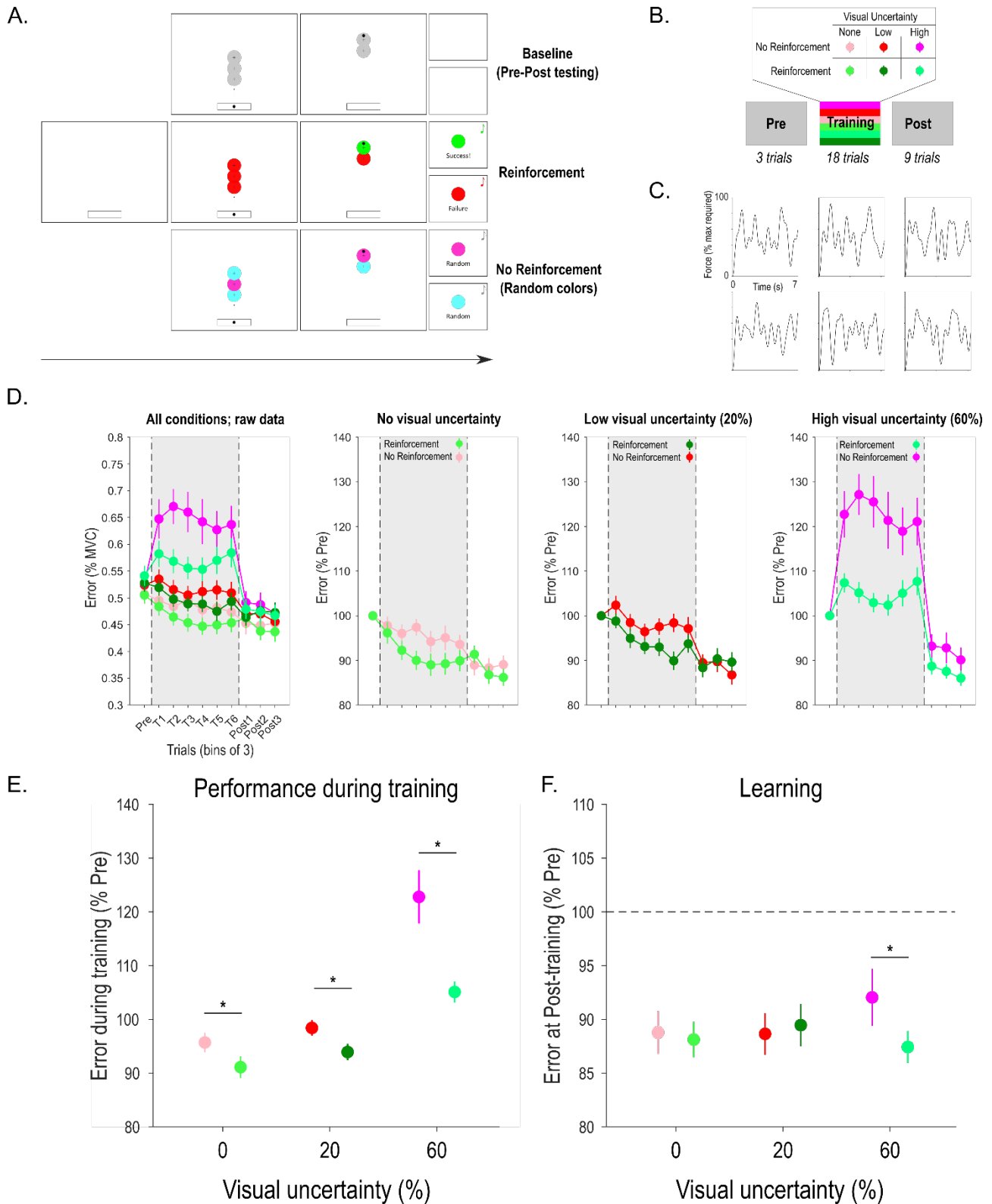


Figure 2. Effect of visual uncertainty on reinforcement motor learning. **A) Representation of the force-tracking task.** Participants ($n = 24$) had to follow a moving target by adjusting the force applied on a hand-grip force sensor. In separate blocks of training they had to practice either with reinforcement feedback indicating real-time performance or with random feedback. Visual uncertainty was manipulated by providing full visual feedback (0% uncertainty) or by removing vision of

the cursor for 21.4% (1.5 / 7s; 20% condition) or 64.3% (4.5 / 7s; 60% condition) of the total duration of the trial. Pre- and Post evaluation were performed with full visual feedback and no reinforcement (grey target). **B) Structure of the learning blocks.** Participants performed 6 learning blocks (one per condition) in a randomized order. **C) Sequences of motion of the target.** Each experimental block was associated to a new pattern of movement of the target that had similar mathematical properties and were selected based on pilot experiment to have similar difficulty. **D) Learning curves.** Performance was evaluated based on the tracking Error (i.e., the absolute force difference between the target force and the applied force). On the left side, we show raw Error data in each condition, averaged by bins of 3 trials. The three plots on the right side show evolution of the Error in the three uncertainty conditions in the 0%, 20% and 60% visual uncertainty condition. Note the large disruption of tracking with high visual uncertainty that was partially counteracted by reinforcement. **D) Effect of reinforcement on motor performance.** A significant Visual uncertainty x Reinforcement interaction in the LMM was found and related to the fact that reinforcement elicited larger gains in performance in the high uncertainty condition. **E) Effect of reinforcement on motor learning.** Again, a significant Visual uncertainty x Reinforcement was found and was due to the fact that reinforcement significantly improved motor learning but only in the presence of high sensory uncertainty. This result also replicates the finding of reinforcement-related benefits in motor learning under high sensory uncertainty observed in the tTIS_{Sham} condition in Study 3.

Apathy as a potential predictor of reinforcement-related motor gains

Another important aspect of reinforcement motor learning is the large inter-individual variability in the response to performance feedback, suggesting that individual factors shape how much people can benefit from reinforcement during motor learning (Holland et al., 2019). Such large inter-individual variability was also observed in the three studies presented here (note the reduction of inter-individual variability in the presence of reward in Study 1 though). Put differently, some individual factors may determine how much a given participant could benefit from a reinforcement-based training. Hence, an important line of research for future clinical translation is to characterize those factors to be able to predict which patients could benefit from reinforcement-based interventions.

A factor that could strongly influence responsiveness to a reinforcement-based training is apathy. Apathy is a syndrome of impaired motivation and consequent reduced goal-directed behavior, which has a large impact on quality of life, and is characterized by a reduced willingness to invest effort to obtain rewards (Husain and Roiser, 2018). It is recognized to occur not only in clinical contexts, but also, in milder forms, in the general population, especially with ageing (Brodaty et al., 2010; Grool et al., 2014; Bonnelle et al., 2016). Importantly, apathy is extremely prevalent in a variety of neurological disorders including stroke (i.e., more than 30% of stroke patients suffer from apathy; Caeiro et al., 2013) and can be the result of a reduced sensitivity to rewards, increased sensitivity to efforts or a

combination of both. Hence, apathy could limit the effect of reinforcement on motor learning by reducing the willingness to engage in costly motor control strategies to obtain reinforcement. If this is true, apathy could constitute a predictor of whether patients could benefit from reinforcement during motor rehabilitation.

To tackle this issue, I have recently launched a new project aiming at evaluating the effect of reinforcement on motor learning in stroke patients with varying levels of apathy (**Figure 3A, B, C**). Preliminary results ($n = 18$) suggest that stroke patients have indeed a reduced sensitivity to reward (assessed through the SPSRQ questionnaire) compared to healthy young individuals ($n = 252$), previously tested during my PhD ($p < 0.001$). As shown in **Figure 3D**, it also appears that sensitivity to reward strongly varies in our sample of stroke patients (between 17 [minimum of the scale] and 39 [above the median of the healthy cohort]). Interestingly, consistent with our hypothesis, inter-individual variability in sensitivity to reward was associated to reinforcement-related gains in motor learning with the most reward-sensitive patients being also the ones exhibiting improvements in motor learning. As observed in **Figure 3E**, reinforcement rather tended to impair learning (Error change > 0) in patients with low levels of sensitivity to reward. This differential effect of reinforcement on motor learning depending on sensitivity to reward supports the idea that personalization might be crucial aspect for successful implementation of reinforcement in motor rehabilitation. Moreover, the amount of variance in reinforcement-related gains explained by the model increased from 45% to 76% when adding the level of apathy (evaluated by the AES questionnaire) as a predictor. Notably, none of the tested patients exhibited significant depressive symptoms (as evaluated by the HADS score), suggesting that variability in depression levels is unlikely to account for these preliminary results. This study is still ongoing but these results support the idea that levels of sensitivity to reward and apathy might constitute reliable predictors of whether a given patient could benefit from a reinforcement-based training.

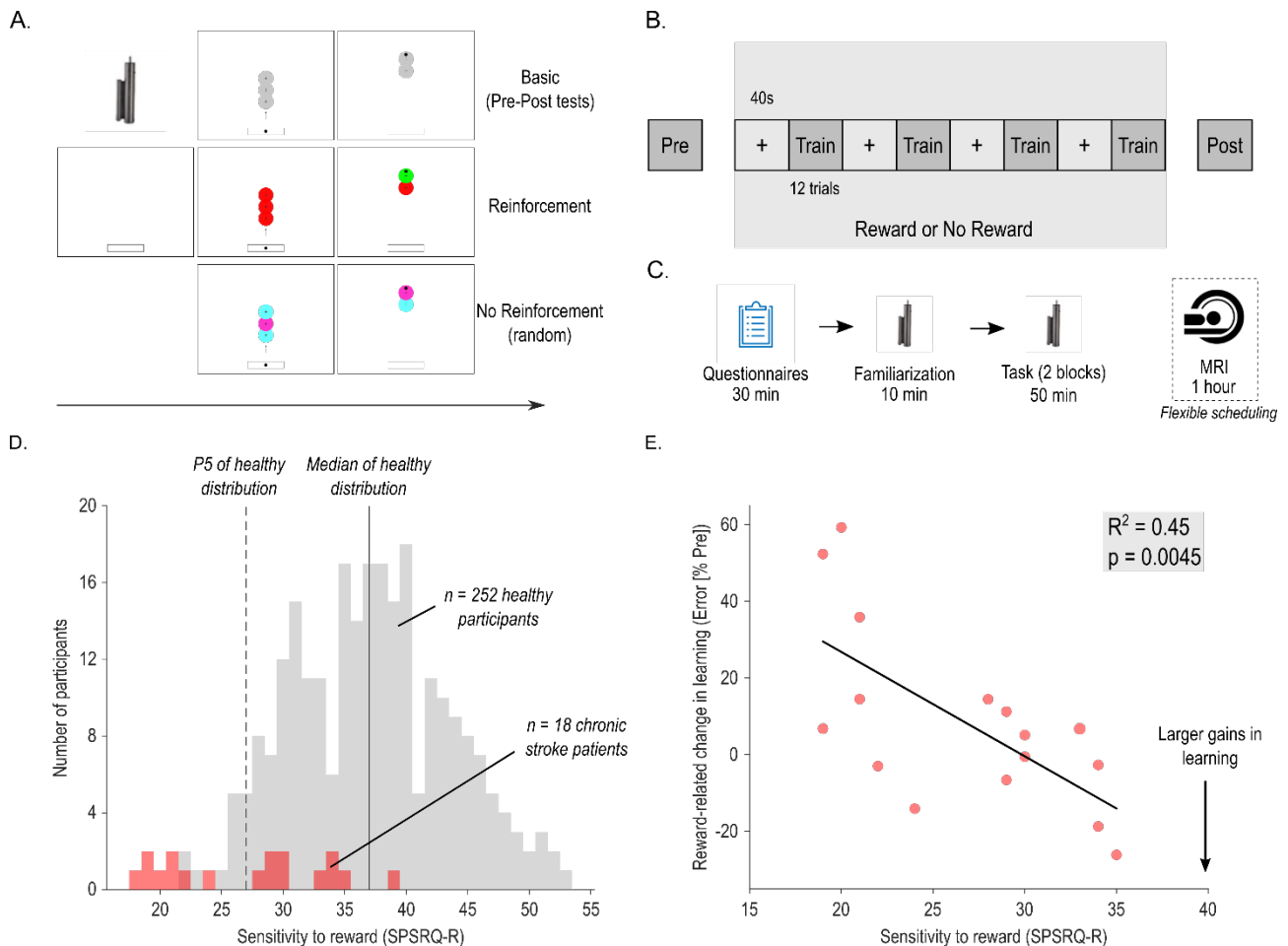


Figure 3. Effect of sensitivity to reward on reinforcement-related gains in motor learning in stroke patients. **A)** Force tracking task. Stroke patients are asked to control a cursor on the screen by squeezing a hand-grip force sensor (represented in the upper left corner) positioned in their affected hand with the goal to follow a moving target. More specifically, they are instructed to position the cursor (black ball) as close as possible to the center of the target (larger circle with a cross) for the whole trial (7 seconds). In basic conditions, the target remains grey for the whole trial. In Reinforcement condition, real-time performance feedback is provided to the participants (Success: Green; Failure: Red). In the No Reward condition, performance feedback is random (Cyan or Magenta independent of performance). Notably, visual properties of the stimuli are matched between conditions (luminance and frequency of color change). Contrary to other studies presented in this PhD, no visual uncertainty was added. **B) Structure of a block.** Patients perform two blocks of the task. Each block is composed as follows: Pre-training (basic conditions; 3 trials), Training (Reward or No Reward conditions; 48 trials; interspersed with 40s resting periods) and Post-training (basic conditions; 6 trials). Notably, a structural and functional (resting-state) MRI is performed, which will allow us to relate lesion localization with reinforcement motor learning. **C) Experimental protocol.** Participants are asked to fill in a battery of questionnaires broadly related to motivation (e.g., AES, SPSRQ, HADS, MFI). Then, they familiarize with the task and perform two blocks of trials (one with each Reinforcement condition). Structural and resting-state functional MRI scans are also performed. **D) Sensitivity to reward in the stroke cohort.** Histogram of sensitivity to reward (assessed with the SPSRQ-R questionnaire) in the tested stroke patients ($n = 18$) and in healthy participants previously tested during my PhD ($n = 252$). The solid and dashed vertical lines represents the median and 5th percentile in the healthy distribution. **E)** Relationship between sensitivity to reward and reinforcement-related benefits in motor learning. Preliminary data on stroke patients suggest a correla-

tion between the individual sensitivity to reward and reinforcement-related changes in learning (Difference of Post-training Error (in % Pre) in Reinforcement vs. No Reinforcement conditions), in line with the hypothesis. Negative numbers indicate an improvement of motor learning with reinforcement.

Interestingly, the results of a separate study performed in healthy participants also supports the idea that sensitivity to reward may influence reinforcement motor learning. As highlighted in the Introduction, motor learning is accompanied by a plastic reorganization of multiple brain regions, including M1 (Pascual-Leone et al., 1995). Such reorganization is also relevant for rehabilitation of motor function following a stroke (Raffin and Hummel, 2018). Interestingly, recent research has suggested that the presence of reinforcement during training may boost M1 plasticity (Uehara et al., 2018; Spampinato et al., 2019), potentially contributing to the beneficial effects of reinforcement on motor learning. Importantly though, another important question relates to the additional effect of motivation by reward on motor learning-induced plasticity. In this study, we evaluated the effect of reinforcement and motivation by reward on M1 plasticity (Lete*, Vassiliadis* et al., 2022 (in prep)) on a subset of the participants evaluated in Study 1. This research was largely inspired by multiple lines of evidence in rodents (Leemburg et al., 2018; Levy et al., 2020; Lee et al., 2021), non-human primates (Ramakrishnan et al., 2017) and humans (Cohen and Ranganath, 2007; Lam et al., 2013) showing reward-related signals in M1, that may boost motor learning-related plasticity (Hosp et al., 2011). We evaluated plastic changes in M1 by measuring the excitability and variability of corticospinal output, as well as by assessing GABAergic short intra-cortical inhibition (SICI) and use-dependent plasticity (UDP) at different time points across training (**Figure 4A**). We found that participants who trained with reward had an early reduction of variability in corticospinal output (**Figure 4B**), suggesting that motivation can indeed modulate learning-induced plasticity in M1. Interestingly, this effect was more pronounced in participants who were more sensitive to rewards (**Figure 4C**), suggesting that the beneficial effect of reward on plasticity is associated to individual sensitivity to incentives, in line with the stroke data reported above. We did not find any effect of reward on the excitability of corticospinal output, neither on SICI or UDP, suggesting that the plastic mechanisms operating in the presence of reward are relatively specific. Overall, these preliminary results suggest that the effect of motivation by reward on reinforcement-based motor learning is related to a specific plasticity mechanism that reduces neural noise in M1, potentially allowing to generate more efficient movements (Manohar et al., 2015, 2019). The

data also suggest that the induction of such plasticity depends on the individual level of sensitivity to reward of participants. To summarize, the stroke and healthy data presented above support the view that individual levels of apathy and sensitivity to reward could be predictors of the behavioral and neural responsiveness to reinforcement-based rehabilitation programs.

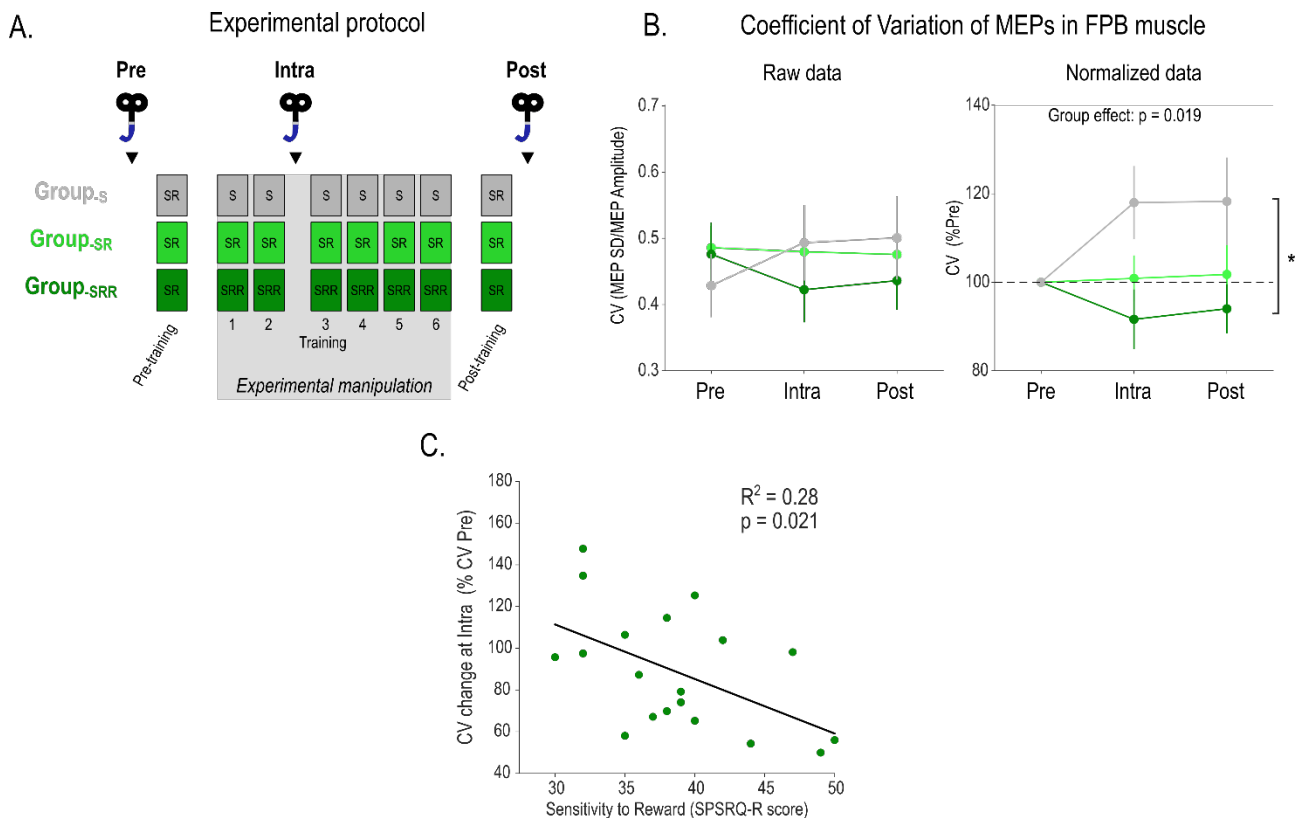


Figure 4. Effect of reward on motor cortex plasticity during motor learning. A) Experimental protocol. The study was conducted on a subset of the participants included in Study 1 ($n = 65$). Motor-evoked potentials (MEPs) induced by TMS applied over M1 were recorded from the Flexor Policis Brevis (FPB) muscle, involved in pinch grip. Peak-to-peak amplitude of MEPs were measured at rest before Pre-training (Pre), after Training 2 (Intra) and after Post-training (Post) in the three experimental groups trained with sensory feedback only (Group-S), sensory and reinforcement feedback (Group-SR) or both types of feedback and monetary reward in the case of good performance (Group-SRR). **B) Variability of resting-state corticospinal output during motor learning.** The coefficient of variation of MEP amplitudes (SD/mean) was computed for each time point and each subject and was used as a proxy of the variability of corticospinal excitability at rest. An ANOVA ran on the Pre-training normalized data showed that the evolution of CVs with training was markedly different between groups ($F_{(2,62)} = 4.22$; $p = 0.019$). Post-hoc tests showed that this effect was driven by a significant difference between CV's in Group-SRR and Group-S (Tukey-corrected $p = 0.016$). No other effect was found. **C) Correlation between training-related plasticity and individual sensitivity to reward.** The reduction of MEP CVs observed early in training in Group-SRR was related to individual sensitivity to reward as evaluated by the Reward questions of the SPSRQ questionnaire.

Towards stratification of patients based on lesion localisation

As mentioned above, Study 3 provides evidence that functional integrity of the striatum and the associated fronto-striatal pathways are required to benefit from reinforcement. This result may therefore have implications for stratification of patients depending on the localisation of their lesion. For instance, stroke patients suffering from striatal lesions that can arise from blockage of the lenticulostriate artery (arising from the middle cerebral artery) or the recurrent artery of Heubner (arising from the anterior cerebral artery), who are known to present strong motor deficits (Liu et al., 2020), may not benefit from reinforcement-based interventions during rehabilitation. Similarly, Parkinson's disease patients may not be able to learn from reinforcement feedback because of striatal dysfunction, as previously suggested (Pekny et al., 2015). In the same vein, the ability to benefit from reinforcement during motor learning is associated to grey-matter volume in the lateral prefrontal cortex (Dayan et al., 2014b). Hence, the presence of fronto-striatal lesions may prevent patients from learning new motor skills through reinforcement. To further explore this idea, an interesting topic for future research will be to relate the ability to benefit from a reinforcement-based motor training with the localisation and type of lesions in specific populations of patients.

Overall, combined with the previous literature discussed above, this knowledge allows us to have a clearer picture of the training parameters that need to be considered for successful integration of reinforcement feedback into future rehabilitation programs (see **Figure 5** for a summary of these ideas). In addition, given the large heterogeneity of clinical profiles undergoing motor rehabilitation, we have started to determine individual factors that determine responsiveness to reinforcement learning of motor skills, a line of research which will be crucial to design patient-tailored rehabilitation strategies.

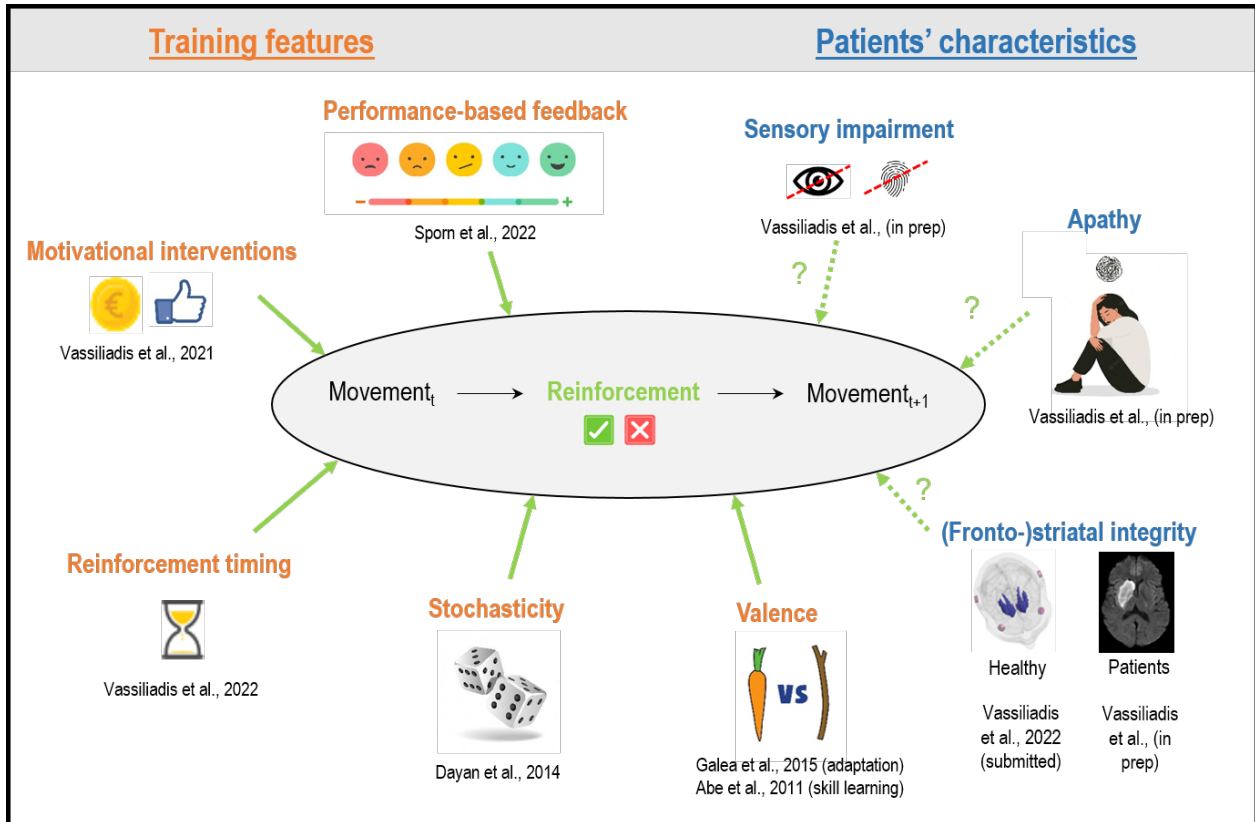


Figure 5. Factors to consider for the use of reinforcement feedback in motor rehabilitation. Summary of the factors that appear to influence reinforcement motor learning, as discussed in this work. These factors are separated in training features and individual characteristics of patients. The dashed arrows are used in the case of factors for which only indirect or preliminary evidence is available.

5.3.2. Transcranial electric temporal interference stimulation: a new non-invasive approach for deep brain stimulation

Study 3 exploited tTIS stimulation in humans to modulate striatal activity. As explained above, tTIS was first validated through physics experiments, computational modeling (Rampersad et al., 2019; Cao et al., 2020; Mirzakhali et al., 2020; Esmailpour et al., 2021; von Conta et al., 2021) and rodents experiments (Grossman et al., 2017; Song et al., 2021). First applications of tTIS in humans have been recently performed in cadavers (Acerbo et al., 2022) and in-vivo on superficial targets (Ma et al., 2022; Zhu et al., 2022). In a previous study from the lab in which I have participated during my PhD, we have applied striatal tTIS stimulation patterned as an intermittent theta burst stimulation (iTBS, a protocol known to support plasticity) while healthy participants were practicing a motor sequence

learning task with the left hand in the MRI (Wessel*, Beanato* et al., 2022, submitted). We found that tTIS_{ITBS} increased activity in the striatum during practice, in particular in a subregion of the striatum that was more activated by the task (i.e., the right putamen). Moreover, these changes of brain activity were accompanied by an improvement of motor performance during training. Combined with the data reported in Study 3, this constitutes converging evidence that tTIS stimulation is able to overcome the depth-focality tradeoff of current non-invasive brain stimulation techniques in humans with sufficient current strength, by showing specific, focal and functionally-relevant modulation of the striatum. In addition, this body of work indicates that tTIS stimulation can modulate different striatal functions depending on the pattern of stimulation, and the state of the underlying network. In both studies, no adverse events were reported and the reported sensations were generally mild and compatible with an efficient blinding. Overall, these elements suggest that tTIS is safe, pain-less and able to stimulate deep brain structures with relevant behavioral effects.

Importantly, despite these promising first results, there is a strong need to develop a careful mechanistic understanding of tTIS effects. First, a core assumption of tTIS is that neurons are not sensitive to electrical currents delivered in the high frequency range, allowing neuromodulation of deep structures without significant stimulation of the overlying cortex. Consistently, the first study applying tTIS on the hippocampus of rodents did not find any evidence for concurrent cortical stimulation (Grossman et al., 2017). However, a previous study did report motor excitability changes in response to high frequency tACS applied on M1 (Chaieb et al., 2011), suggesting that high-frequency currents can indeed modulate cortical activity in humans. Importantly, in both tTIS studies conducted in the context of my PhD, we compared our experimental stimulation with an active control that also contained the high frequency component (i.e., tTIS_{20Hz} for Study 3 and a pure high frequency stimulation in Wessel*, Beanato* et al., 2022). Hence, the comparison between the experimental conditions and the active control allowed us to isolate the unique contribution of the tTIS field. That being said, better understanding the potential influence of high-frequency electrical stimulation on neuronal activity appears as an important line of future research (Mirzakhilili et al., 2020).

Another important question is how electric field magnitudes elicited by tTIS deep in the brain are able to modulate neuronal activity. As such, the simulations performed by our collaborators (team of Prof. Esra Neufeld, ETZ) indicates that electric fields in the striatum

are in the range of 0.5-0.6 V/m with the stimulation parameters used in the present work. Our results therefore support the idea that neuromodulation is possible at these weak electric field magnitudes, in line with previous tACS studies in primates showing neuronal effects with electric fields as low as 0.2-0.3 V/m (Krause et al., 2019, 2022). Importantly, another aspect that may influence the responsiveness to weak electric fields is the concurrent engagement of the stimulated structures in a task. As such, we have found a preferential increase of activity in subparts of the striatum that were already engaged in the task, supporting the view that co-activation is important for tTIS-related neuromodulation (Wessel*, Beanato* et al., 2022). Such co-activation may bring the neurons closer to a critical threshold for depolarization, ultimately increasing their sensitivity to weak electrical fields. Overall, our work indicates that the putatively weak electric fields generated by tTIS are sufficient for deep brain neuromodulation and further suggests that concurrent co-activation may be a key element in this process.

In summary, tTIS is a promising new technology allowing to neuromodulate deep brain regions non-invasively. Beyond its clear relevance in research to establish causal links between deep brain regions and particular behaviors (as done in Study 3), this technique may also be relevant to support recovery of neuro-psychiatric disorders characterized by abnormal processing in deep brain regions in a non-invasive and cost-effective way.

5.4. Conclusion

The present work characterized key mechanisms underlying the effect of reinforcement on motor learning using a combination of behavioral analyses, brain stimulation and neuroimaging. More specifically, we have shown that motivation, timing of reinforcement and striatal neuromodulation can all strongly influence reinforcement learning of motor skills. As I have argued in this discussion, personalized reinforcement feedback appears to be a promising tool to improve current motor rehabilitation protocols.

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. *Curr Biol* 21:557–562 Available at: <http://dx.doi.org/10.1016/j.cub.2011.02.030>.
- Acerbo E, Jegou A, Luff C, Dzialecka P, Botzanowski B, Missey F, Ngom I, Lagarde S, Bartolomei F, Cassara A, Neufeld E, Jirsa V, Carron R, Grossman N, Williamson A (2022) Focal non-invasive deep-brain stimulation with temporal interference for the suppression of epileptic biomarkers. *Front Neurosci* 16:1–12.
- Albert NB, Robertson EM, Miall RC (2009) The Resting Human Brain and Motor Learning. *Curr Biol* 19:1023–1027 Available at: <http://dx.doi.org/10.1016/j.cub.2009.04.028>.
- Albouy G, Fogel S, King BR, Laventure S, Benali H, Karni A, Carrier J, Robertson EM, Doyon J (2015) Maintaining vs. enhancing motor sequence memories: Respective roles of striatal and hippocampal systems. *Neuroimage* 108:423–434 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2014.12.049>.
- Albouy G, King BR, Maquet P, Doyon J (2013a) Hippocampus and striatum: Dynamics and interaction during acquisition and sleep-related motor sequence memory consolidation. *Hippocampus* 23:985–1004.
- Albouy G, Sterpenich V, Baiteau E, Vandewalle G, Desseilles M, Dang-Vu T, Darsaud A, Ruby P, Luppi PH, Degueldre C, Peigneux P, Luxen A, Maquet P (2008) Both the Hippocampus and Striatum Are Involved in Consolidation of Motor Sequence Memory. *Neuron* 58:261–272.
- Albouy G, Sterpenich V, Vandewalle G, Darsaud A, Gais S, Rauchs G, Desseilles M, Boly M, Dang-Vu T, Baiteau E, Degueldre C, Phillips C, Luxen A, Maquet P (2012) Neural correlates of performance variability during motor sequence acquisition. *Neuroimage* 60:324–331 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2011.12.049>.
- Albouy G, Sterpenich V, Vandewalle G, Darsaud A, Gais S, Rauchs G, Desseilles M, Boly M, Dang-Vu T, Baiteau E, Degueldre C, Phillips C, Luxen A, Maquet P (2013b) Interaction between Hippocampal and Striatal Systems Predicts Subsequent Consolidation of Motor Sequence Memory. *PLoS One* 8:12–14.
- Algoet M, Duque J, Iannetti GD, Mouraux A (2018) Temporal Profile and Limb-specificity of Phasic Pain-Evoked Changes in Motor Excitability. *Neuroscience* 386:240–255 Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0306452218304573>.
- Allen GI, Tsukahara N (1974) Cerebrocerebellar communication systems. *Physiol Rev* 54:957–1006 Available at: <https://www.physiology.org/doi/10.1152/physrev.1974.54.4.957>.
- Ambrose RE, Pfeiffer BE, Foster DJ (2016) Reverse Replay of Hippocampal Place Cells Is Uniquely Modulated by Changing Reward. *Neuron* 91:1124–1136 Available at: <http://dx.doi.org/10.1016/j.neuron.2016.07.047>.
- Anderson BJ, Li X, Alcantara AA, Isaacs KR, Black JE, Greenough WT (1994) Glial hypertrophy is associated with synaptogenesis following motor-skill learning, but not with angiogenesis following exercise. *Glia* 11:73–80.
- Antal A et al. (2017) Low intensity transcranial electric stimulation: Safety, ethical, legal regulatory and application guidelines. *Clin Neurophysiol* 128:1774–1809 Available at: <http://dx.doi.org/10.1016/j.clinph.2017.06.001>.
- Areshenkoff CN, de Brouwer AJ, Gale DJ, Nashed JY, Gallivan JP (2022) SEPARATE AND SHARED LOW - DIMENSIONAL NEURAL ARCHITECTURES FOR ERROR - BASED AND REINFORCEMENT.
- Atkinson-Clement C, Cavazzini É, Zénon A, Witjas T, Fluchère F, Azulay JP, Baunez C, Eusebio A (2019) Effects of subthalamic nucleus stimulation and levodopa on decision-making in Parkinson's disease. *Mov Disord* 34:377–385.
- Averbeck B, O'Doherty JP (2022) Reinforcement-learning in fronto-striatal circuits. *Neuropsychopharmacology* 47:147–162.
- Avraham G, Taylor JA, Ivry RB, McDougle SD (2020) An associative learning account of sensorimotor adaptation. *bioRxiv*.
- Balleine BW, Delgado MR, Hikosaka O (2007) The role of the dorsal striatum in reward and decision-making. *J Neurosci* 27:8161–8165.
- Balleine BW, O'Doherty JP (2010) Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35:48–69.

- Barron AB, Søvik E, Cornish JL (2010) The roles of dopamine and related compounds in reward-seeking behavior across animal phyla. *Front Behav Neurosci* 4:1–9.
- Bartra O, McGuire JT, Kable JW (2013) The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* 76:412–427 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2013.02.063>.
- Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67.
- Battaglia F, Quartarone A, Ghilardi MF, Dattola R, Bagnato S, Rizzo V, Morgante L, Girlanda P (2006) Unilateral cerebellar stroke disrupts movement preparation and motor imagery. *Clin Neurophysiol* 117:1009–1016.
- Beliaeva V, Savvateev I, Zerbi V, Polania R (2021) Toward integrative approaches to study the causal role of neural oscillations via transcranial electrical stimulation. *Nat Commun* 12:1–12 Available at: <http://dx.doi.org/10.1038/s41467-021-22468-7>.
- Ben-Shachar M, Lüdtke D, Makowski D (2020) effectsize: Estimation of Effect Size Indices and Standardized Parameters. *J Open Source Softw* 5:2815.
- Berke JD (2009) Fast oscillations in cortical-striatal networks switch frequency following rewarding events and stimulant drugs. *Eur J Neurosci* 30:848–859.
- Berke JD (2018) What does dopamine mean? *Nat Neurosci* 21:787–793 Available at: <http://dx.doi.org/10.1038/s41593-018-0152-y>.
- Bernardi NF, Darainy M, Ostry DJ (2015) Somatosensory contribution to the initial stages of human motor learning. *J Neurosci* 35:14316–14326.
- Bian T, Wolpert DM, Jiang ZP (2020) Model-free robust optimal feedback mechanisms of biological motor control. *Neural Comput* 32:562–595.
- Bliss, T.V.P. & Collingridge GL (1993) A synaptic model of memory: LTP in the hippocampus. *Nature* 361:31–39.
- Boele HJ, Koekkoek SKE, De Zeeuw CI, Ruigrok TJH (2013) Axonal sprouting and formation of terminals in the adult cerebellum during associative motor learning. *J Neurosci* 33:17897–17907.
- Bologna M, Rocchi L, Paparella G, Nardella A, Li P, Conte A, Kojovic M, Rothwell JC, Berardelli A (2015) Brain Stimulation Reversal of Practice-related Effects on Corticospinal Excitability has no Immediate Effect on Behavioral Outcome. *Brain Stimul* 8:603–612 Available at: <http://dx.doi.org/10.1016/j.brs.2015.01.405>.
- Bongioanni A, Folloni D, Verhagen L, Sallet J, Klein-Flügge MC, Rushworth MFS (2021) Activation and disruption of a neural mechanism for novel choice in monkeys. *Nature*.
- Bonnelle V, Manohar S, Behrens T, Husain M (2016) Individual Differences in Premotor Brain Systems Underlie Behavioral Apathy. *Cereb Cortex* 26:807–819.
- Bönstrup M, Iturrate I, Hebart MN, Censor N, Cohen LG (2020) Mechanisms of offline motor learning at a microscale of seconds in large-scale crowdsourced data. *npj Sci Learn* 5:1–10 Available at: <http://dx.doi.org/10.1038/s41539-020-0066-9>.
- Bönstrup M, Iturrate I, Thompson R, Cruciani G, Censor N, Cohen LG (2019) A Rapid Form of Offline Consolidation in Skill Learning. *Curr Biol* 29:1346–1351.e4 Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0960982219302192>.
- Boraud T, Leblois A, Rougier NP (2018) A natural history of skills. *Prog Neurobiol* 171:114–124 Available at: <https://doi.org/10.1016/j.pneurobio.2018.08.003>.
- Bostan AC, Strick PL (2018) The basal ganglia and the cerebellum: Nodes in an integrated network. *Nat Rev Neurosci* 19:338–350 Available at: <http://dx.doi.org/10.1038/s41583-018-0002-7>.
- Bowles S, Hickman J, Peng X, Williamson WR, Huang R, Washington K, Donegan D, Welle CG (2022) Vagus nerve stimulation drives selective circuit modulation through cholinergic reinforcement. *Neuron* 110:2867–2885.e7 Available at: <https://doi.org/10.1016/j.neuron.2022.06.017>.
- Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436 Available at: <http://booksandjournals.brillonline.com/content/journals/10.1163/156856897x00357>.
- Branscheidt M, Kassavetis P, Anaya M, Rogers D, Huang HD, Lindquist MA, Celnik P (2019) Fatigue induces long-lasting detrimental

changes in motor-skill learning. *Elife* 8 Available at: <https://elifesciences.org/articles/40578>.

- Brodaty H, Altendorf A, Withall A, Sachdev P (2010) Do people become more apathetic as they grow older? A longitudinal study in healthy individuals. *Int Psychogeriatrics* 22:426–436.
- Brown P (2007) Abnormal oscillatory synchronisation in the motor system leads to impaired movement. *Curr Opin Neurobiol* 17:656–664.
- Brücke C, Huebl J, Schönecker T, Neumann WJ, Yarrow K, Kupsch A, Blahak C, Lütjens G, Brown P, Krauss JK, Schneider GH, Kühn AA (2012) Scaling of movement is related to pallidal γ oscillations in patients with dystonia. *J Neurosci* 32:1008–1019.
- Buch ER, Claudino L, Quentin R, Bönstrup M, Cohen LG (2021) Consolidation of human skill linked to waking hippocampo-neocortical replay. *Cell Rep* 35.
- Bush G, Vogt BA, Holmes J, Dale AM, Greve D, Jenike MA, Rosen BR (2002) Dorsal anterior cingulate cortex: A role in reward-based decision making. *Proc Natl Acad Sci U S A* 99:523–528.
- Butefisch CM, Davis BC, Wise SP, Sawaki L, Kopylev L, Classen J, Cohen LG (2000) Mechanisms of use-dependent plasticity in the human motor cortex. *Proc Natl Acad Sci* 97:3661–3665 Available at: <http://www.pnas.org/cgi/doi/10.1073/pnas.97.7.3661>.
- Caeiro L, Ferro JM, Costa J (2013) Apathy secondary to stroke: A systematic review and meta-analysis. *Cerebrovasc Dis* 35:23–39.
- Cantarero G, Lloyd A, Celnik P (2013) Reversal of Long-Term Potentiation-Like Plasticity Processes after Motor Learning Disrupts Skill Retention. *J Neurosci* 33:12862–12869.
- Cao J, Doiron B, Goswami C, Grover P (2020) The mechanics of temporal interference stimulation. *bioRxiv*:1–6.
- Carr MF, Jadhav SP, Frank LM (2011) Hippocampal replay in the awake state: A potential substrate for memory consolidation and retrieval. *Nat Neurosci* 14:147–153.
- Carroll TJ, McNamee D, Ingram JN, Wolpert DM (2019) Rapid visuomotor responses reflect value-based decisions. *J Neurosci* 39:3906–3920.
- Carta I, Chen CH, Schott AL, Dorizan S, Khodakhah K (2019) Cerebellar modulation of the reward circuitry and social behavior. *Science* (80-) 363.
- Cashaback JGA, McGregor HR, Mohatarem A, Gribble PL (2017) Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. *PLoS Comput Biol* 13:1–28.
- Catanese J, Carmichael JE, van der Meer MAA (2016) Low- and high-gamma oscillations deviate in opposite directions from zero-phase synchrony in the limbic corticostriatal loop. *J Neurophysiol* 116:5–17.
- Chaieb L, Antal A, Paulus W (2011) Transcranial alternating current stimulation in the low kHz range increases motor cortex excitability. *Restor Neurol Neurosci* 29:167–175.
- Chen X, Holland P, Galea JM (2017) The effects of reward and punishment on motor skill learning. *Curr Opin Behav Sci* 20:83–88 Available at: <http://dx.doi.org/10.1016/j.cobeha.2017.11.011>.
- Christiansen L, Madsen MJ, Bojsen-møller E, Thomas R, Nielsen JB (2018) Brain Stimulation Progressive practice promotes motor learning and repeated transient increases in corticospinal excitability across multiple days. *Brain Stimul* 11:346–357 Available at: <https://doi.org/10.1016/j.brs.2017.11.005>.
- Churchland MM, Cunningham JP, Kaufman MT, Foster JD, Nuyujukian P, Ryu SI, Shenoy K V., Shenoy K V. (2012) Neural population dynamics during reaching. *Nature* 487:51–56.
- Cincotta M, Borgheresi A, Jung P, Balestrieri F, Giovannelli F, Zaccara G, Ziemann U (2005) Physical interactions between induced electrical fields can have substantial effects on neuronal excitation during simultaneous TMS of two brain areas. *Clin Neurophysiol* 116:1733–1742.
- Cisek P (2007) Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos Trans R Soc Lond B Biol Sci* 362:1585–1599.
- Cisek P (2019) Resynthesizing behavior through phylogenetic refinement.

- Clark K, Gregory Appelbaum L, van den Berg B, Mitroff SR, Woldorff MG (2015) Improvement in visual search with practice: Mapping learning-related changes in neurocognitive stages of processing. *J Neurosci* 35:5351–5359.
- Classen J, Liepert J, Wise SP, Hallett M, Cohen LG (1998) Rapid Plasticity of Human Cortical Movement Representation Induced by Practice. *J Neurophysiol* 79:1117–1123 Available at: <https://www.physiology.org/doi/10.1152/jn.1998.79.2.1117>.
- Clemente F, D'Alonzo M, Controzzi M, Edin BB, Cipriani C (2016) Non-Invasive, Temporally Discrete Feedback of Object Contact and Release Improves Grasp Control of Closed-Loop Myoelectric Transradial Prostheses. *IEEE Trans Neural Syst Rehabil Eng* 24:1314–1322.
- Codol O, Holland PJ, Galea JM (2018) The relationship between reinforcement and explicit control during visuomotor adaptation. *Sci Rep* 8:1–11 Available at: <http://dx.doi.org/10.1038/s41598-018-27378-1>.
- Codol O, Holland PJ, Manohar SG, Galea JM (2020) Reward-Based Improvements in Motor Control Are Driven by Multiple Error-Reducing Mechanisms. *J Neurosci* 40:3604–3620 Available at: <http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.2646-19.2020>.
- Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N (2012) Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* 482:85–88.
- Cohen MX, Axmacher N, Lenartz D, Elger CE, Sturm V, Schlaepfer TE (2009) Good vibrations: Cross-frequency coupling in the human nucleus accumbens during reward processing. *J Cogn Neurosci* 21:875–889.
- Cohen MX, Ranganath C (2007) Reinforcement Learning Signals Predict Future Decisions. *J Neurosci* 27:371–378 Available at: <http://www.jneurosci.org/cgi/content/abstract/27/2/371>.
- Comite A De, Crevecoeur F, Lefèvre P (2021) Reward-dependent selection of feedback gains impacts rapid motor decisions. *bioRxiv Neurosci* Available at: http://biorxiv.org/cgi/content/short/2021.07.25.453678v1?rss=1&utm_source=researcher_app&utm_medium=referral&utm_campaign=RESR_MRKT_Researcher_inbound.
- Connell LA (2008) Somatosensory impairment after stroke: Frequency of different deficits and their recovery.
- Costa RM, Lin SC, Sotnikova TDD, Cyr M, Gainetdinov RRR, Caron MGG, Nicolelis MAAL (2006) Rapid Alterations in Corticostriatal Ensemble Coordination during Acute Dopamine-Dependent Motor Dysfunction. *Neuron* 52:359–369.
- Courtemanche R, Fujii N, Graybiel AM (2003) Synchronous, Focally Modulated β -Band Oscillations Characterize Local Field Potential Activity in the Striatum of Awake Behaving Monkeys. *J Neurosci* 23:11741–11752.
- Crevecoeur F, Scott SH, Cluff T (2019) Robust Control in Human Reaching Movements: A Model-Free Strategy to Compensate for Unpredictable Disturbances. *J Neurosci* 39:8135–8148.
- Dang MT, Yokoi F, Yin HH, Lovinger DM, Wang Y, Li Y (2006) Disrupted motor learning and long-term synaptic plasticity in mice lacking NMDAR1 in the striatum. *Proc Natl Acad Sci U S A* 103:15254–15259.
- Dayan E, Averbeck BB, Richmond BJ, Cohen LG (2014a) Stochastic reinforcement benefits skill acquisition. *Learn Mem* 21:140–142.
- Dayan E, Cohen LG (2011) Neuroplasticity Subservicing Motor Skill Learning. *Neuron* 72:443–454 Available at: <http://dx.doi.org/10.1016/j.neuron.2011.10.008>.
- Dayan E, Hamann XJM, Averbeck BB, Cohen LG (2014b) Brain Structural Substrates of Reward Dependence during Behavioral Performance. *J Neurosci* 34:16433–16441.
- Dayan E, Herszage J, Laor-Maayan R, Sharon H, Censor N (2018) Neuromodulation of reinforced skill learning reveals the causal function of prefrontal cortex. *Hum Brain Mapp* 39:4724–4732.
- Dayan P, Niv Y (2008) Reinforcement learning: The Good, The Bad and The Ugly. *Curr Opin Neurobiol* 18:185–196.
- Debas K, Carrier J, Orban P, Barakat M, Lungu O, Vandewalle G, Tahar AH, Bellec P, Karni A, Ungerleider LG, Benali H, Doyon J (2010) Brain plasticity related to the consolidation of motor sequence learning and motor adaptation. *Proc Natl Acad Sci U S A* 107:17839–17844.
- Del Arco A, Park J, Wood J, Kim Y, Moghaddam B (2017) Adaptive encoding of outcome prediction by prefrontal cortex ensembles supports behavioral flexibility. *J Neurosci* 37:8363–8373.

- Deng Z De, Lisanby SH, Peterchev A V. (2013) Electric field depth-focality tradeoff in transcranial magnetic stimulation: Simulation comparison of 50 coil designs. *Brain Stimul* 6:1–13 Available at: <http://dx.doi.org/10.1016/j.brs.2012.02.005>.
- Derosiere G (2018) A Dynamical System Framework for Theorizing Preparatory Inhibition. *J Neurosci* 38:3391–3393 Available at: <http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.0069-18.2018>.
- Derosière G, Alexandre F, Bourdillon N, Mandrick K, Ward TE, Perrey S (2014) Similar scaling of contralateral and ipsilateral cortical responses during graded unimanual force generation. *Neuroimage* 85:471–477 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2013.02.006>.
- Derosière G, Billot M, Ward ET, Perrey S (2015) Adaptations of motor neural structures' activity to lapses in attention. *Cereb Cortex* 25:66–74.
- Derosiere G, Duque J (2020) Tuning the Corticospinal System: How Distributed Brain Circuits Shape Human Actions. *Neurosci*:107385841989675 Available at: <http://journals.sagepub.com/doi/10.1177/1073858419896751>.
- Derosiere G, Perrey S (2012) Relationship between submaximal handgrip muscle force and NIRS-measured motor cortical activation. *Adv Exp Med Biol* 737:269–274.
- Derosiere G, Thura D, Cisek P, Duque J (2019) Motor cortex disruption delays motor processes but not deliberation about action choices. *J Neurophysiol* 122:1566–1577 Available at: <https://www.physiology.org/doi/10.1152/jn.00163.2019>.
- Derosiere G, Vassiliadis P, Demaret S, Zénon A, Duque J (2017a) Learning stage-dependent effect of M1 disruption on value-based motor decisions. *Neuroimage* Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1053811917307164>.
- Derosiere G, Vassiliadis P, Demaret S, Zénon A, Duque J (2017b) Learning stage-dependent effect of M1 disruption on value-based motor decisions. *Neuroimage* 162:173–185.
- Derosiere G, Vassiliadis P, Duque J (2020) Advanced TMS approaches to probe corticospinal excitability during action preparation. *Neuroimage* 213:116746 Available at: <https://doi.org/10.1016/j.neuroimage.2020.116746>.
- Derosiere G, Zénon A, Alamia A, Duque J (2017c) Primary motor cortex contributes to the implementation of implicit value-based rules during motor decisions. *Neuroimage* 146:1115–1127 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2016.10.010>.
- Dhawale AK, Miyamoto YR, Smith MA, Ölveczky BP (2019) Adaptive Regulation of Motor Variability. *Curr Biol* 29:3551–3562.e7.
- Dhawale AK, Smith MA (2017) The Role of Variability in Motor Learning.
- Dhawale AK, Smith MA, Ölveczky BP (2017) The Role of Variability in Motor Learning. *Annu Rev Neurosci* 40:479–498 Available at: <http://www.annualreviews.org/doi/10.1146/annurev-neuro-072116-031548>.
- Diedrichsen J, White O, Newman D, Lally N (2010) Use-Dependent and Error-Based Learning of Motor Behaviors. *J Neurosci* 30:5159–5166 Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.5406-09.2010>.
- Donnelly NA, Holtzman T, Rich PD, Nevado-Holgado AJ, Fernando ABP, Van Dijck G, Holzhammer T, Paul O, Ruther P, Paulsen O, Robbins TW, Dalley JW (2014) Oscillatory activity in the medial prefrontal cortex and nucleus accumbens correlates with impulsivity and reward outcome. *PLoS One* 9:14–17.
- Doumas I, Everard G, Dehem S, Lejeune T (2021) Serious games for upper limb rehabilitation after stroke: a meta-analysis. *J Neuroeng Rehabil* 18:1–16.
- Doyon J, Benali H (2005) Reorganization and plasticity in the adult brain during learning of motor skills. *Curr Opin Neurobiol* 15:161–167.
- Doyon J, Gabbitov E, Vahdat S, Lungu O, Boutin A (2018) Current issues related to motor sequence learning in humans. *Curr Opin Behav Sci* 20:89–97 Available at: <https://doi.org/10.1016/j.cobeha.2017.11.012>.
- Doyon J, Korman M, Morin A, Dostie V, Tahar AH, Benali H, Karni A, Ungerleider LG, Carrier J (2009) Contribution of night and day sleep vs. simple passage of time to the consolidation of motor sequence and visuomotor adaptation learning. *Exp Brain Res* 195:15–26.
- Draaisma LR, Wessel MJ, Moyne M, Morishita T, Hummel FC (2022) Targeting the frontoparietal network using bifocal transcranial alternating current stimulation during a motor sequence learning task in healthy older adults. *Brain Stimul* 15:968–979

Available at: <https://doi.org/10.1016/j.brs.2022.06.012>.

- Du Y, Krakauer JW, Haith AM (2022) The relationship between habits and motor skills in humans. *Trends Cogn Sci*:1–17 Available at: <https://doi.org/10.1016/j.tics.2022.02.002>.
- Dudek SM, Bear MF (1992) Homosynaptic long-term depression in area CA1 of hippocampus and effects of N-methyl-D-aspartate receptor blockade. *Proc Natl Acad Sci U S A* 89:4363–4367.
- Dupont-Hadwen J, Bestmann S, Stagg CJ (2018) Motor training modulates intracortical inhibitory dynamics in motor cortex during movement preparation. *Brain Stimul* Available at: <https://linkinghub.elsevier.com/retrieve/pii/S1935861X18303681>.
- Duque J, Greenhouse I, Labruna L, Ivry RB (2017) Physiological Markers of Motor Inhibition during Human Behavior. *Trends Neurosci* 40:219–236 Available at: <http://dx.doi.org/10.1016/j.tins.2017.02.006>.
- Duque J, Ivry RB (2009) Role of corticospinal suppression during motor preparation. *Cereb Cortex* 19:2013–2024.
- Duque J, Labruna L, Cazaes C, Ivry RB (2014) Dissociating the influence of response selection and task anticipation on corticospinal suppression during response preparation. *Neuropsychologia* 65:287–296 Available at: <http://dx.doi.org/10.1016/j.neuropsychologia.2014.08.006>.
- Duque J, Lew D, Mazzocchio R, Olivier E, Richard B (2010) Evidence for two concurrent inhibitory mechanisms during response preparation. *J Neurosci* 30:3793–3802.
- Duque J, Mazzocchio R, Stefan K, Hummel F, Olivier E, Cohen LG (2008) Memory formation in the motor cortex ipsilateral to a training hand. *Cereb Cortex* 18:1395–1406.
- Duque J, Petitjean C, Swinnen SP (2016) Effect of aging on motor inhibition during action preparation under sensory conflict. *Front Aging Neurosci* 8:1–14.
- Dwiel LL, Khokhar JY, Connerney MA, Green AI, Doucette WT (2019) Finding the balance between model complexity and performance: Using ventral striatal oscillations to classify feeding behavior in rats. *PLoS Comput Biol* 15:1–19.
- Efron B (1979) Bootstrap Methods: Another Look at the Jackknife. *Ann Stat* 7:1–26.
- Ekhtiari H et al. (2022) A checklist for assessing the methodological quality of concurrent tES-fMRI studies (ContES checklist): a consensus study and statement. *Nat Protoc* 17:596–617.
- Elbert T, Pantev C, Wienbruch C, Rockstroh B, Taub E (1995) Increased Cortical Representation of the Fingers of the Left Hand in String Players. *Science* (80-) 270:305–307 Available at: <https://www.science.org/doi/10.1126/science.270.5234.305>.
- Engelhard B, Finkelstein J, Cox J, Fleming W, Jang HJ, Ornelas S, Koay SA, Thiberge SY, Daw ND, Tank DW, Witten IB (2019) Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature* 570:509–513 Available at: <http://dx.doi.org/10.1038/s41586-019-1261-9>.
- Esmailpour Z, Kronberg G, Reato D, Parra LC, Bikson M (2021) Temporal interference stimulation targets deep brain regions by modulating neural oscillations. *Brain Stimul* 14:55–65 Available at: <https://doi.org/10.1016/j.brs.2020.11.007>.
- Ferbert A, Priori A, Rothwell JC, Day BL, Colebatch JG, Marsden CD (1992) Interhemispheric inhibition of the human motor cortex. *J Physiol* 453:525–546.
- Fernández-Seara MA, Aznárez-Sanado M, Mengual E, Loayza FR, Pastor MA (2009) Continuous performance of a novel motor sequence leads to highly correlated striatal and hippocampal perfusion increases. *Neuroimage* 47:1797–1808 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2009.05.061>.
- Fiorillo CD, Newsome WT, Schultz W (2008) The temporal precision of reward prediction in dopamine neurons. *Nat Neurosci* 11:966–973.
- Fischer S, Nitschke MF, Melchert UH, Erdmann C, Born J (2005) Motor memory consolidation in sleep shapes more effective neuronal representations. *J Neurosci* 25:11248–11255.
- Flanagan JR, Wing AM (1997) The role of internal models in motion planning and control: Evidence from grip force adjustments during movements of hand-held loads. *J Neurosci* 17:1519–1528.
- Floyer-Lea A, Matthews PM (2004) Changing brain networks for visuomotor control with increased movement automaticity. *J*

Neurophysiol 92:2405–2412.

Floyer-Lea A, Matthews PM (2005) Distinguishable brain activation networks for short- and long-term motor skill learning. *J Neurophysiol* 94:512–518.

Foerde K, Braun EK, Shohamy D (2012) A Trade-Off between feedback-based learning and episodic memory for feedback events: Evidence from Parkinson's disease. *Neurodegener Dis* 11:93–101.

Foerde K, Race E, Verfaellie M, Shohamy D (2013) A role for the medial temporal lobe in feedback-driven learning: Evidence from amnesia. *J Neurosci* 33:5698–5704.

Foerde K, Shohamy D (2011) Feedback timing modulates brain systems for learning in humans. *J Neurosci* 31:13157–13167.

Foster DJ, Wilson MA (2006) Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature* 440:680–683.

Freedberg M, Toader AC, Wassermann EM, Voss JL (2020) Competitive and cooperative interactions between medial temporal and striatal learning systems. *Neuropsychologia* 136:107257 Available at: <https://doi.org/10.1016/j.neuropsychologia.2019.107257>.

Gabay Y, Shahbari-Khateb E, Mendelsohn A (2018) Feedback Timing Modulates Probabilistic Learning in Adults with ADHD. *Sci Rep* 8:1–11 Available at: <http://dx.doi.org/10.1038/s41598-018-33551-3>.

Gabitov E, Lungu O, Albouy G, Doyon J (2019) Weaker Inter-hemispheric and Local Functional Connectivity of the Somatomotor Cortex During a Motor Skill Acquisition Is Associated With Better Learning. *Front Neurol* 10:1–19.

Gabitov E, Manor D, Karni A (2015) Patterns of Modulation in the Activity and Connectivity of Motor Cortex during the Repeated Generation of Movement Sequences. *J Cogn Neurosci* 27:736–751 Available at: <https://direct.mit.edu/jocn/article/27/4/736/28336/Patterns-of-Modulation-in-the-Activity-and->

Galea JM, Celnik P (2009) Brain Polarization Enhances the Formation and Retention of Motor Memories. *J Neurophysiol* 102:294–301 Available at: <http://www.physiology.org/doi/10.1152/jn.00184.2009>.

Galea JM, Mallia E, Rothwell J, Diedrichsen J (2015a) The dissociable effects of punishment and reward on motor learning. *Nat Neurosci* 18:597–602 Available at: <http://dx.doi.org/10.1038/nn.3956><http://www.nature.com/neuro/journal/v18/n4/abs/nn.3956.html#supplementary-information>.

Galea JM, Mallia E, Rothwell J, Diedrichsen J (2015b) The dissociable effects of punishment and reward on motor learning. *Nat Neurosci* 18:597–602.

Galea JM, Ruge D, Buijink A, Bestmann S, Rothwell JC (2013) Punishment-induced behavioral and neurophysiological variability reveals dopamine-dependent selection of kinematic movement parameters. *J Neurosci* 33:3981–3988.

Garrison J, Erdeniz B, Done J (2013) Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neurosci Biobehav Rev* 37:1297–1310 Available at: <http://dx.doi.org/10.1016/j.neubiorev.2013.03.023>.

Gershman SJ, Daw ND (2017) Reinforcement Learning and Episodic Memory in Humans and Animals : An Integrative Framework.

Ghahramani Z, Wolpert DM (1997) Modular decomposition in visuomotor learning. *Nature* 386:392–395 Available at: <http://www.nature.com/articles/386392a0>.

Giordano N, Iemolo A, Mancini M, Cacace F, De Risi M, Latagliata EC, Ghiglieri V, Bellenchi GC, Puglisi-Allegra S, Calabresi P, Picconi B, De Leonibus E (2018) Motor learning and metaplasticity in striatal neurons: Relevance for Parkinson's disease. *Brain* 141:505–520.

Goodman RN, Rietschel JC, Roy A, Jung BC, Diaz J, Macko RF, Forrester LW (2014) Increased reward in ankle robotics training enhances motor control and cortical efficiency in stroke. *J Rehabil Res Dev* 51:213–228 Available at: <http://www.rehab.research.va.gov/jour/2014/512/pdf/JRRD-2013-02-0050.pdf>.

Grandjean J, Derosiere G, Vassiliadis P, Quemener L, Wilde Y de, Duque J (2018) Towards assessing corticospinal excitability bilaterally: Validation of a double-coil TMS method. *J Neurosci Methods* 293:162–168 Available at: <http://linkinghub.elsevier.com/retrieve/pii/S016502701730345X>.

- Grandjean J, Duque J (2020) A TMS Study of Preparatory Suppression in Binge Drinkers. *NeuroImage Clin* 28:102383 Available at: <https://linkinghub.elsevier.com/retrieve/pii/S2213158220302205>.
- Grandjean J, Quoilin C, Duque J (2019) Investigating the effect of anticipating a startling acoustic stimulus on preparatory inhibition. *Neurophysiol Clin* 49:137–147 Available at: <https://doi.org/10.1016/j.neucli.2018.11.002>.
- Graybiel AM, Grafton ST (2015) The striatum: Where skills and habits meet. *Cold Spring Harb Perspect Biol* 7:1–14.
- Greenhouse I, Sias a., Labruna L, Ivry RB (2015) Nonspecific Inhibition of the Motor System during Response Preparation. *J Neurosci* 35:10675–10684 Available at: <http://www.jneurosci.org/content/35/30/10675.abstract?etoc>.
- Greenhouse XI, King XM, Noah XS, Maddock XRJ, Ivry RB (2017) Individual Differences in Resting Corticospinal Excitability Are Correlated with Reaction Time and GABA Content in Motor Cortex. *J Neurosci* 37:2686–2696.
- Gribble PL, Mullin LI, Cothros N, Mattar A (2003) Role of cocontraction in arm movement accuracy. *J Neurophysiol* 89:2396–2405.
- Grogan JP, Sandhu TR, Hu MT, Manohar SG (2020a) Dopamine promotes instrumental motivation, but reduces reward-related vigour. *bioRxiv*.
- Grogan JP, Sandhu TR, Hu MT, Manohar SG (2020b) Dopamine promotes instrumental motivation, but reduces reward-related vigour. *Elife* 9:1–20.
- Grool AM, Geerlings MI, Sigurdsson S, Eiriksdottir G, Jonsson P V., Garcia ME, Siggeirsdottir K, Harris TB, Sigmundsson T, Gudnason V, Launer LJ (2014) Structural MRI correlates of apathy symptoms in older persons without dementia: AGES-Reykjavik Study. *Neurology* 82:1628–1635.
- Groppa S, Oliviero A, Eisen A, Quartarone A, Cohen LG, Mall V, Kaelin-Lang A, Mima T, Rossi S, Thickbroom GW, Rossini PM, Ziemann U, Valls-Solé J, Siebner HR (2012) A practical guide to diagnostic transcranial magnetic stimulation: Report of an IFCN committee. *Clin Neurophysiol* 123:858–882 Available at: <http://dx.doi.org/10.1016/j.clinph.2012.01.010>.
- Grossman N (2018) Modulation without surgical intervention. *Science* (80-) 361:461–462.
- Grossman N, Bono D, Dedic N, Kodandaramaiah SB, Rudenko A, Suk H, Cassara AM, Neufeld E, Kuster N, Tsai L, Pascual-Leone A, Boyden ES (2017) Noninvasive Deep Brain Stimulation via Temporally Interfering Electric Fields. *Cell* 169:1029–1041.e16 Available at: <http://dx.doi.org/10.1016/j.cell.2017.05.024>.
- Haber SN (2016) Corticostriatal circuitry. *Dialogues Clin Neurosci* 18:7–21.
- Haber SN, Kim KS, Maily P, Calzavara R (2006) Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. *J Neurosci* 26:8368–8376.
- Hadjiiosif AM, Krakauer JW, Haith AM (2021) Did we get sensorimotor adaptation wrong? implicit adaptation as direct policy updating rather than forward-model-based learning. *J Neurosci* 41:2747–2761.
- Haith AM, Pakpoor J, Krakauer JW (2016) Independence of Movement Preparation and Movement Initiation. *J Neurosci* 36:3007–3015 Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.3245-15.2016> <http://www.ncbi.nlm.nih.gov/pubmed/26961954>.
- Hamel R, Savoie FA, Lacroix A, Whittingstall K, Trempe M, Bernier PM (2018) Added value of money on motor performance feedback: Increased left central beta-band power for rewards and fronto-central theta-band power for punishments. *Neuroimage* 179:63–78.
- Hampton WH, Alm KH, Venkatraman V, Nugiel T, Olson IR (2017) Dissociable frontostriatal white matter connectivity underlies reward and motor impulsivity. *Neuroimage* 150:336–343 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2017.02.021>.
- Hamzei F, Glauche V, Schwarzwald R, May A (2012) Dynamic gray matter changes within cortex and striatum after short motor skill training are associated with their increased functional interaction. *Neuroimage* 59:3364–3372 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2011.10.089>.
- Hanajima R, Ugawa Y, Machii K, Mochizuki H, Terao Y, Enomoto H, Furubayashi T, Shiio Y, Uesugi H, Kanazawa I (2001) Interhemispheric facilitation of the hand motor area in humans. *J Physiol* 531:849–859.
- Hannah R, Cavanagh SE, Tremblay S, Simeoni S, Rothwell JC (2018) Selective suppression of local interneuron circuits in human motor cortex contributes to movement preparation. *J Neurosci* 38:2869–17 Available at:

<http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.2869-17.2017>.

- Hardwick RM, Rajan VA, Bastian AJ, Krakauer JW, Celnik PA (2017) Motor Learning in Stroke. *Neurorehabil Neural Repair* 31:178–189.
- Hardwick RM, Rottschy C, Miall RC, Eickhoff SB (2013) A quantitative meta-analysis and review of motor learning in the human brain. *Neuroimage* 67:283–297 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2012.11.020>.
- Hare TA, Schultz W, Camerer CF, O'Doherty JP, Rangel A (2011) Transformation of stimulus value signals into motor commands during simple choice. *Proc Natl Acad Sci* 108:18120–18125 Available at: <http://www.pnas.org/cgi/doi/10.1073/pnas.1109322108>.
- Harris JA, Harris IM, Diamond ME (2001) The topography of tactile working memory. *J Neurosci* 21:8262–8269.
- Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In: *Power Technology and Engineering*, pp 139–183 Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0166411508623869>.
- Hasbroucq T, Kaneko H, Akamatsu M, Possamai CA (1997) Preparatory inhibition of cortico-spinal excitability: A transcranial magnetic stimulation study in man. *Cogn Brain Res* 5:185–192.
- Hasegawa M, Majima K, Itokazu T, Maki T, Albrecht UR, Castner N, Izumo M, Sohya K, Sato TK, Kamitani Y, Sato TR (2017) Selective Suppression of Local Circuits during Movement Preparation in the Mouse Motor Cortex. *Cell Rep* 18:2676–2686 Available at: <http://dx.doi.org/10.1016/j.celrep.2017.02.043>.
- Hasgall P, Di Gennaro F, Baumgartner F, Neufeld E, Lloyd B, Gosselin M, Payne D, Klingensböck A, Kuster N (2022) “IT’IS Database for thermal and electromagnetic parameters of biological tissues,” Version 4.1. Available at: itis.swiss/database.
- Heffley W, Song EY, Xu Z, Taylor BN, Hughes MA, McKinney A, Joshua M, Hull C (2018) Coordinated cerebellar climbing fiber activity signals learned sensorimotor predictions. *Nat Neurosci* 21:1431–1441 Available at: <http://dx.doi.org/10.1038/s41593-018-0228-8>.
- Hepworth LR, Rowe FJ, Walker MF, Rockliffe J (2016) Post-stroke Visual Impairment : A Systematic Literature Review of Types and Recovery of Visual Conditions Post-stroke Visual Impairment : A Systematic Literature Review of Types and Recovery of Visual Conditions.
- Herrojo-Ruiz M, Huebl J, Schönecker T, Kupsch A, Yarrow K, Krauss JK, Schneider GH, Kühn AA (2014) Involvement of human internal globus pallidus in the early modulation of cortical error-related activity. *Cereb Cortex* 24:1502–1517.
- Herrojo Ruiz M, Maudrich T, Kalloch B, Sammler D, Kenville R, Villringer A, Sehm B, Nikulin V V. (2021) Modulation of neural activity in frontopolar cortex drives reward-based motor learning. *Sci Rep* 11:20303 Available at: [10.1038/s41598-021-98571-y](https://doi.org/10.1038/s41598-021-98571-y).
- Herzfeld DJ, Kojima Y, Soetedjo R, Shadmehr R (2018) Encoding of error and learning to correct that error by the Purkinje cells of the cerebellum. *Nat Neurosci* 21:736–743 Available at: <http://dx.doi.org/10.1038/s41593-018-0136-y>.
- Hill CM, Stringer M, Waddell DE, Del Arco A (2020) Punishment Feedback Impairs Memory and Changes Cortical Feedback-Related Potentials During Motor Learning. *Front Hum Neurosci* 14:1–14.
- Holland P, Codol X, Oxley E, Taylor M, Hamshire E, Joseph S, Huffer L, Gale XM (2019) Domain-specific working memory, but not dopamine-related genetic variability, shapes reward-based motor learning. *J Neurosci* 39:9383–9396.
- Hori Y, Ihara N, Sugai C, Ogura J, Honda M, Kato K, Isomura Y, Hanakawa T (2019) Ventral striatum links motivational and motor networks during operant-conditioned movement in rats. *Neuroimage* 184:943–953 Available at: <https://doi.org/10.1016/j.neuroimage.2018.10.018>.
- Hosp JA, Luft AR (2013) Dopaminergic meso-cortical projections to M1: Role in motor learning and motor cortex plasticity. *Front Neurol* 4 OCT:1–7.
- Hosp JA, Nolan HE, Luft AR (2015) Topography and collateralization of dopaminergic projections to primary motor cortex in rats. *Exp Brain Res* 233:1365–1375.
- Hosp JA, Pekanovic A, Rioult-Pedotti MS, Luft AR (2011) Dopaminergic Projections from Midbrain to Primary Motor Cortex Mediate Motor Skill Learning. *J Neurosci* 31:2481–2487 Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.5411-10.2011>.

- Huang VS, Haith A, Mazzone 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 Available at: <http://dx.doi.org/10.1016/j.neuron.2011.04.012>.
- Hummel F, Celnik P, Giraux P, Floel A, Wu W, Gerloff C, Cohen LG (2005) Effects of non-invasive cortical stimulation on skilled motor function in chronic stroke. *1*:490–499.
- Husain M, Roiser JP (2018) Neuroscience of apathy and anhedonia: A transdiagnostic approach. *Nat Rev Neurosci* 19:470–484 Available at: <http://dx.doi.org/10.1038/s41583-018-0029-9>.
- Hussain SJ, Vollmer MK, Stimely J, Norato G, Zrenner C, Ziemann U, Buch ER, Cohen LG (2021) Phase-dependent offline enhancement of human motor memory. *Brain Stimul* 14:873–883 Available at: <https://doi.org/10.1016/j.brs.2021.05.009>.
- Hutcheon B, Yarom Y (2000) Resonance, oscillation and the intrinsic frequency preferences of neurons. *Trends Neurosci* 23:216–222 Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0166223600015472>.
- Iacono MI, Neufeld E, Akinagbe E, Bower K, Wolf J, Oikonomidis IV, Sharma D, Lloyd B, Wilm BJ, Wyss M, Pruessmann KP, Jakab A, Makris N, Cohen ED, Kuster N, Kainz W, Angelone LM (2015) MIDA: A multimodal imaging-based detailed anatomical model of the human head and neck. *PLoS One* 10.
- Igata H, Ikegaya Y, Sasaki T (2021) Prioritized experience replays on a hippocampal predictive map for learning. *Proc Natl Acad Sci U S A* 118:1–9.
- Igloi K, Gaggioni G, Sterpenich V, Schwartz S (2015) A nap to recap or how reward regulates hippocampal-prefrontal memory networks during daytime sleep in humans. *Elife* 4:1–16.
- Ito M (2001) Cerebellar long-term depression: Characterization, signal transduction, and functional roles. *Physiol Rev* 81:1143–1195.
- Itthipuripat S, Vo VA, Sprague TC, Serences JT (2019) Value-driven attentional capture enhances distractor representations in early visual cortex. *PLoS Biol* 17.
- Izawa J, Shadmehr R (2011) Learning from sensory and reward prediction errors during motor adaptation. *PLoS Comput Biol* 7:1–12.
- Jacobacci F, Armony JL, Yeffal A, Lerner G, Amaro E, Jovicich J, Doyone J, Della-Maggiore V (2020) Rapid hippocampal plasticity supports motor sequence learning. *Proc Natl Acad Sci U S A* 117:23898–23903.
- Jäncke L, Koeneke S, Hoppe A, Rominger C, Hänggi J (2009) The Architecture of the Golfer's Brain Earley RL, ed. *PLoS One* 4:e4785 Available at: <https://dx.plos.org/10.1371/journal.pone.0004785>.
- Jayaram G, Galea JM, Bastian AJ, Celnik P (2011) Human locomotor adaptive learning is proportional to depression of cerebellar excitability. *Cereb Cortex* 21:1901–1909.
- Jenkinson N, Brown P (2011) New insights into the relationship between dopamine, beta oscillations and motor function. *Trends Neurosci* 34:611–618 Available at: <http://dx.doi.org/10.1016/j.tins.2011.09.003>.
- Jiang T (2013) Brainnetome: A new -ome to understand the brain and its disorders. *Neuroimage* 80:263–272 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2013.04.002>.
- Johnson BP, Cohen LG (2022) Reward and plasticity: Implications for neurorehabilitation. In, pp 331–340 Available at: <https://linkinghub.elsevier.com/retrieve/pii/B9780128194102000187>.
- Johnson BP, Cohen LG, Westlake KP (2021) The Intersection of Offline Learning and Rehabilitation. *Front Hum Neurosci* 15:1–6.
- Johnson L, Alekseichuk I, Krieg J, Doyle A, Yu Y, Vitek J, Johnson M, Opitz A (2020) Dose-dependent effects of transcranial alternating current stimulation on spike timing in awake nonhuman primates. *Sci Adv* 6:1–9 Available at: <https://www.science.org/doi/10.1126/sciadv.aaz2747>.
- Kalenscher T, Lansink CS, Lankelma J V., Pennartz CMA (2010) Reward-associated gamma oscillations in ventral striatum are regionally differentiated and modulate local firing activity. *J Neurophysiol* 103:1658–1672.
- Kami A, Meyer G, Jezzard P, Adams MM, Turner R, Ungerleider LG (1995) Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature* 377:155–158 Available at: <https://www.nature.com/articles/377155a0.pdf>.

- Kao MH, Doupe AJ, Brainard MS (2005) Contribution of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature* 433:638–643.
- Kaplan BA, Amlung M, Reed DD, Jarmolowicz DP, McKerchar TL, Lemley SM (2016) Automating Scoring of Delay Discounting for the 21- and 27-Item Monetary Choice Questionnaires. *Behav Anal* 39:293–304 Available at: <http://dx.doi.org/10.1007/s40614-016-0070-9>.
- Kastner S, De Weerd P, Desimone R, Ungerleider LG (1998) Mechanisms of Directed Attention in the Human Extrastriate Cortex as Revealed by Functional MRI. *Science* (80-) 282:108–111 Available at: <https://www.sciencemag.org/lookup/doi/10.1126/science.282.5386.108>.
- Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG (1999) Increased Activity in Human Visual Cortex during Directed Attention in the Absence of Visual Stimulation. *Neuron* 22:751–761 Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0896627300807345>.
- Kim HE, Parvin DE, Ivry RB (2019) The influence of task outcome on implicit motor learning. *Elife* 8:1–28.
- Kirby KN, Petry NM, Bickel WK (1999) Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *J Exp Psychol Gen* 128:78–87.
- Kitago T, Krakauer JW (2013) Motor learning principles for neurorehabilitation, 1st ed. Elsevier B.V. Available at: <http://dx.doi.org/10.1016/B978-0-444-52901-5.00008-3>.
- Kleim JA, Hogg TM, VandenBerg PM, Cooper NR, Bruneau R, Remple M (2004) Cortical Synaptogenesis and Motor Map Reorganization Occur during Late, but Not Early, Phase of Motor Skill Learning. *J Neurosci* 24:628–633.
- Klein-Flügge MC, Hunt LT, Bach DR, Dolan RJ, Behrens TEJ (2011) Dissociable Reward and Timing Signals in Human Midbrain and Ventral Striatum. *Neuron* 72:654–664.
- Klein P-A, Olivier E, Duque J (2012) Influence of Reward on Corticospinal Excitability during Movement Preparation. *J Neurosci* 32:18124–18136.
- Klein PA, Petitjean C, Olivier E, Duque J (2014) Top-down suppression of incompatible motor activations during response selection under conflict. *Neuroimage* 86:138–149 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2013.08.005>.
- Kobayashi S, Schultz W (2008) Influence of reward delays on responses of dopamine neurons. *J Neurosci* 28:7837–7846.
- Kondabolu K, Roberts EA, Bucklin M, McCarthy MM, Kopell N, Han X (2016) Striatal cholinergic interneurons generate beta and gamma oscillations in the corticostriatal circuit and produce motor deficits. *Proc Natl Acad Sci U S A* 113:3159–3168.
- Kostadinov D, Beau M, Pozo MB, Häusser M (2019) Predictive and reactive reward signals conveyed by climbing fiber inputs to cerebellar Purkinje cells. *Nat Neurosci* 22 Available at: <http://dx.doi.org/10.1038/s41593-019-0381-8>.
- Kostadinov D, Häusser M (2022) Reward signals in the cerebellum: origins, targets, and functional implications. *Neuron*:1–14.
- Krakauer JW, Ghilardi MF, Mentis M, Barnes A, Veytsman M, Eidelberg D, Ghez C (2004) Differential Cortical and Subcortical Activations in Learning Rotations and Gains for Reaching: A PET Study. *J Neurophysiol* 91:924–933.
- Krakauer JW, Hadjiosif AM, Xu J, Wong AL, Haith AM (2019) Motor Learning. *9*:613–663.
- Krakauer JW, Mazzoni P, Ghazizadeh A, Ravindran R, Shadmehr R (2006) Generalization of motor learning depends on the history of prior action. *PLoS Biol* 4:1798–1808.
- Krause MR, Vieira PG, Csorba BA, Pilly PK, Pack CC (2019) Transcranial alternating current stimulation entrains single-neuron activity in the primate brain. *Proc Natl Acad Sci U S A* 116:5747–5755.
- Krause MR, Vieira PG, Thivierge JP, Pack CC (2022) Brain stimulation competes with ongoing oscillations for control of spike timing in the primate brain. *PLoS Biol* 20:1–27 Available at: <http://dx.doi.org/10.1371/journal.pbio.3001650>.
- Kupferschmidt DA, Juczewski K, Cui G, Johnson KA, Lovinger DM (2017) Parallel, but Dissociable, Processing in Discrete Corticostriatal Inputs Encodes Skill Learning. *Neuron* 96:476–489.e5 Available at: <https://doi.org/10.1016/j.neuron.2017.09.040>.
- Labruna L, Fernández-Del-Olmo M, Ivry RB (2011) Comparison of different baseline conditions in evaluating factors that influence

- motor cortex excitability. *Brain Stimul* 4:152–155 Available at: <http://dx.doi.org/10.1016/j.brs.2010.09.010>.
- Lam JM, Wächter T, Globas C, Karnath HO, Luft AR (2013) Predictive value and reward in implicit classification learning. *Hum Brain Mapp* 34:176–185.
- Lardi C, Billieux J, D’Acromont M, Linden M Van der (2008) A French adaptation of a short version of the Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ). *Pers Individ Dif* 45:722–725 Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0191886908002717>.
- Lee C, Harkin E, Naud R, Chen SX (2021) Cell-Type Specific Responses to Associative Learning in the Primary Motor Cortex. *bioRxiv*:2021.08.08.455571 Available at: <https://www.biorxiv.org/content/10.1101/2021.08.08.455571v1%0Ahttps://www.biorxiv.org/content/10.1101/2021.08.08.455571v1.abstract>.
- Lee D, Seo H, Jung M (2012) Neural Basis of Reinforcement Learning and Decision Making. *Annu Rev Neurosci* 35:287–308.
- Lee J, Wang W, Sabatini BL (2020) Anatomically segregated basal ganglia pathways allow parallel behavioral modulation. *Nat Neurosci* 23:1388–1398 Available at: <http://dx.doi.org/10.1038/s41593-020-00712-5>.
- Leemburg S, Canonica T, Luft A (2018) Motor skill learning and reward consumption differentially affect VTA activation. *Sci Rep* 8:1–12 Available at: <http://dx.doi.org/10.1038/s41598-017-18716-w>.
- Leow LA, Marinovic W, de Rugy A, Carroll TJ (2018) Task errors contribute to implicit aftereffects in sensorimotor adaptation. *Eur J Neurosci*:0–2.
- Levy S, Lavzin M, Benisty H, Ghanayim A, Dubin U, Achvat S, Brosh Z, Aeed F, Mensh BD, Schiller Y, Meir R, Barak O, Talmon R, Hantman AW, Schiller J (2020) Cell-Type-Specific Outcome Representation in the Primary Motor Cortex. *Neuron* 107:954–971.e9 Available at: <https://doi.org/10.1016/j.neuron.2020.06.006>.
- Lewthwaite R, Wulf G (2017) Optimizing motivation and attention for motor performance and learning. *Curr Opin Psychol* 16:38–42 Available at: <http://dx.doi.org/10.1016/j.copsyc.2017.04.005>.
- Lindström B, Golkar A, Jangard S, Tobler PN, Olsson A (2019) Social threat learning transfers to decision making in humans. *Proc Natl Acad Sci U S A* 116:4732–4737.
- Littman ML (2015) Reinforcement learning improves behaviour from evaluative feedback. *Nature* 521:445–451.
- Liu H, Peng X, Dahmani L, Wang H, Zhang M, Shan Y, Rong D, Guo Y, Li J, Li N, Wang L, Lin Y, Pan R, Lu J, Wang D (2020) Patterns of motor recovery and structural neuroplasticity after basal ganglia infarcts. *Neurology* 95:E1174–E1187.
- Liu Y, Dolan RJ, Kurth-Nelson Z, Behrens TEJ (2019) Human Replay Spontaneously Reorganizes Experience. *Cell* 178:640–652.e14 Available at: <https://doi.org/10.1016/j.cell.2019.06.012>.
- Liu Y, Mattar MG, Behrens TEJ, Daw ND, Dolan RJ (2021) Experience replay is associated with efficient nonlocal learning. *Science* (80-) 372.
- Luke SG (2017) Evaluating significance in linear mixed-effects models in R. *Behav Res Methods* 49:1494–1502.
- Ma I, van Holstein M, Mies GW, Mennes M, Buitelaar J, Cools R, Cillessen AHN, Krebs RM, Scheres A (2016) Ventral striatal hyperconnectivity during rewarded interference control in adolescents with ADHD. *Cortex* 82:225–236 Available at: <http://dx.doi.org/10.1016/j.cortex.2016.05.021>.
- Ma R, Xia X, Zhang W, Lu Z, Wu Q, Cui J, Song H, Fan C, Chen X, Zha R, Wei J, Ji G-J, Wang X, Qiu B, Zhang X (2022) High Gamma and Beta Temporal Interference Stimulation in the Human Motor Cortex Improves Motor Functions. *Front Neurosci* 15:1–11.
- Mahon S, Deniau JM, Charpier S (2004) Corticostriatal plasticity: Life after the depression. *Trends Neurosci* 27:460–467.
- Makino H, Ren C, Liu H, Kim AN, Kondapaneni N, Liu X, Kuzum D, Komiyama T (2017) Transformation of Cortex-wide Emergent Properties during Motor Learning. *Neuron* 94:880–890.e8 Available at: <http://dx.doi.org/10.1016/j.neuron.2017.04.015>.
- Malenka RC, Bear MF (2004) LTP and LTD: An embarrassment of riches. *Neuron* 44:5–21.
- Mandelblat-cerf Y, Paz R, Vaadia E (2009) Trial-to-Trial Variability of Single Cells in Motor Cortices Is Dynamically Modified during Visuomotor Adaptation. *J Neurosci* 29:15053–15062.

- Manohar SG, Chong TT, Apps MAJ, Jarman PR, Bhatia KP, Husain M, Manohar SG, Chong TT, Apps MAJ, Batla A, Stamelou M (2015) Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control Article Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control. *Curr Biol* 25:1707–1716 Available at: <http://dx.doi.org/10.1016/j.cub.2015.05.038>.
- Manohar SG, Muhammed K, Fallon SJ, Husain M (2019) Motivation dynamically increases noise resistance by internal feedback during movement. *Neuropsychologia* 123:19–29 Available at: <https://doi.org/10.1016/j.neuropsychologia.2018.07.011>.
- Mas-Herrero E, Dagher A, Farrés-Franch M, Zatorre RJ (2021) Unraveling the temporal dynamics of reward signals in music-induced pleasure with TMS. *J Neurosci* 41:3889–3899.
- Mathis MW, Mathis A, Uchida N (2017) Somatosensory Cortex Plays an Essential Role in Forelimb Motor Adaptation in Mice. *Neuron* 93:1493–1503.e6 Available at: <http://dx.doi.org/10.1016/j.neuron.2017.02.049>.
- Matsumoto H, Tian J, Uchida N, Watabe-Uchida M (2016) Midbrain dopamine neurons signal aversion in a reward-context-dependent manner. *Elife* 5:1–24.
- Matsumoto J, Urakawa S, Hori E, de Araujo MFP, Sakuma Y, Ono T, Nishijo H (2012) Neuronal responses in the nucleus accumbens shell during sexual behavior in male rats. *J Neurosci* 32:1672–1686.
- Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459:837–841.
- Mawase F, Lopez D, Celnik PA, Haith AM (2018) Movement Repetition Facilitates Response Preparation. *Cell Rep* 24:801–808 Available at: <https://doi.org/10.1016/j.celrep.2018.06.097>.
- Mawase F, Uehara S, Bastian AJ, Celnik P (2017) Motor Learning Enhances Use-Dependent Plasticity. *J Neurosci* 37:2673–2685 Available at: <http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.3303-16.2017>.
- McCarthy MM, Moore-Kochlacs C, Gu X, Boyden ES, Han X, Kopell N (2011) Striatal origin of the pathologic beta oscillations in Parkinson's disease. *Proc Natl Acad Sci U S A* 108:11620–11625.
- McDougle SD, Boggess MJ, Crossley MJ, Parvin D, Ivry RB, Taylor JA (2016) Credit assignment in movement-dependent reinforcement learning. *Proc Natl Acad Sci* 113:6797–6802 Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1523669113>.
- McDougle SD, Bond KM, Taylor JA (2015) Explicit and Implicit Processes Constitute the Fast and Slow Processes of Sensorimotor Learning. *J Neurosci* 35:9568–9579 Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.5061-14.2015>.
- McDougle SD, Butcher PA, Parvin D, Mushtaq F, Niv Y, Ivry RB, Taylor JA (2018) Neural signatures of prediction errors in a decision-making task are modulated by action execution failures. :1–35.
- McDougle SD, Wilterson SA, Turk-Browne NB, Taylor JA (2022) Revisiting the Role of the Medial Temporal Lobe in Motor Learning. *J Cogn Neurosci* 34:532–549 Available at: <https://direct.mit.edu/jocn/article/34/3/532/108899/Revisiting-the-Role-of-the-Medial-Temporal-Lobe-in>.
- McGrane N, Galvin R, Cusack T, Stokes E (2015) Addition of motivational interventions to exercise and traditional Physiotherapy: A review and meta-analysis. *Physiother (United Kingdom)* 101:1–12 Available at: <http://dx.doi.org/10.1016/j.physio.2014.04.009>.
- McLaren DG, Ries ML, Xu G, Johnson SC (2012) A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *Neuroimage* 61:1277–1286 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2012.03.068>.
- Mehler-Wex C, Riederer P, Gerlach M (2006) Dopaminergic dysbalance in distinct basal ganglia neurocircuits: Implications for the pathophysiology of parkinson's disease, schizophrenia and attention deficit hyperactivity disorder. *Neurotox Res* 10:167–179.
- Mehrkanoon S, Boonstra TW, Breakspear M, Hinder M, Summers JJ (2016) Upregulation of cortico-cerebellar functional connectivity after motor learning. *Neuroimage* 128:252–263 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2015.12.052>.
- Miall RC, Afanasyeva D, Cole JD, Mason P (2021) The role of somatosensation in automatic visuo-motor control: a comparison of congenital and acquired sensory loss. *Exp Brain Res* 239:2043–2061 Available at: <https://doi.org/10.1007/s00221-021-06110->

y.

- Miall RC, Weir DJ, Stein JF (1985) Visuomotor tracking with delayed visual feedback. *Neuroscience* 16:511–520.
- Mirdamadi JL, Block HJ (2020) Somatosensory changes associated with motor skill learning. *J Neurophysiol* 123:1052–1062.
- Mirzakhilili E, Barra B, Capogrosso M, Lempka SF (2020) Biophysics of Temporal Interference Stimulation. *Cell Syst* 11:557-572.e5 Available at: <https://doi.org/10.1016/j.cels.2020.10.004>.
- Mitchell JM, Fields HL, D’Esposito M, Boettiger CA (2005) Impulsive responding in alcoholics. *Alcohol Clin Exp Res* 29:2158–2169.
- Mitchell MR, Potenza MN (2014) Recent Insights into the Neurobiology of Impulsivity. *Curr Addict Reports* 1:309–319.
- Mohebi A, Pettibone JR, Hamid AA, Wong JT, Vinson LT, Patriarchi T, Tian L, Kennedy RT, Berke JD (2019) Dissociable dopamine dynamics for learning and motivation. *Nature* 570:65–70 Available at: <http://dx.doi.org/10.1038/s41586-019-1235-y>.
- Morehead JR, Orban De Xivry JJ (2021) A Synthesis of the Many Errors and Learning Processes of Visuomotor Adaptation. *bioRxiv*:1–50.
- Morton SM, Bastian AJ (2006) Cerebellar contributions to locomotor adaptations during splitbelt treadmill walking. *J Neurosci* 26:9107–9116.
- Mosley PE, Paliwal S, Robinson K, Coyne T, Silburn P, Tittgemeyer M, Stephan KE, Breakspear M, Perry A (2019) The structural connectivity of discrete networks underlies impulsivity and gambling in Parkinson’s disease. *Brain* 142:3917–3935.
- Muellbacher W, Ziemann U, Wissel J, Dang N, Kofler M, Facchini S, Boroojerdi B, Poewe W, Hallett M (2002) Early consolidation in human primary motor cortex. *Nature* 415:640–644.
- Muratori LM, Lamberg EM, Quinn L, Duff S V. (2013) Applying principles of motor learning and control to upper extremity rehabilitation. *J Hand Ther* 26:94–103 Available at: <http://dx.doi.org/10.1016/j.jht.2012.12.007>.
- Mure H, Tang CC, Argyelan M, Ghilardi MF, Kaplitt MG, Dhawan V, Eidelberg D (2012) Improved sequence learning with subthalamic nucleus deep brain stimulation: Evidence for treatment-specific network modulation. *J Neurosci* 32:2804–2813.
- Nabavi S, Fox R, Proulx CD, Lin JY, Tsien RY, Malinow R (2014) Engineering a memory with LTD and LTP. *Nature* 511:348–352 Available at: <http://dx.doi.org/10.1038/nature13294>.
- Nakamura K, Hikosaka O (2006) Facilitation of saccadic eye movements by postsaccadic electrical stimulation in the primate caudate. *J Neurosci* 26:12885–12895.
- Neftci EO, Averbeck BB (2019) Reinforcement learning in artificial and biological systems. *Nat Mach Intell* 1:133–143 Available at: <http://dx.doi.org/10.1038/s42256-019-0025-4>.
- Nickchen K, Boehme R, del Mar Amador M, Hälbig TD, Dehnicke K, Panneck P, Behr J, Prass K, Heinz A, Deserno L, Schlagenhauf F, Priller J (2017) Reversal learning reveals cognitive deficits and altered prediction error encoding in the ventral striatum in Huntington’s disease. *Brain Imaging Behav* 11:1862–1872.
- Nieuwenhuis R, te Grotenhuis M, Pelzer B (2012) Influence.ME: Tools for detecting influential data in mixed effects models. *R J* 4:38–47.
- Nikooyan AA, Ahmed AA (2015) Reward feedback accelerates motor learning. *J Neurophysiol* 113:633–646 Available at: <http://jn.physiology.org/lookup/doi/10.1152/jn.00032.2014>.
- Nikooyan AA, Ahmed AA, Nikooyan AA, Ahmed AA (2015) Reward feedback accelerates motor learning Reward feedback accelerates motor learning. :633–646.
- Norman SL, Lobo-Prat J, Reinkensmeyer DJ (2017) How do strength and coordination recovery interact after stroke? A computational model for informing robotic training. *IEEE Int Conf Rehabil Robot*:181–186.
- O’Doherty JP (2004) Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Curr Opin Neurobiol* 14:769–776.
- O’Doherty JP, Cockburn J, Pauli WM (2017) Learning, Reward, and Decision Making. *Annu Rev Psychol* 68:73–100 Available at: <http://www.annualreviews.org/doi/10.1146/annurev-psych-010416-044216>.

- Ohmae S, Medina JF (2015) Climbing fibers encode a temporal-difference prediction error during cerebellar learning in mice. *Nat Neurosci* 29:1–7.
- Okamoto M, Dan H, Sakamoto K, Takeo K, Shimizu K, Kohno S, Oda I, Isobe S, Suzuki T, Kohyama K, Dan I (2004) Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10-20 system oriented for transcranial functional brain mapping. *Neuroimage* 21:99–111.
- Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9:97–113 Available at: <https://linkinghub.elsevier.com/retrieve/pii/0028393271900674>.
- Ölveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol* 3:0902–0909.
- Ozen S, Sirota A, Belluscio MA, Anastassiou CA, Stark E, Koch C, Buzsáki G (2010) Transcranial electric stimulation entrains cortical neuronal populations in rats. *J Neurosci* 30:11476–11485.
- Padoa-Schioppa C (2011) Neurobiology of Economic Choice: A Good-Based Model. *Annu Rev Neurosci* 34:333–359 Available at: <http://www.annualreviews.org/doi/10.1146/annurev-neuro-061010-113648>.
- Palidis DJ, Cashaback JGA, Gribble PL (2019) Neural signatures of reward and sensory error feedback processing in motor learning. *J Neurophysiol* 121:1561–1574.
- Palidis DJ, McGregor HR, Vo A, MacDonald PA, Gribble PL (2021) Null effects of levodopa on reward- and error-based motor adaptation, savings, and anterograde interference. *J Neurophysiol* 126:47–67 Available at: <https://journals.physiology.org/doi/10.1152/jn.00696.2020>.
- Palminteri S, Lebreton M, Worbe Y, Hartmann A, Lehéricy S, Vidailhet M, Grabi D, Pessiglione M (2011) Dopamine-dependent reinforcement of motor skill learning: Evidence from Gilles de la Tourette syndrome. *Brain* 134:2287–2301.
- Pantev C, Engelien A, Candia V, Elbert T (2001) Representational cortex in musicians plastic alterations in response to musical practice. *Ann N Y Acad Sci* 930:300–314.
- Parvin DE, McDougale SD, Taylor JA, Ivry RB (2018) Credit assignment in a motor decision making task is influenced by agency and not sensorimotor prediction errors. *J Neurosci* 38:3601–3617 Available at: <http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.3601-17.2018>.
- Pascual-Leone A et al. (1995) Modulation of muscle responses evoked by transcranial magnetic stimulation during the acquisition of new fine motor skills. *J Neurophysiol* 74:1037–1045 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7500130>.
- Pascual-Leone A, Amedi A, Fregni F, Merabet LB (2005) The plastic human brain cortex. *Annu Rev Neurosci* 28:377–401.
- Pavlov I (1927) *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex*. Oxford Univ Press.
- Paz R, Boraud T, Natan C, Bergman H, Vaadia E (2003) Preparatory activity in motor cortex reflects learning of local visuomotor skills. *6:882–890*.
- Pearson K, Filon L (1898) VII. Mathematical contributions to the theory of evolution. — IV. On the probable errors of frequency constants and on the influence of random selection on variation and correlation. *Philos Trans R Soc London* 191:229–311.
- Pekny SE, Izawa J, Shadmehr R (2015) Reward-Dependent Modulation of Movement Variability. *J Neurosci* 35:4015–4024 Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.3244-14.2015>.
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442 Available at: <http://booksandjournals.brillonline.com/content/journals/10.1163/156856897x00366>.
- Peterburs J, Kobza S, Bellebaum C (2016) Feedback delay gradually affects amplitude and valence specificity of the feedback-related negativity (FRN). *Psychophysiology* 53:209–215.
- Piray P, den Ouden HEM, van der Schaaf ME, Toni I, Cools R (2017) Dopaminergic modulation of the functional ventrodorsal architecture of the human striatum. *Cereb Cortex* 27:485–495.
- Polanía R, Nitsche MA, Ruff CC (2018) Studying and modifying brain function with non-invasive brain stimulation. *Nat Neurosci* 21:174–187 Available at: <http://dx.doi.org/10.1038/s41593-017-0054-4>.

- Pool ER, Pauli WM, Kress CS, O'Doherty JP (2019) Behavioural evidence for parallel outcome-sensitive and outcome-insensitive Pavlovian learning systems in humans. *Nat Hum Behav* 3:284–296 Available at: <http://dx.doi.org/10.1038/s41562-018-0527-9>.
- Popp NJ, Hernandez-Castillo CR, Gribble PL, Diedrichsen J (2022) The role of feedback in the production of skilled finger sequences. *J Neurophysiol* 127:829–839.
- Popp NJ, Yokoi A, Gribble PL, Diedrichsen J (2020) The effect of instruction on motor skill learning. *J Neurophysiol* 124:1449–1457.
- Quattrocchi G, Greenwood R, Rothwell JC, Galea JM, Bestmann S (2017) Reward and punishment enhance motor adaptation in stroke. *J Neurol Neurosurg Psychiatry:jnnp-2016-314728* Available at: <http://jnnp.bmj.com/lookup/doi/10.1136/jnnp-2016-314728>.
- Quattrocchi G, Monaco J, Ho A, Irmen F, Strube W, Ruge D, Bestmann S, Galea JM (2018) Pharmacological Dopamine Manipulation Does Not Alter Reward-Based Improvements in Memory Retention during a Visuomotor Adaptation Task. *Eneuro* 5:ENEURO.0453-17.2018 Available at: <http://eneuro.sfn.org/lookup/doi/10.1523/ENEURO.0453-17.2018>.
- Quoilin C, Fievez F, Duque J (2019) Preparatory inhibition: Impact of choice in reaction time tasks. *Neuropsychologia* 129:212–222 Available at: <https://doi.org/10.1016/j.neuropsychologia.2019.04.016>.
- Quoilin C, Grandjean J, Duque J (2020) Considering Motor Excitability During Action Preparation in Gambling Disorder: A Transcranial Magnetic Stimulation Study. *Front Psychiatry* 11:1–13.
- Quoilin C, Lambert J, Jacob B, Klein PA, Duque J (2016) Comparison of motor inhibition in variants of the instructed-delay choice reaction time task. *PLoS One* 11:1–16.
- Quoilin C, Wilhelm E, Maurage P, Timary P De, Duque J (2018) Deficient inhibition in alcohol-dependence : let's consider the role of the motor system ! *Neuropsychopharmacology*:0–8 Available at: <http://dx.doi.org/10.1038/s41386-018-0074-0>.
- Raffin E, Hummel FC (2018) Restoring Motor Functions After Stroke: Multiple Approaches and Opportunities. *Neuroscientist* 24:400–416.
- Ramakrishnan A, Byun YW, Rand K, Pedersen CE, Lebedev MA, Nicolelis MAL (2017) Cortical neurons multiplex reward-related signals along with sensory and motor information. *Proc Natl Acad Sci* 114:E4841–E4850 Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1703668114>.
- Ramkumar P, Dekleva B, Cooler S, Miller L, Kording K (2016) Premotor and motor cortices encode reward. *PLoS One* 11:1–13.
- Ramnani N, Elliott R, Athwal BS, Passingham RE (2004) Prediction error for free monetary reward in the human prefrontal cortex. *Neuroimage* 23:777–786.
- Rampersad S, Roig-Solvas B, Yarossi M, Kulkarni PP, Santarnecchi E, Dorval AD, Brooks DH (2019) Prospects for transcranial temporal interference stimulation in humans: A computational study. *Neuroimage* 202:116124 Available at: <https://doi.org/10.1016/j.neuroimage.2019.116124>.
- Ranganathan R, Tomlinson AD, Lokesh R, Lin T, Patel P (2019) A tale of too many tasks : Fragmentation of tasks in motor learning and a call for model task paradigms.
- Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci* 9:545–556.
- Raspopovic S et al. (2014) Restoring Natural Sensory Feedback in Real-Time Bidirectional Hand Prostheses. 6.
- Reis J, Schambra HM, Cohen LG, Buch ER, Fritsch B, Zarah E, Celnik PA, Krakauer JW (2009) Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. *Proc Natl Acad Sci* 106:1590–1595 Available at: <http://www.pnas.org/cgi/doi/10.1073/pnas.0805413106>.
- Reis J, Swayne OB, Vandermeeren Y, Camus M, Dimyan MA, Harris-Love M, Perez MA, Ragert P, Rothwell JC, Cohen LG (2008) Contribution of transcranial magnetic stimulation to the understanding of cortical mechanisms involved in motor control. *J Physiol* 586:325–351 Available at: <http://doi.wiley.com/10.1113/jphysiol.2007.144824>.
- Reppert TR, Lempert KM, Glimcher PW, Shadmehr R (2015) Modulation of saccade vigor during value-based decision making. *J Neurosci* 35:15369–15378.

- Rescorla R, Wagner A (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning II: Current research. and theory*:pp64-99.
- Rieckmann A, Fischer H, Bäckman L (2010) Activation in striatum and medial temporal lobe during sequence learning in younger and older adults: Relations to performance. *Neuroimage* 50:1303–1312 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2010.01.015>.
- Riout-Pedotti MS, Friedman D, Donoghue JP (2000) Learning-induced LTP in neocortex. *Science* (80-) 290:533–536.
- Riout-Pedotti MS, Friedman D, Hess G, Donoghue JP (1998) Strengthening of horizontal cortical connections following skill learning. *Nat Neurosci* 1:230–234.
- Robertson EM, Pascual-leone A, Miall RC (2004) Current concepts in procedural consolidation. *Nat Rev Neurosci* 5:1–7.
- Robertson EM, Press DZ, Pascual-Leone A (2005) Off-line learning and the primary motor cortex. *J Neurosci* 25:6372–6378.
- Roemmich RT, Bastian AJ (2018) Closing the Loop: From Motor Neuroscience to Neurorehabilitation. *Annu Rev Neurosci* 41:415–429 Available at: <https://www.annualreviews.org/doi/10.1146/annurev-neuro-080317-062245>.
- Rose M, Haider H, Salari N, Büchel C (2011) Functional dissociation of Hippocampal mechanism during implicit learning based on the domain of associations. *J Neurosci* 31:13739–13745.
- Rosenkranz K, Kacar A, Rothwell JC (2007) Differential modulation of motor cortical plasticity and excitability in early and late phases of human motor learning. *J Neurosci* 27:12058–12066.
- Rothé M, Quilodran R, Sallet J, Procyk E (2011) Coordination of high gamma activity in anterior cingulate and lateral prefrontal cortical areas during adaptation. *J Neurosci* 31:11110–11117.
- Rubin DB, Hosman T, Kelemen JN, Kapitonava A, Willett FR, Coughlin BF, Halgren E, Kimchi EY, Williams ZM, Simeral JD, Hochberg LR, Cash SS (2022) Learned Motor Patterns Are Replayed in Human Motor Cortex during Sleep. *J Neurosci* 42:5007–5020.
- Sampaio-Baptista C, Sanders ZB, Johansen-Berg H (2018) Structural plasticity in adulthood with motor learning and stroke rehabilitation. *Annu Rev Neurosci* 41:25–40.
- Schendan HE, Searl MM, Melrose RJ, Stern CE (2003) An fMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron* 37:1013–1025.
- Schlerf J, Ivry RB (2012) Encoding of Sensory Prediction Errors in the Human Cerebellum. 32:4913–4922.
- Schlerf JE, Galea JM, Bastian AJ, Celnik PA (2012) Dynamic Modulation of Cerebellar Excitability for Abrupt , But Not Gradual , Visuomotor Adaptation. 32:11610–11617.
- Schmidt L, D’Arc BF, Lafargue G, Galanaud D, Czernecki V, Grabli D, Schüpbach M, Hartmann A, Lévy R, Dubois B, Pessiglione M (2008) Disconnecting force from money: Effects of basal ganglia damage on incentive motivation. *Brain* 131:1303–1310.
- Schmidt RA, Bjork RA (1992) New Conceptualizations of Practice: Common Principles in Three Paradigms Suggest New Concepts for Training. *Psychol Sci* 3:207–217.
- Schonewille M, Belmeguenai A, Koekkoek SK, Houtman SH, Boele HJ, van Beugen BJ, Gao Z, Badura A, Ohtsuki G, Amerika WE, Hosy E, Hoebeek FE, Elgersma Y, Hansel C, De Zeeuw CI (2010) Purkinje cell-specific knockout of the protein phosphatase PP2B impairs potentiation and cerebellar motor learning. *Neuron* 67:618–628 Available at: <http://dx.doi.org/10.1016/j.neuron.2010.07.009>.
- Schonewille M, Girasole AE, Rostaing P, Mailhes-Hamon C, Ayon A, Nelson AB, Triller A, Casado M, De Zeeuw CI, Bouvier G (2021) NMDARs in granule cells contribute to parallel fiber-Purkinje cell synaptic plasticity and motor learning. *Proc Natl Acad Sci U S A* 118:1–9.
- Schubert C, Dabbagh A, Classen J, Krämer UM, Tzvi E (2021) Alpha oscillations modulate premotor-cerebellar connectivity in motor learning: Insights from transcranial alternating current stimulation. *Neuroimage* 241:118410 Available at: <https://doi.org/10.1016/j.neuroimage.2021.118410>.
- Schultz W (2015) Neuronal Reward and Decision Signals: From Theories to Data. *Physiol Rev* 95:853–951 Available at: <http://physrev.physiology.org/lookup/doi/10.1152/physrev.00023.2014>.

- Schultz W, Dayan P, Montague PR (1997) A Neural Substrate of Prediction and Reward. *Science* (80-) 275:1593–1599 Available at: <http://www.sciencemag.org/cgi/doi/10.1126/science.275.5306.1593>.
- Searle SR, Speed FM, Milliken GA (1980) Population marginal means in the linear model: An alternative to least squares means. *Am Stat* 34:216–221.
- Seeck M, Koessler L, Bast T, Leijten F, Michel C, Baumgartner C, He B, Beniczky S (2017) The standardized EEG electrode array of the IFCN. *Clin Neurophysiol* 128:2070–2077 Available at: <http://dx.doi.org/10.1016/j.clinph.2017.06.254>.
- Sendhilnathan N, Ipata AE, Goldberg ME (2020) Neural Correlates of Reinforcement Learning in Mid-lateral Cerebellum. *Neuron* 106:188-198.e5 Available at: <https://doi.org/10.1016/j.neuron.2019.12.032>.
- Sepe-Forrest L, Carver FW, Quentin R, Holroyd T, Nugent AC (2021) Basal ganglia activation localized in MEG using a reward task. *Neuroimage: Reports* 1:100034 Available at: <https://doi.org/10.1016/j.ynirp.2021.100034>.
- Seymour B, Barbe M, Dayan P, Shiner T, Dolan R, Fink GR (2016) Deep brain stimulation of the subthalamic nucleus modulates sensitivity to decision outcome value in Parkinson's disease. *Sci Rep* 6:1–12 Available at: <http://dx.doi.org/10.1038/srep32509>.
- Seymour B, O'Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS (2004) ID 1627 letters to nature Temporal difference models describe higher-order learning in humans. *Nature* 429:664–667.
- Shadmehr R (2017) Learning to Predict and Control the Physics of Our Movements. *J Neurosci* 37:1663–1671 Available at: <http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.1675-16.2016>.
- Shadmehr R (2018) Motor Learning: A Cortical System for Adaptive Motor Control. *Curr Biol* 28:R793–R795 Available at: <https://doi.org/10.1016/j.cub.2018.05.071>.
- Shadmehr R, De Xivry JJO, Xu-Wilson M, Shih TY (2010a) Temporal discounting of reward and the cost of time in motor control. *J Neurosci* 30:10507–10516.
- Shadmehr R, Holcomb HH (1997) Neural correlates of motor memory consolidation. *Science* (80-) 277:821–825.
- Shadmehr R, Huang HJ, Ahmed AA (2016) A Representation of Effort in Decision-Making and Motor Control. *Curr Biol* 26:1929–1934.
- Shadmehr R, Reppert TR, Summerside EM, Yoon T, Ahmed AA (2019) Movement Vigor as a Reflection of Subjective Economic Utility. *Trends Neurosci* xx:1–14 Available at: <https://doi.org/10.1016/j.tins.2019.02.003>.
- Shadmehr R, Smith MA, Krakauer JW (2010b) Error Correction , Sensory Prediction , and Adaptation in Motor Control.
- Sharma DA, Chevidikunnan MF, Khan FR, Gaowgzeh RA (2016) Effectiveness of knowledge of result and knowledge of performance in the learning of a skilled motor activity by healthy young adults. *J Phys Ther Sci* 28:1482–1486.
- Shima K, Tanji J (1998) Role for cingulate motor area cells in voluntary movement selection based on reward. *Science* 282:1335–1338 Available at: <http://www.sciencemag.org/cgi/doi/10.1126/science.282.5392.1335><http://www.ncbi.nlm.nih.gov/pubmed/9812901>.
- Shirer WR, Ryali S, Rykhlevskaia E, Menon V, Greicius MD (2012) Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cereb Cortex* 22:158–165.
- Shmuelof L, Huang VS, Haith AM, Delnicki RJ, Mazzoni P, Krakauer JW (2012) Overcoming Motor “Forgetting” Through Reinforcement Of Learned Actions. *J Neurosci* 32:14617-14621a Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.2184-12.2012>.
- Shmuelof L, Yang J, Caffo B, Mazzoni P, Krakauer JW (2014) The neural correlates of learned motor acuity. :971–980.
- 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.
- Sidarta XA, Vugt XFT Van, Ostry DJ (2018) Somatosensory working memory in human reinforcement-based motor learning. :3275–3286.
- Singer AC, Frank LM (2009) Rewarded Outcomes Enhance Reactivation of Experience in the Hippocampus. *Neuron* 64:910–921

Available at: <http://dx.doi.org/10.1016/j.neuron.2009.11.016>.

- Smith MA, Ghazizadeh A, Shadmehr R (2006) Interacting Adaptive Processes with Different Timescales Underlie Short-Term Motor Learning. *PLoS Biol* 4:e179.
- Soderstrom NC, Bjork RA (2015) Learning Versus Performance: An Integrative Review. *Perspect Psychol Sci* 10:176–199.
- Song S, Zhang J, Tian Y, Wang L, Wei P (2021) Temporal Interference Stimulation Regulates Eye Movements and Neural Activity in the Mice Superior Colliculus. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS*:6231–6234.
- Spampinato D, Celnik P (2018) Deconstructing skill learning and its physiological mechanisms. *Cortex* 104:90–102 Available at: <https://doi.org/10.1016/j.cortex.2018.03.017>.
- Spampinato D, Celnik P (2021) Multiple Motor Learning Processes in Humans: Defining Their Neurophysiological Bases. *Neuroscientist* 27:246–267.
- Spampinato DA, Block HJ, Celnik PA (2017) Cerebellar–M1 Connectivity Changes Associated with Motor Learning Are Somatotopic Specific. *J Neurosci* 37:2377–2386 Available at: <http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.2511-16.2017>.
- Spampinato DA, Celnik PA, Rothwell JC (2020) Cerebellar-motor cortex connectivity: One or two different networks? *J Neurosci* 40:4230–4239.
- Spampinato DA, Satar Z, Rothwell JC (2019) Combining reward and M1 transcranial direct current stimulation enhances the retention of newly learnt sensorimotor mappings. *Brain Stimul*:1–8 Available at: <https://doi.org/10.1016/j.brs.2019.05.015>.
- Sporn S, Chen X, Galea JM (2020) Reward-based invigoration of sequential reaching. *bioRxiv*:1–38.
- Sporn S, Chen X, Galea JM (2022) The dissociable effects of reward on sequential motor behavior. *J Neurophysiol* 128:86–104.
- Starkweather CK, Babayan BM, Uchida N, Gershman SJ (2017) Dopamine reward prediction errors reflect hidden-state inference across time. *Nat Neurosci* 20:581–589.
- Steel A, Silson EH, Stagg CJ, Baker CI (2016) The impact of reward and punishment on skill learning depends on task demands. *Sci Rep* 6:1–9 Available at: <http://dx.doi.org/10.1038/srep36056>.
- Steel A, Silson EH, Stagg CJ, Baker CI (2019) Differential impact of reward and punishment on functional connectivity after skill learning. *Neuroimage* 189:95–105 Available at: <https://doi.org/10.1016/j.neuroimage.2019.01.009>.
- Steele CJ, Penhune VB (2010) Specific increases within global decreases: A functional magnetic resonance imaging investigation of five days of motor sequence learning. *J Neurosci* 30:8332–8341.
- Sterpenich V, van Schie MKM, Catsiannis M, Ramyea A, Perrig S, Yang HD, Van De Ville D, Schwartz S (2021) Reward biases spontaneous neural reactivation during sleep. *Nat Commun* 12 Available at: <http://dx.doi.org/10.1038/s41467-021-24357-5>.
- Strafella AP, Paus T, Fraraccio M, Dagher A (2003) Striatal dopamine release induced by repetitive transcranial magnetic stimulation of the human motor cortex. *Brain* 126:2609–2615.
- Sugawara SK, Tanaka S, Okazaki S, Watanabe K, Sadato N (2012) Social Rewards Enhance Offline Improvements in Motor Skill. *PLoS One* 7:3–8.
- Sutton RS, Barton AG (2018) Reinforcement learning: An introduction, 2nd ed. The MIT Press.
- Swain RA, Kerr AL, Thompson RF (2011) The Cerebellum: A Neural System for the Study of Reinforcement Learning. *Front Behav Neurosci* 5:1–6 Available at: <http://journal.frontiersin.org/article/10.3389/fnbeh.2011.00008/abstract>.
- Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S (2004) Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. 7:1–2.
- Taubert M, Lohmann G, Margulies DS, Villringer A, Ragert P (2011) Long-term effects of motor training on resting-state networks and underlying brain structure. *Neuroimage* 57:1492–1498 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2011.05.078>.
- Taylor JA, Ivry RB (2014) Cerebellar and Prefrontal Cortex Contributions to Adaptation, Strategies, and Reinforcement Learning, 1st ed. Elsevier B.V. Available at: <http://dx.doi.org/10.1016/B978-0-444-63356-9.00009-1>.

- Taylor JA, Krakauer JW, Ivry RB (2014) Explicit and implicit contributions to learning in a sensorimotor adaptation task. *J Neurosci* 34:3023–3032.
- Therrien AS, Statton MA, Bastian AJ (2020) Reinforcement Signaling Can Be Used to Reduce Elements of Cerebellar Reaching Ataxia. *Cerebellum*.
- Therrien AS, Wolpert DM, Bastian AJ (2016) Effective Reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. *Brain* 139:101–114.
- Therrien AS, Wolpert DM, Bastian AJ (2018) Increasing Motor Noise Impairs Reinforcement Learning in Healthy Individuals. *Eneuro* 5:ENEURO.0050-18.2018 Available at: <http://eneuro.sfn.org/lookup/doi/10.1523/ENEURO.0050-18.2018>.
- Therrien AS, Wong AL (2022) Mechanisms of Human Motor Learning Do Not Function Independently. *Front Hum Neurosci* 15:1–9.
- Thoma P, Bellebaum C, Koch B, Schwarz M, Daum I (2008) The cerebellum is involved in reward-based reversal learning. *Cerebellum* 7:433–443.
- Tobler P, O'Doherty J (2006) Human neural learning depends on reward prediction errors in the blocking paradigm. *J Neurophysiol*:301–310 Available at: <http://jn.physiology.org/content/95/1/301.short>.
- Torrubia R, Ávila C, Moltó J, Caseras X (2001) The Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. *Pers Individ Dif* 31:837–862 Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0191886900001835>.
- Trevartha KM, Garcia A, Wolpert DM, Flanagan JR (2014) Fast but fleeting: Adaptive motor learning processes associated with aging and cognitive decline. *J Neurosci* 34:13411–13421.
- Tseng Y -w., Diedrichsen J, Krakauer JW, Shadmehr R, Bastian AJ (2007) Sensory Prediction Errors Drive Cerebellum-Dependent Adaptation of Reaching. *J Neurophysiol* 98:54–62 Available at: <http://jn.physiology.org/cgi/doi/10.1152/jn.00266.2007>.
- Tumer EC, Brainard MS (2007) Performance variability enables adaptive plasticity of “crystallized” adult birdsong. *Nature* 450:1240–1244.
- Uehara S, Mawase F, Celnik P (2018) Learning similar actions by reinforcement or sensory-prediction errors rely on distinct physiological mechanisms. *Cereb Cortex* 28:3478–3490.
- Uehara S, Mawase F, Therrien AS, Cherry-allen KM, Celnik P, Hopkins J (2019) Interactions between motor exploration and reinforcement learning.
- van der Kooij K, van Mastrigt NM, Crowe EM, Smeets JBJ (2021) Learning a reach trajectory based on binary reward feedback. *Sci Rep* 11:1–15 Available at: <https://doi.org/10.1038/s41598-020-80155-x>.
- van der Meer MAA, Kalenscher T, Lansink CS, Pennartz CMA, Berke JD, Redish AD (2010) Integrating early results on ventral striatal gamma oscillations in the rat. *Front Neurosci* 4:1–12.
- van der Meer MAA, Redish AD (2009) Low and high gamma oscillations in rat ventral striatum have distinct relationships to behavior, reward, and spiking activity on a learned spatial decision task. *Front Integr Neurosci* 3:1–19.
- van Mastrigt NM, van der Kooij K, Smeets JBJ (2021) Pitfalls in quantifying exploration in reward-based motor learning and how to avoid them. *Biol Cybern* 115:365–382 Available at: <https://doi.org/10.1007/s00422-021-00884-8>.
- Vassiliadis P, Derosiere G (2020) Selecting and executing actions for rewards. *J Neurosci* 40:6474–6476.
- Vassiliadis P, Derosiere G, Dubuc C, Lete A, Crevecoeur F, Hummel FC, Duque J (2021) Reward boosts reinforcement-based motor learning. *iScience* 24:102821 Available at: <https://doi.org/10.1016/j.isci.2021.102821>.
- Vassiliadis P, Derosiere G, Duque J (2019) Beyond Motor Noise : Considering Other Causes of Impaired Reinforcement Learning in Cerebellar Patients. *Eneuro* 6:1–4.
- Vassiliadis P, Derosiere G, Grandjean J, Duque J (2020a) Motor training strengthens corticospinal suppression during movement preparation. *J Neurophysiol* 124:1656–1666.
- Vassiliadis P, Derosiere G, Grandjean J, Duque J (2020b) Motor training strengthens corticospinal suppression during movement preparation. *bioRxiv*:2020.02.14.948877.

- Vassiliadis P, Grandjean J, Derosiere G, Wilde Y De (2018) Using a Double-Coil TMS Protocol to Assess Preparatory Inhibition Bilaterally. *Front Neurosci* 12:1–14.
- Vassiliadis P, Lete A, Duque J, Derosiere G (2022) Reward timing matters in motor learning. *iScience*.
- Verbruggen F, Best M, Bowditch WA, Stevens T, McLaren IPL (2014) The inhibitory control reflex. *Neuropsychologia* 65:263–278.
- Violante IR, Alania K, Cassarà AM, Neufeld E, Acerbo E, Carron R, Williamson A, Kurtin DL, Hampshire A, Kuster N, Boyden ES, Pascual-Leone A, Grossman N, Marcus A (2022) Non-invasive temporal interference electrical stimulation of the human hippocampus.
- Voloh B, Oemisch M, Womelsdorf T (2020) Phase of firing coding of learning variables across the fronto-striatal network during feature-based learning. *Nat Commun* 11:1–16 Available at: <http://dx.doi.org/10.1038/s41467-020-18435-3>.
- von Conta J, Kasten FH, Ćurčić-Blake B, Aleman A, Thielscher A, Herrmann CS (2021) Interindividual variability of electric fields during transcranial temporal interference stimulation (tTIS). *Sci Rep* 11:1–12 Available at: <https://doi.org/10.1038/s41598-021-99749-0>.
- Vyas S, Even-Chen N, Stavisky SD, Ryu SI, Nuyujukian P, Shenoy K V. (2018) Neural Population Dynamics Underlying Motor Learning Transfer. *Neuron* 97:1177–1186.e3 Available at: <https://doi.org/10.1016/j.neuron.2018.01.040>.
- Vyas S, O'Shea DJ, Ryu SI, Shenoy K (2020) Causal role of motor preparation during error-driven learning. *Neuron*:1–11 Available at: <https://doi.org/10.1016/j.neuron.2020.01.019>.
- Wachter T, Lungu O V., Liu T, Willingham DT, Ashe J (2009) Differential Effect of Reward and Punishment on Procedural Learning. *J Neurosci* 29:436–443 Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.4132-08.2009>.
- Wagner MJ, Kim TH, Savall J, Schnitzer MJ, Luo L (2017) Cerebellar granule cells encode the expectation of reward. *Nature* 544:96–100 Available at: <http://dx.doi.org/10.1038/nature21726>.
- Wang Q, Chen C, Cai Y, Li S, Zhao X, Zheng L, Zhang H, Liu J, Chen C, Xue G (2016) Dissociated neural substrates underlying impulsive choice and impulsive action. *Neuroimage* 134:540–549 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2016.04.010>.
- Waters S, Wiestler T, Diedrichsen J (2017) Cooperation not competition: Bihemispheric tDCS and fMRI show role for ipsilateral hemisphere in motor learning. *J Neurosci* 37:7500–7512.
- Weismüller B, Ghio M, Logmin K, Hartmann C, Schnitzler A, Pollok B, Südmeyer M, Bellebaum C (2018) Effects of feedback delay on learning from positive and negative feedback in patients with Parkinson's disease off medication. *Neuropsychologia* 117:46–54 Available at: <https://doi.org/10.1016/j.neuropsychologia.2018.05.010>.
- Wenger E, Kühn S, Verrel J, Mårtensson J, Bodammer NC, Lindenberg U, Lövdén M (2017) Repeated Structural Imaging Reveals Nonlinear Progression of Experience-Dependent Volume Changes in Human Motor Cortex. *Cereb Cortex* 27:2911–2925.
- Wessel MJ, Beanato E, Popa T, Windel F, Menoud P, Beliaeva V, Vassiliadis PT, Grossman N, Neufeld E, Hummel FC (2021) Evidence for temporal interference (TI) stimulation effects on motor striatum. *Brain Stimul* 14:1684 Available at: <https://linkinghub.elsevier.com/retrieve/pii/S1935861X21005532>.
- Wessel MJ, Draaisma LR, de Boer AFW, Park C hyun, Maceira-Elvira P, Durand-Ruel M, Koch PJ, Morishita T, Hummel FC (2020) Cerebellar transcranial alternating current stimulation in the gamma range applied during the acquisition of a novel motor skill. *Sci Rep* 10:1–12 Available at: <https://doi.org/10.1038/s41598-020-68028-9>.
- Whitlock JR, Heynen AJ, Shuler MG, Bear MF (2006) Learning induces long-term potentiation in the hippocampus. *Science* (80-) 313:1093–1097.
- Widmer M, Held JPO, Wittmann F, Valladares B, Lamercy O, Sturzenegger C, Palla A, Lutz K, Luft AR (2021) Reward During Arm Training Improves Impairment and Activity After Stroke: A Randomized Controlled Trial. *Neurorehabil Neural Repair* 0:154596832110628.
- Widmer M, Lutz K, Luft AR (2019) Reduced striatal activation in response to rewarding motor performance feedback after stroke. *NeuroImage Clin* 24:102036 Available at: <https://doi.org/10.1016/j.nicl.2019.102036>.
- Widmer M, Ziegler N, Held J, Luft A, Lutz K (2016) Rewarding feedback promotes motor skill consolidation via striatal activity. In: *Progress in Brain Research*, pp 303–323 Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0079612316300590>.

- Wiestler T, Diedrichsen J (2013) Skill learning strengthens cortical representations of motor sequences. *Elife* 2013:1–20.
- Wilkinson L, Steel A, Mooshagian E, Zimmermann T, Keisler A, Lewis JD, Wassermann EM (2015) Online feedback enhances early consolidation of motor sequence learning and reverses recall deficit from transcranial stimulation of motor cortex. *Cortex* 71:134–147 Available at: <http://dx.doi.org/10.1016/j.cortex.2015.06.012>.
- Williams ZM, Eskandar EN (2006) Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nat Neurosci* 9:562–568.
- Wolff SBE, Ko R, Ölteczky BP (2022) Distinct roles for motor cortical and thalamic inputs to striatum during motor skill learning and execution. *Sci Adv* 8:1–15.
- Wolpert DM, Diedrichsen J, Flanagan JR (2011) Principles of sensorimotor learning. *Nat Rev Neurosci* 12 Available at: <http://dx.doi.org/10.1038/nrn3112>.
- Wolpert DM, Ghahramani Z, Jordan M (1995) An internal model for sensorimotor integration - Wolpert et al. (1995).pdf. *Science* (80-) 269:1880–1882 Available at: <http://dx.doi.org/10.1038/nphys2178>.
- Wu HG, Miyamoto YR, Castro LNG, Ölteczky BP, Smith MA (2014) Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nat Neurosci* 17:312–321 Available at: <http://dx.doi.org/10.1038/nn.3616>.
- Wu T, Kansaku K, Hallett M (2004) How Self-Initiated Memorized Movements Become Automatic: A Functional MRI Study. *J Neurophysiol* 91:1690–1698.
- Yamazaki Y, Hikishima K, Saiki M, Inada M, Sasaki E, Lemon RN, Price CJ, Okano H, Iriki A (2016) Neural changes in the primate brain correlated with the evolution of complex motor skills. *Nat Publ Gr*:1–10 Available at: <http://dx.doi.org/10.1038/srep31084>.
- Yang G, Lai CSW, Cichon J, Ma L, Li W, Gan WB (2014) Sleep promotes branch-specific formation of dendritic spines after learning. *Science* (80-) 344:1173–1178.
- Yokoi A, Diedrichsen J (2019) Neural Organization of Hierarchical Motor Sequence Representations in the Human Neocortex. *Neuron* 103:1178–1190.e7.
- Yoon T, Jaleel A, Ahmed AA, Shadmehr R (2019) Saccade vigor as an implicit measure of subjective economic value. :1–15.
- Yoshimoto A, Shibata Y, Kudara M, Ikegaya Y, Matsumoto N (2022) Enhancement of Motor Cortical Gamma Oscillations and Sniffing Activity by Medial Forebrain Bundle Stimulation Precedes Locomotion. *eNeuro* 9:1–14.
- Zhu Z, Xiong Y, Chen Y, Jiang Y, Qian Z, Lu J (2022) Temporal Interference (TI) Stimulation Boosts Functional Connectivity in Human Motor Cortex : A Comparison Study with Transcranial Direct Current Stimulation (tDCS). 2022:1–7.

Annex 1: Motor training strengthens corticospinal suppression during movement preparation

Pierre Vassiliadis^{1,2*}, Gerard Derosiere¹, Julien Grandjean¹, Julie Duque¹

¹ Institute of Neuroscience, Université Catholique de Louvain, 1200, Brussels, Belgium.

² Defitech Chair for Clinical Neuroengineering, Center for Neuroprosthetics (CNP) and Brain Mind Institute (BMI), Swiss Federal Institute of Technology (EPFL), 1202, Geneva, Switzerland

Reference: Vassiliadis P, Derosiere G, Grandjean J, Duque J (2020). Motor training strengthens corticospinal suppression during movement preparation. Journal of Neurophysiology, 124: 1656–1666.

Personal contribution: Conceptualization, Methodology, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Funding acquisition.

* Correspondence

Pierre Vassiliadis

Institute of Neuroscience

Université catholique de Louvain

53, Avenue Mounier

COSY- B1.53.04

B-1200 BRUSSELS

Belgium

Email: pierre.vassiliadis@uclouvain.be

Web: <http://www.coactionslab.com>

Abstract

Training can improve motor skills and modify neural activity at rest and during movement execution. Learning-related modulations may also concern motor preparation but the neural correlates and the potential behavioral relevance of such adjustments remain unclear. In humans, preparatory processes have been largely investigated using transcranial magnetic stimulation (TMS) with several studies reporting decreased corticospinal excitability (CSE) relative to a baseline measure at rest; a phenomenon called preparatory suppression. Here, we investigated the effect of motor training on such preparatory suppression, in relation to resting CSE, in humans. We trained participants to initiate quick movements in an instructed-delay reaction time (RT) task and used TMS to investigate changes in CSE over the practice blocks. Training on the task speeded up RTs, with no repercussion on error rates. Training also increased resting CSE. Most interestingly, we found that CSE during action preparation did not mirror the training-related increase observed at rest. Rather, compared to the rising baseline, the degree of preparatory suppression strengthened with practice. This training-related change in preparatory suppression (but not the changes in baseline CSE) predicted RT gains: the subjects showing a greater strengthening of preparatory suppression were also those exhibiting larger gains in RTs. Finally, such relationship between RTs and preparatory suppression was also evident at the single-trial level, though only in the non-selected effector: RTs were generally faster in trials where preparatory suppression was deeper. These findings suggest training induces changes in motor preparatory processes that are linked to an enhanced ability to initiate fast movements.

New and Noteworthy

Movement preparation involves a broad suppression in the excitability of the corticospinal pathway, a phenomenon called preparatory suppression. Here, we show that motor training strengthens preparatory suppression and that this strengthening is associated with faster reaction times. Our findings highlight a key role of preparatory suppression in training-driven behavioral improvements.

Keywords: motor training; action preparation; transcranial magnetic stimulation; motor-evoked potentials; primary motor cortex; corticospinal excitability; corticospinal suppression, reaction time.

1. Introduction

Motor training improves the speed and/or accuracy at which movements are selected, initiated and executed. Significant research has been devoted to unveiling the functional changes at the basis of such improvements (Krakauer et al., 2019). At the neural level, neuroimaging (e.g., Wiestler & Diedrichsen, 2013; Wenger et al., 2017; Yokoi & Diedrichsen, 2019) and transcranial magnetic stimulation (TMS) studies (e.g., Rosenkranz et al., 2007; Reis et al., 2008; Mawase et al., 2017) have shown that training is accompanied by a plastic reorganization of the motor system, supporting the formation of new motor memories. Specifically, training amplifies resting motor activity (e.g., Pascual-Leone et al., 1995; Butefisch et al., 2000; Duque et al., 2008; Galea & Celnik, 2009; Christiansen et al., 2018) and induces learning-specific changes of motor activity during movement execution (Krakauer et al., 2004; Steele and Penhune, 2010; Shmuelof et al., 2014). Animal studies also show learning-related modulations of motor activity during action preparation (Paz et al., 2003; Makino et al., 2017; Vyas et al., 2018, 2020) that could reflect an optimization of preparatory processes with training (Mawase et al., 2018). Yet, the behavioral relevance of the effects of training on action preparation remain unclear.

In humans, the excitability of the motor system can be assessed by applying TMS over primary motor cortex (M1), eliciting motor-evoked potentials (MEPs), whose amplitude reflects the excitability of the corticospinal pathway (Derosiere and Duque, 2020; Derosiere et al., 2020). When applied during reaction time (RT) tasks, TMS elicits MEPs that are used to assess corticospinal excitability (CSE) changes associated with action preparation and initiation. CSE is often suppressed during action preparation when compared to a baseline, measured at rest. The function of this preparatory suppression (or inhibition) remains unclear (e.g., Greenhouse et al., 2015; Duque et al., 2017; Derosiere, 2018; Hannah et al., 2018). A prominent view is that it assists action selection processes, by preventing the release of premature or incorrect responses (Duque et al., 2010; Quoilin et al., 2018, 2020; Grandjean and Duque, 2020). Indeed, the amount of preparatory suppression seems to scale with the complexity of selection processes (Klein et al., 2014; Duque et al., 2016). Another hypothesis is that preparatory suppression eases movement initiation (Greenhouse et al., 2015; Hasegawa et al., 2017). In this line, a study showed a dependence of RTs on the amount of preparatory suppression on a single-trial basis: the stronger the suppression,

the faster the initiation of the ensuing movement (Hannah et al., 2018). Importantly, both hypotheses could be valid as they focus on different levels of control: while the “selection hypothesis” suggests that suppression originates from processes that help select accurate actions (*i.e.*, therefore reducing the error rate), the “motor hypothesis” entails that preparatory suppression is also generated by processes speeding up movement initiation (*i.e.*, therefore reducing RTs).

Here, we investigated the impact of motor training on preparatory suppression, while subjects practiced an instructed-delay RT task. The selection aspects were clear-cut, as evident from the low error rates, even before training. Hence, in such task, there is no room for improvement at the selection level and subjects can only become more skilled at the motor level, by initiating their response faster. Based on this, we predicted that RTs would shorten over the course of practice but that error rates would remain marginal. In addition, we expected resting CSE to increase with training, in accordance with previous work (Pascual-Leone et al., 1995; Butefisch et al., 2000; Duque et al., 2008; Galea and Celnik, 2009; Christiansen et al., 2018). Based on the motor hypothesis (*i.e.*, that preparatory suppression fastens RTs), we expected that the suppression of CSE during action preparation would strengthen along with the RT decrease over the practice blocks.

2. Materials and Methods

2.1. Participants

Fifteen right-handed healthy subjects participated in the present study ($n=15$; 10 women; 22.4 ± 1.63 years old). Handedness was assessed via Edinburgh Handedness inventory (Oldfield, 1971). Participants filled out a TMS safety questionnaire to look for any contra-indication and gave written informed consent in accordance with the Ethics Committee of the Université Catholique de Louvain (approval number: 2012/22MAR/119) and the principles of the Declaration of Helsinki. We had to exclude one subject because we encountered a technical problem during the experiment; hence, analyses were run on the fourteen remaining subjects. Part of the data reported here has been exploited in a separate study (Vassiliadis et al., 2018). All of the data are expressed as $\text{mean}\pm\text{SE}$.

2.2. Task

Subjects were seated in front of a computer screen with the hands on response devices (**Figure 1A**, Grandjean et al. 2019; Grandjean and Duque 2020; Quoilin et al. 2016, 2018, 2019, 2020). They performed an instructed-delay RT task, which required them to choose between abduction movements of the left or right index finger according to the position of a preparatory cue (*i.e.*, a left- or right-sided ball separated from a goal by a gap). Participants had to prepare their movement once the ball appeared but to withhold responding until the onset of a “Go” signal (*i.e.*, a bridge). When the bridge appeared on the screen, subjects had to respond as fast as possible, allowing the ball to roll on the bridge and to reach the goal (**Figure 1B**). To reduce anticipation of the “Go” signal, the bridge did not appear in some of the trials (5%). Subjects were required not to respond on these trials and were penalized if they did so.

Trials always ended with a feedback score reflecting performance. On correct trials, scores ranged from 1 to 100 points and were displayed in green. Participants were informed that the score was inversely proportional to the RT: the faster the response, the higher the score. In order to homogenize the score across subjects, scores on correct trials were individualized according to RTs measured during a familiarization block just before the main experiment (Vassiliadis et al., 2018; Grandjean et al., 2019). Incorrect responses were penalized with negative scores displayed in red. They involved responses occurring too early ($\text{RT}<100$ ms), referred to as “anticipation errors” (-75 points), responses occurring too late

(RT>500 ms), referred to as “time-out errors” (-50 points), responses provided with the incorrect hand (-20 points), referred to as “selection errors” and responses provided on catch trials (-12 points), referred as “catch errors”. When subjects succeeded not to respond on a catch trial, they were rewarded by +12 points. The total score was displayed at the end of each block.

2.3 TMS Protocol

Monophasic pulses were delivered through one or two small figure-of-eight shaped coils (internal diameter: 3.5 cm), each connected either to a Magstim 200² magnetic stimulator (Magstim, Whitland, Dyfed, UK). The TMS machine used to stimulate each hemisphere was counterbalanced between subjects. Pulses could be triggered in one (*i.e.*, single-coil TMS) or in the two coils (*i.e.*, double-coil TMS, **Figure 1C**) because the dataset was initially acquired for a separate study to establish the reliability of double-coil TMS to probe CSE bilaterally (Grandjean et al., 2018; Vassiliadis et al., 2018). In double-coil trials, a 1-ms interval separated the onset of the two pulses, eliciting MEPs in both hands at a near simultaneous time (Algoet et al., 2018; Grandjean et al., 2018; Vassiliadis et al., 2018; Quoilin et al., 2019). This interval was used to avoid direct electromagnetic interference between the two coils (Cincotta et al., 2005), while preventing transcallosal interactions that would occur between motor areas with longer delays (Ferber et al., 1992; Hanajima et al., 2001). Notably, in double-coil trials, half of the trials involved a pulse over left M1 first whereas the other half involved a pulse over right M1 first (1ms delay). These data were assembled because a prior analysis reported elsewhere showed that the order of stimulation does not influence double-coil MEP amplitudes, which are equivalent to single-coil MEPs (Vassiliadis et al., 2018).

Each TMS coil was placed tangentially over one M1 with the handle pointing backward, laterally at a 45° angle away from the midline to induce a posterior-anterior current in the underlying cortical tissue (**Figure 1C**). TMS was applied over the hotspot of the first dorsal interosseous muscle (FDI), which was the prime-mover in our task (Duque et al., 2014; Derosiere et al., 2017a, 2017b). The resting Motor Threshold (rMT) was determined for each M1. It was defined as the minimal intensity required to evoke MEPs of 50µV at rest in at least 5 out of 10 stimulations (Groppa et al., 2012; Wessel et al., 2020). The rMTs equalled 41.7±5.05% and 40.8±6.39% of the maximum stimulator output for the left and the

right FDI, respectively. For each hemisphere, the intensity used throughout the experiment was set at 115% of the individual rMT (Derosiere et al., 2019).

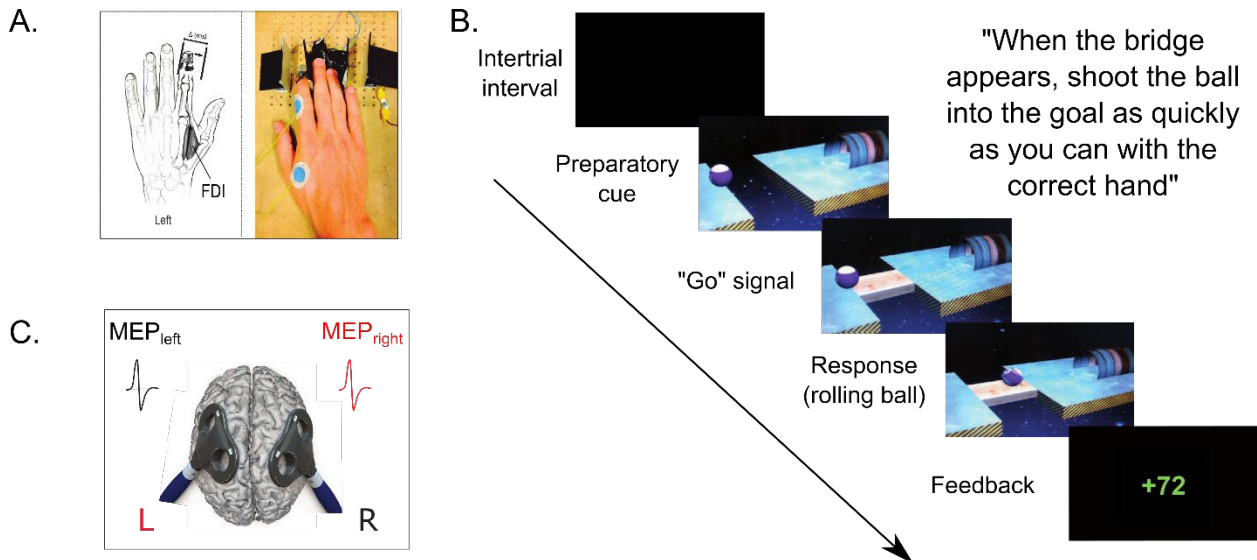


Figure 1. A, The response device. Index finger responses were recorded using a home-made device composed of two pairs of metal edges fixed on a wooden platform and positioned under the left (graphic representation) and right (photographic representation) hands. **B, "Rolling Ball" task.** Subjects were asked to choose between responding with the left or right index finger according to the position of a ball (Preparatory cue) appearing on the left or right part of the screen (left in the current example). They had to wait until the onset of a bridge ("Go" signal) to release their response as quickly as possible. The ball then rolled on the bridge (when the subjects answered correctly) to reach a goal located on the other side of the gap. A feedback reflecting how fast and accurate the subjects were concluded each trial. **C, TMS protocol.** Two small figure-eight-shaped coils were placed over the subject's M1, eliciting MEPs in the left and/or right FDI.

2.4 Experimental procedure

The experiment started with two familiarization blocks. The first block included 20 trials and allowed subjects to become acquainted with the task. The second block involved 40 trials with TMS (as in the experimental blocks) and served to compute the median RT for each subject. The latter was used to individualize the feedback scores on correct trials according to the initial performance.

Then, subjects performed 400 trials of the task, divided in 10 experimental blocks. Each block involved an equal combination of single- and double-coil stimulations, occurring in a random order (*i.e.*, subjects could not anticipate the type of stimulation they would face). Given that both techniques produce equivalent MEPs (Grandjean et al., 2018; Vassiliadis et al., 2018), these data were considered regardless of the protocol used to elicit them.

TMS could occur in three different settings. First, some TMS pulses were delivered outside the blocks (TMS_{baseline-out}), providing MEPs reflecting baseline CSE at complete rest. TMS_{baseline-out} pulses occurred every other block, starting before block 1 and ending after block 8 (5 time points; **Figure 2A**). Each time point involved 10 double-coil pulses (i.e. 10 pulses on both M1) and 10 single-coil pulses (5 pulses on each M1), allowing us to collect 15 MEPs in each hand at each of the 5 different stages of the experiment. Second, TMS occurred during the intertrial interval, 300 ms before the beginning of the trial (**Figure 2B**). MEPs recorded at this time provided another baseline measure of CSE, with subjects at rest but engaged in the task (TMS_{baseline-in} Labruna et al., 2011). In each block, TMS_{baseline-in} occurred in 4 double-coil trials (i.e. 4 pulses on both M1) and 4 single-coil trials (2 pulses on each M1), allowing us to collect a total of 6 baseline MEPs per block in each hand (i.e., 12 for each Training_{STAGE}). Finally, TMS occurred at 900 or 950 ms after the onset of the preparatory cue (TMS_{preparation}). Since no difference was found between MEPs recorded at these two timings in our previous analysis of the same data set (Vassiliadis et al., 2018), these MEPs were pooled together. TMS_{preparation} occurred in 16 double-coil trials (i.e., 16 pulses on both M1) and 16 single-coil trials (8 pulses on each M1), allowing us to collect a total of 24 preparatory MEPs per block in each hand (48 MEPs for each Training_{STAGE}). Half of these MEPs fell in left response trials, while the other half occurred in right response trials. Hence, MEPs could either fall in a hand that was selected for the forthcoming response (MEP_{selected}; e.g., left MEPs preceding a left index finger response) or in a hand that was non-selected (MEP_{non-selected}).

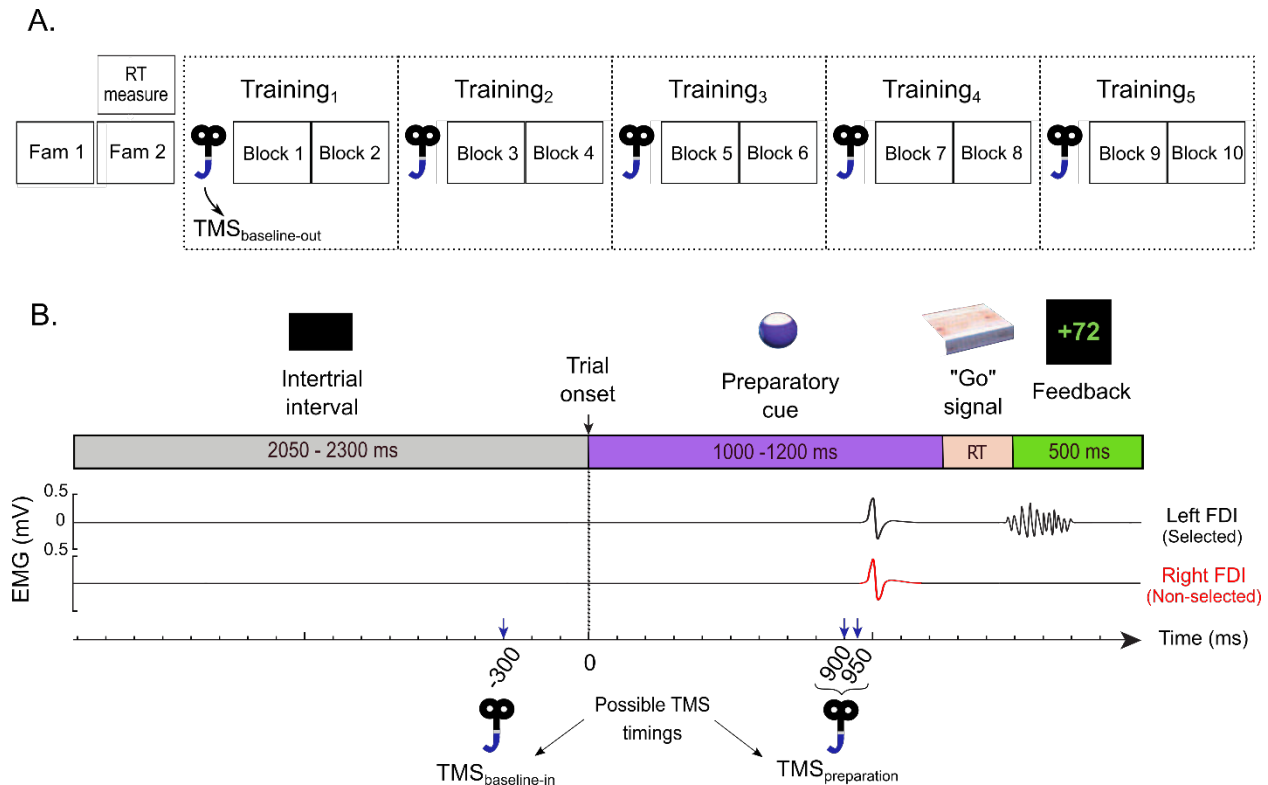


Figure 2. A, Time-course of the experiment. After two familiarization blocks, subjects executed ten blocks of forty trials during which MEPs were elicited at $TMS_{baseline-in}$ or $TMS_{preparation}$. The effect of training was assessed by comparing five sets of data (Training₁ to Training₅), each involving MEPs pooled over two consecutive blocks. MEPs were also elicited outside the blocks ($TMS_{baseline-out}$) at five points in time, before block 1 and after blocks 2, 4, 6 and 8, categorized as Training₁ to Training₅, similar to the MEPs elicited during the blocks. Comparing these data sets allowed us to consider potential training-related changes in resting CSE outside the context of the task. **B, Time course of a trial.** Trials were separated by a blank screen (intertrial interval; 2050 to 2300 ms) and always started with a preparatory cue appearing for a variable delay period (1000 to 1200 ms). Variable delays were sampled from uniform distributions to induce temporal uncertainty and therefore reduce anticipation of the pulses that could emerge with the repetition of trials. Then, a “Go” signal was presented and remained on the screen until a response was detected, hence for the duration of the reaction time (RT). The feedback was presented at the end of each trial for 500 ms and depended on the RT on correct trials. TMS pulses occurred either during the intertrial interval (300 ms before the beginning of the trial; $TMS_{baseline-in}$), or during the delay period (900 or 950 ms after the preparatory cue onset, timings pooled; $TMS_{preparation}$). In double-coil trials, motor-evoked potentials (MEPs) were elicited in the first dorsal interosseous (FDI) of both hands at a near simultaneous time (1 ms delay); in single-coil trials, MEPs were elicited in the left or right hand. The figure displays a left hand trial with double-coil TMS at $TMS_{preparation}$.

2.5 Data processing and statistical analyses

The purpose of the study was twofold: (1) to characterize changes in CSE at rest and during action preparation occurring along with training in a basic instructed-delay RT task, (2) to assess whether modulations in CSE were correlated to training-related improvements in RTs. To do so, the behavioral and MEP data were evaluated according to the block within

which they were elicited and data from two consecutive blocks were pooled together. Given the 10 blocks, we obtained five data sets reflecting five training stages (Training_{STAGE}: Training₁ to Training₅; **Figure 2A**).

Statistical analyses were carried out with Matlab 2018a (the Mathworks, Natick, Massachusetts, USA) and Statistica 10 (StatSoft Inc., Tulsa, Oklahoma, USA). All data were systematically tested for the sphericity assumption using Maunchley's tests. The Greenhouse–Geisser (GG) correction was used for sphericity when necessary. Post-hocs comparisons were always conducted using the Fisher's LSD procedure. The significance level was set at $p \leq 0.05$.

2.5.1 RTs and errors

Left and right hand RTs were computed as the difference between the onset of the “Go” signal and movement onset (when the finger quitted the outer metal edge of the device). Trials where subjects made an error were removed from the data set for the RT analysis. An average of 34.9 left and 33.6 right response trials remained for each subject at each Training_{STAGE} (Table 1). We computed the mean RT for left and right responses separately and then averaged these data together. We choose this two-step method to make sure that left and right hand RTs would have the same weight in the averaged data for each Training_{STAGE}, regardless of the number of errors in each hand. Besides, we also assessed response accuracy over training, by computing, for each Training_{STAGE}, the amount of anticipation, time-out and catch errors as well as the total error rate. For each of these variables, we expressed the number of incorrect trials in percentage of the total amount of trials, regardless of the responding hand. Selection errors were not analysed because they were rare (4 selection errors across all subjects). For the statistical analysis of RTs and errors (*i.e.*, anticipation, time-out, catch and global errors), we used one-way analyses of variance for repeated measures (ANOVA_{RM}) with the factor Training_{STAGE} (Training₁ to Training₅).

2.5.2 MEP amplitudes

MEPs were obtained by recording electromyography (EMG) bilaterally from surface electrodes (Neuroline, Medicotest, Oelstykke, Denmark) placed over the FDI. The signals were amplified (x1000), bandpass filtered (10-500Hz; NeuroloLog; Digitimer), digitalized at

2000 Hz and collected with Signal (Signal 3.0, Cambridge, UK) for offline analysis. Trials with background EMG activity in the 200 ms window preceding the pulse exceeding 3 SDs above the mean were discarded ($1.68 \pm 0.30\%$ removal; Vassiliadis et al., 2018; Grandjean et al., 2018, 2019). This was done to prevent contamination of the MEP measurements by significant fluctuations in background EMG.

To assess training-related changes in resting CSE based on MEPs elicited at TMS_{baseline-out} and TMS_{baseline-in}, we averaged separately left and right hand MEPs for each Training_{STAGE} before computing the mean of these averages. These data were analysed using a two-way ANOVA_{RM} with TMS_{TIMING} (TMS_{baseline-out} or TMS_{baseline-in}), and Training_{STAGE} (Training₁ to Training₅) as within-subject factors. To assess training-related changes in preparatory suppression based on MEPs at TMS_{preparation} (expressed in percentage of MEPs at TMS_{baseline-in}), we first removed the trials in which subjects made a mistake ($10.78 \pm 1.81\%$ removal) and then grouped left and right hand MEPs according to whether they corresponded to a MEP_{selected} or MEP_{non-selected}. Within these categories, we averaged the separate means of left and right hand MEPs for each Training_{STAGE}. The number of trials that remained in each condition after data cleaning is provided in Table 1. To analyse these data, we first focused on percentage MEPs at Training₁, assessing with t-tests (against a constant value of 100%) the significance of preparatory suppression at the beginning of training. Then, we analyzed all training stages using a two-way ANOVA_{RM} with the factors MEP_{SELECTION} (MEP_{selected} or MEP_{non-selected}) and Training_{STAGE} (Training₁, to Training₅). This ANOVA was also run on absolute MEP amplitudes (in mV).

	Selected	Non-Selected	Selected	Non-Selected
MEP at TMS _{preparation}				

Table 1: Number of trials included per condition in the main analysis (mean [range]).

2.5.3 Relationship between training-related changes in RTs and CSE

As described in the Result section, training influenced RTs and CSE. We studied the relationship between changes at these two levels, with CSE considered separately at rest and during action preparation. We computed ratios reflecting training-related changes. Based on the RT data, we realized that the subjects' behavior improved substantially during the first practice stage (Training₁ to Training₃) but then, RTs remained quite stable (from Training₃ to Training₅; Result section). For this reason, we considered ratios for these two phases of training separately, providing us with an indication of early (Training_{ratio-early}: Training₃/Training₁) and late (Training_{ratio-late}: Training₅/Training₃) training-related changes in RTs and CSE. For the latter, we computed separate ratios for MEPs at TMS_{baseline-out}, TMS_{baseline-in} and TMS_{preparation} (expressed in percentage of MEPs at TMS_{baseline-in}). We then examined the correlation between the RT and MEP Training_{ratios} by using least squares linear regressions.

Finally, we compared the strength of the RT relationship to training-related changes in MEP amplitudes at TMS_{baseline-in} (reflecting resting CSE) and changes in percentage MEPs at TMS_{preparation} (reflecting preparatory suppression of CSE). To do so, in order to obtain a robust estimate of the absolute Pearson's R, we ran a bootstrap analysis with 10000 resamples and calculated a median R for each correlation (Efron, 1979). These R-values were then compared to each other using Pearson and Fillon's z test (Pearson and Filon, 1898).

2.5.4 Single-trial relationship between RTs and preparatory suppression

The correlation analyses revealed a relationship between RTs and preparatory suppression: the subjects who showed the greatest training-related reduction in RTs were also those who displayed the strongest deepening in preparatory suppression (see Result section). To better understand the dependency of RTs to the strength of preparatory suppression, we investigated whether this relationship was evident on a single-trial basis, as suggested previously (Hannah et al., 2018). We selected the MEPs elicited at TMS_{preparation} and again, expressed them as a percentage of TMS_{baseline-in}. We only used the double-coil trials,

to consider a homogeneous set of data, with preparatory MEPs falling in both hands systematically. For each trial, we extracted the RT, as well as the MEPs recorded at TMS_{preparation} in both selected (MEP_{selected}) and non-selected (MEP_{non-selected}) hands. Hence, for each trial, we obtained one RT measure linked to two different MEPs.

To examine the relationship between RTs and preparatory suppression, we pooled the trials from all 10 blocks together and sorted them according to the amplitude of MEPs within each trial. Given that there were two MEPs in each trial, we repeated this procedure twice, providing us with two different orderings of the trials according to the MEP_{selected} or MEP_{non-selected}. Within each arrangement, trials were grouped into 6 consecutive percentile bins (MEP_{BIN}: MEP_{BIN-1} = 0 to 16.7%, MEP_{BIN-2} = 16.7 to 33.3% ... MEP_{BIN-6} = 83.3 to 100% of the data). MEP_{BIN-1} contained the trials with the stronger preparatory suppression whereas the MEP_{BIN-6} included the trials with the weaker preparatory suppression. We then computed the mean RT of trials within each MEP_{BIN} (23.5 trials per condition on average and never less than 19 trials). Notably, because each MEP_{BIN} involved a limited number of trials in this analysis, RTs were pooled together regardless of whether they were obtained in a left or right hand trial. Hence, we obtained six average RT values (*i.e.*, one for each MEP_{BIN}) for each of the trial arrangements based on the two MEP types. These two sets of RT data were analysed using two separate ANOVA_{RM} with the factor MEP_{BIN} (MEP_{BIN-1} to MEP_{BIN-6}).

3. Results

3.1. RTs and errors

Figure 3A shows the evolution of RTs with training. The ANOVA_{RM} revealed a significant influence of Training_{STAGE} on RT ($F_{(4,52)}=4.31$, $p=0.0043$). Post-hoc tests showed that RTs measured from Training₃ to Training₅ were shorter than at Training₁ (all $p<0.004$). In contrast, the total error rate remained stable over the blocks ($F_{(4,52)}=0.82$, $p=0.52$, **Figure 3B**). We did not observe any modification of the percentage of anticipation ($F_{(4,52)}=1.12$, $p=0.36$), time-out (GG-corrected $F_{(2.50,32.50)}=0.90$, $p=0.44$) or catch errors ($F_{(4,52)}=1.73$, $p=0.16$). Hence, training enabled subjects to respond more quickly while maintaining the same accuracy level.

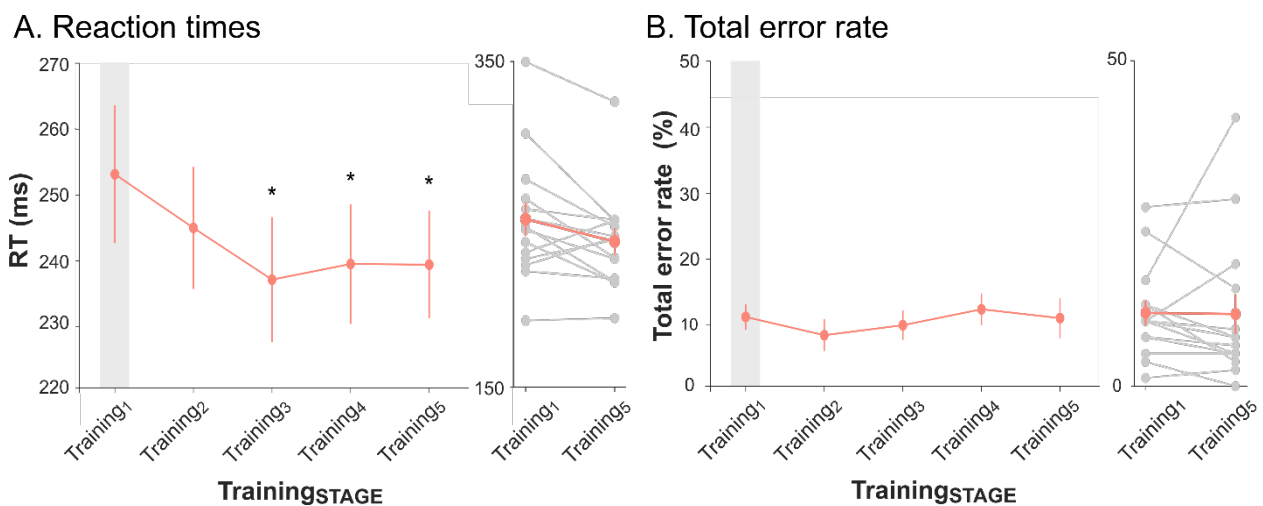


Figure 3. Evolution of reaction times (RTs) and total error rate throughout training. The mean RTs (**A**, in ms) and total error rate (**B**, in % of all trials) are represented for each Training_{STAGE}, regardless of the responding hand. Stars denote a significant difference between a given Training_{STAGE} and Training₁ ($p<0.05$). Individual data for Training₁ and Training₅.

3.2. MEP amplitude

First, we evaluated the effect of training on MEPs acquired at rest. As evident on **Figure 4**, MEPs were larger when assessed in the context of the task (TMS_{baseline-in}: 1.79 ± 0.17 mV) compared to when subjects were fully at rest (TMS_{baseline-out}: 1.34 ± 0.17 mV), as supported by the significant factor TMS_{TIMING} ($F_{(1,13)}=28.43$, $p<0.001$) and consistent with previous studies (Labruna et al., 2011; Derosière et al., 2015). The ANOVA_{RM} also revealed

an effect of Training_{STAGE} on baseline MEPs ($F_{(4,52)}=6.34$, $p<0.001$). MEPs recorded at Training₂ to Training₅ were larger than at Training₁ (all $p<0.03$). This training effect on MEPs occurred independently of the TMS_{TIMING}: there was a parallel increase in the amplitude of MEPs elicited at TMS_{baseline-out} and TMS_{baseline-in} (Training_{STAGE}×TMS_{TIMING}: $F_{(4,52)}=0.18$, $p=0.95$).

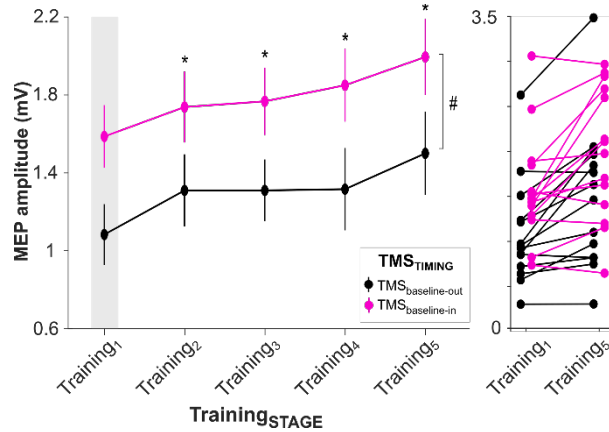


Figure 4. Evolution of baseline MEPs throughout training. MEP amplitudes (in mV) elicited at TMS_{baseline-out} (black) and TMS_{baseline-in} (pink) at the different Training_{STAGES}. Hash signs indicate a TMS_{TIMING} effect. Stars denote a significant difference between a given Training_{STAGE} and Training₁ ($p<0.05$). Individual data for Training₁ and Training₅ are also displayed.

Second, we analyzed the effect of training on preparatory suppression by considering MEPs elicited at TMS_{preparation} (expressed in percentage of TMS_{baseline-in}). As evident on **Figure 5A**, percentage FDI MEPs were initially suppressed at Training₁ (MEPs smaller than 100%), consistent with the presence of preparatory suppression in the prime-mover, whether selected for the forthcoming response (MEP_{selected}: $73.98\pm4.00\%$; $t_{(13)}=-6.50$, $p<0.0001$) or not (MEP_{non-selected}: $76.26\pm4.36\%$; $t_{(13)}=-5.44$, $p<0.001$). Interestingly, preparatory suppression became more prominent with training: the ANOVA_{RM} revealed a significant decrease in percentage MEP amplitudes over the Training_{STAGES} ($F_{(4,52)}=2.79$, $p=0.036$). This change was marginal at Training₄ (*i.e.*, Training₄: $p=0.058$ when compared to Training₁) and became significant at Training₅ ($p=0.006$). It concerned MEPs obtained from the selected and non-selected hands (Training_{STAGE}×MEP_{SELECTION}: $F_{(4,52)}=0.56$, $p=0.70$). To further our understanding of training-related changes of preparatory activity, we ran another set of ANOVA_{RM} on absolute MEP amplitudes (rather than percentages) at TMS_{preparation} (**Figure 5B**). These MEPs did not show any fluctuation over the Training_{STAGES} ($F_{(4,52)}=1.30$,

$p=0.28$). Moreover, we did not find any $\text{MEP}_{\text{SELECTION}}$ effect ($F_{(1,13)}=1.69$, $p=0.22$) or $\text{Training}_{\text{STAGE}} \times \text{MEP}_{\text{SELECTION}}$ interaction ($F_{(4,52)}=0.85$, $p=0.50$).

In conclusion, our results indicate that training did not produce even modulatory changes in motor activity at rest and during action preparation: while resting CSE increased, preparatory activity remained flat over the blocks, thus revealing an augmenting drop in preparatory activity (i.e. a strengthening of preparatory suppression) with respect to the rising baseline excitability state. These changes in CSE occurred in parallel with an acceleration of RTs.

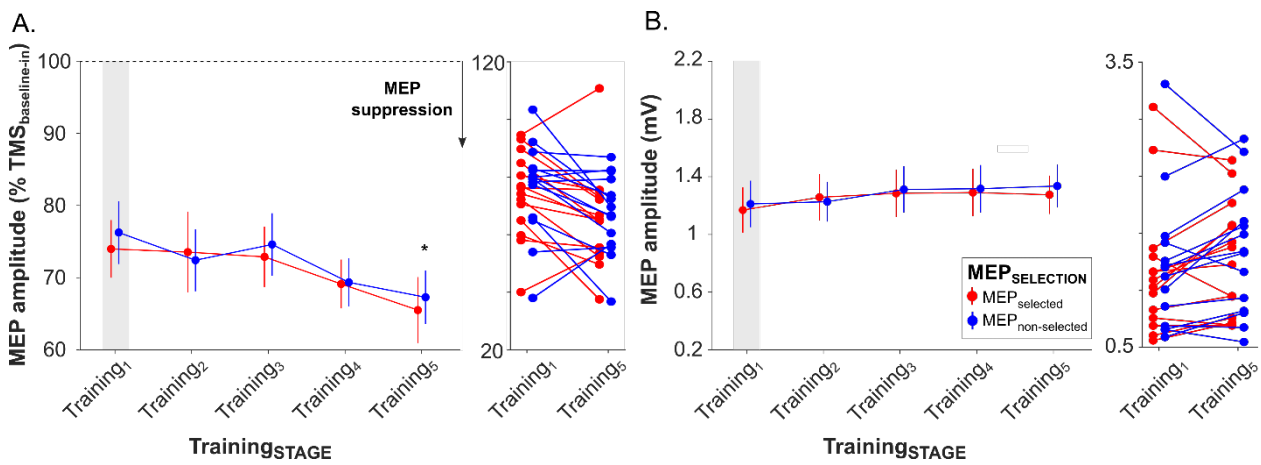


Figure 5. Evolution of preparatory MEPs throughout training. Normalized MEP amplitudes recorded at TMS_{preparation} (in percentage of MEPs elicited at TMS_{baseline-in}) muscles at the different Training_{STAGES} (A). Absolute MEP data (in mV) are also represented muscles (B). The star denotes a significant difference between a Training₅ and Training₁ ($p<0.05$). Note that the change in preparatory suppression was close to significance at Training₄ (i.e., $p=0.058$ when compared to Training₁). Individual data for Training₁ and Training₅ are also displayed.

Because RTs became shorter over the blocks, one may argue that MEPs at TMS_{preparation} were not recorded in a comparable preparatory state throughout training; that is, the delay between TMS and movement onset ($\text{Delay}_{\text{TMS-TO-MOVE}}$) may have decreased over the blocks. Importantly, we shuffled the delay between the pulse and the “Go” signal in the present study (see Methods), in order to prevent changes in RT to convert into equivalent changes in the $\text{Delay}_{\text{TMS-TO-MOVE}}$. However, because TMS fell on average closer to movement onset at Training₅ ($399.70 \pm 8.48\text{ms}$) than Training₁ ($419.99 \pm 9.99\text{ms}$, $t_{(13)}=-3.10$, $p=0.008$), we performed an additional analysis to control for a potential bias of the $\text{Delay}_{\text{TMS-TO-MOVE}}$. We conducted a response-locked analysis whereby we classified MEP data at

TMS_{preparation} (regardless of the Training_{STAGE}) according to the Delay_{TMS-TO-MOVE} in 5 consecutive bins of trials (Delay_{BIN} = Delay_{BIN-1} = 0 to 20%, Delay_{BIN-2} = 20 to 40%, ..., Delay_{BIN-5} = 80 to 100% of the Delay_{TMS-TO-MOVE} data). An ANOVA_{RM} ran on these data did not reveal any effect of Delay_{BIN} ($F_{(4,52)}=1.45$; $p=0.23$), nor was there any significant MEP_{SELECTION} × Delay_{BIN} interaction ($F_{(4,52)}=0.40$; $p=0.81$; **Figure 6**). These results indicate that MEPs elicited preceding a “Go” signal remain quite unaffected by the delay separating the TMS_{preparation} pulse and movement onset.

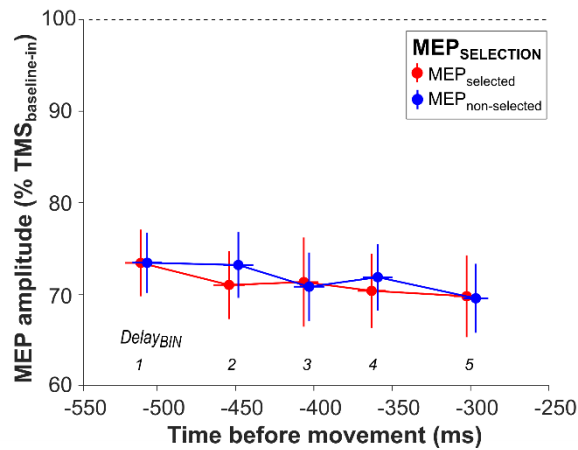


Figure 6. Preparatory suppression according to time before movement onset. MEP amplitudes recorded at TMS_{preparation} (in percentage of MEPs elicited at TMS_{baseline-in}) are represented for each Delay_{BIN} in a selected (red) or non-selected (blue) muscle.

Moreover, as an additional safety check, we also considered whether potential changes in background EMG activity could be responsible for the training-related changes in CSE reported here. To do so, we ran two additional analyses on pre-TMS EMG RMS amplitudes. First, we analysed RMS amplitude before the TMS_{baseline-in} pulses by means of a one-way ANOVA with the factor Training_{STAGE}. This analysis revealed a stable EMG background at TMS_{baseline-in} over the Training_{STAGE} (GG-corrected $F_{(1.65, 21.45)}=0.89$; $p=0.41$). In a second step, in order to make sure that a change in the relationship between EMG background activity at TMS_{preparation} with respect to TMS_{baseline-in} was not responsible for the observed training-related change in preparatory suppression, we expressed the RMS computed at TMS_{preparation} in percentage of the values obtained at TMS_{baseline-in} for each Training_{STAGE}. As in the main analysis, we did so separately for the selected and non-selected FDIs (FDI_{SELECTION}: FDI_{selected} and FDI_{non-selected}). We found that RMS amplitudes at TMS_{preparation} were very similar to the ones obtained at TMS_{baseline-in} for all Training_{STAGES}, with the

RMS recorded at TMS_{preparation} ranging from 99.79 \pm 0.76 to 101.85 \pm 1.77 % of values obtained at TMS_{baseline-in}. Consistently, the two-way ANOVA run on these data did not show any effect of Training_{STAGE} (GG-corrected $F_{(1.88, 24.49)}=0.91$; $p=0.41$), FDI_{SELECTION} ($F_{(1,13)}=1.40$; $p=0.26$) nor any Training_{STAGE} \times FDI_{SELECTION} interaction (GG-corrected $F_{(1.02, 13.20)}=1.12$; $p=0.31$). Hence, changes in EMG background activity are unlikely to be responsible for the training-related changes in preparatory suppression observed in the present study.

3.3 Relationship between training-related changes in RTs and CSE

Given that training influenced RTs and CSE, we studied the relationship between changes at these two levels, with CSE considered separately at rest and during action preparation. To assess the relationship between RTs and resting CSE, we ran correlations between training-related changes in RTs and changes in MEPs at TMS_{baseline-in} and TMS_{baseline-out}. These analyses did not reveal any link between variations in resting measures of CSE and changes observed in RTs, neither at Training_{ratio-early} (**Figure 7A**, $R=-0.27$, $p=0.36$ and $R=0.079$, $p=0.79$ for TMS_{baseline-in} and TMS_{baseline-out}, respectively) nor at Training_{ratio-late} ($R=-0.28$, $p=0.33$ and $R=-0.16$, $p=0.59$).

In contrast, changes in RTs at Training_{ratio-early} were linked to variations in preparatory suppression observed in the selected (**Figure 7B**; $R=0.55$, $p=0.043$) and non-selected FDI (**Figure 7C**; $R=0.74$, $p=0.0027$): subjects showing a greater training-related strengthening of preparatory suppression also showed larger improvements in RTs. This correlation was not significant at Training_{ratio-late}, neither for the selected ($R=0.12$, $p=0.67$) nor for the non-selected effectors ($R=0.48$, $p=0.084$). Our results suggest that RT improvements were related to early changes in preparatory suppression.

This conclusion is further supported by an additional analysis showing that the strength of the correlation between RTs and CSE at Training_{ratio-early} was significantly higher when considering percentage MEPs at TMS_{preparation} (*i.e.*, preparatory suppression) in the non-selected FDI (bootstrap estimate of absolute $R=0.76$), than when MEPs were considered at TMS_{baseline-in} ($R=0.29$; $z=1.75$; $p=0.040$, **Figure 7D**). This difference was not significant when taking preparatory suppression in the selected FDI ($R=0.56$; $z\text{-score}=0.85$,

$p=0.20$). Hence, training-related changes in preparatory suppression of the non-selected effector turned out to be the best predictor of RT improvements.

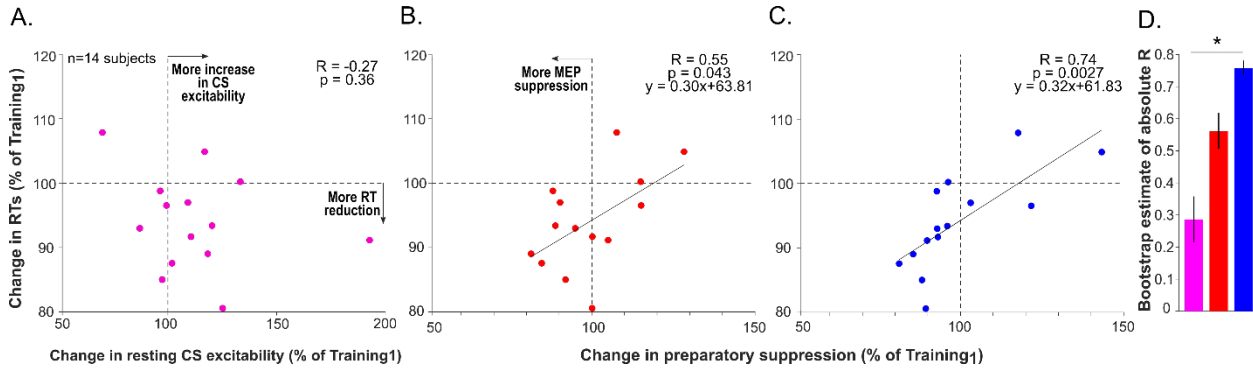


Figure 7. Correlation between early training-related changes in RTs and CSE. Changes in RTs as a function of changes in MEP amplitudes at TMS_{baseline-in} (reflecting resting CSE, **A**) and changes in percentage MEPs at TMS_{preparation} (reflecting preparatory suppression of CSE) in the selected (**B**) and non-selected FDI muscle (**C**) during the early Training_{stage}. For this analysis, changes in RTs and MEPs were assessed by computing percentage ratios between the values obtained at Training₃ and Training₁. (**D**) Bootstrap estimates of absolute R values are also displayed (\pm standard deviation of the samples) for each condition. These R values were compared by means of a Pearson and Fillon's z test. One tail p-values were used given our a priori hypothesis concerning the directionality of the effect ($p < 0.05$).

3.4 Single-trial relationship between RTs and preparatory suppression

Finally, we asked whether the dependency of RTs to preparatory suppression is also evident on a single-trial basis. This was the case for MEPs recorded from the non-selected hand: the greater the preparatory suppression in that hand, the shorter the following RT (**Figure 8**, right panel), as supported by the ANOVA_{RM} revealing an effect of the factor MEP_{BIN} on RTs ($F_{(5,65)}=2.57$, $p=0.035$). Post-hoc tests revealed that RTs in MEP_{BIN-1} and MEP_{BIN-2} (*i.e.*, strongest preparatory suppression) were systematically shorter than those in MEP_{BIN-6} ($p=0.0021$ and $p=0.0090$). We did not observe any relationship between RTs and MEPs obtained in the selected hand (MEP_{BIN}: GG-corrected $F_{(2.22,28.88)}=0.85$, $p=0.45$; **Figure 8**, left panel). Hence, the training-related effects and the single-trial relationship indicates that preparatory suppression in the non-selected (non-responding) hand is a predictor of the following RT. The lower this activity, the faster the response.

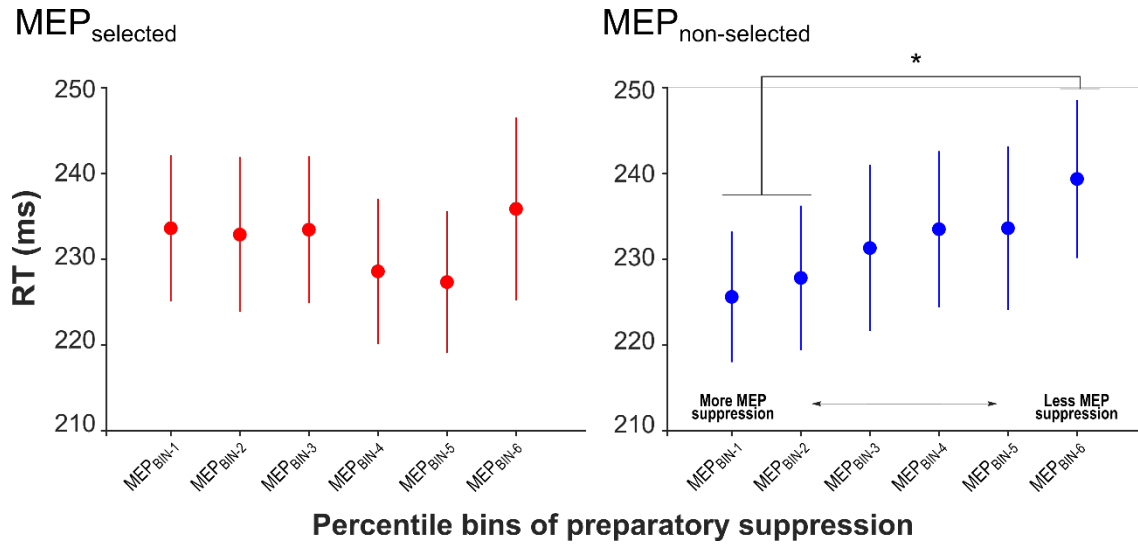


Figure 8. Single-trial MEP-RT relationship. Averaged RTs as a function of the preceding preparatory suppression in a selected (left panel) or non-selected muscle (right panel). For this analysis, the MEP data were divided in 6 MEP_{BIN} of increasing amplitude and the RTs corresponding to each MEP_{BIN} were averaged. The star denotes a significant difference between RTs at MEP_{BIN-1} and MEP_{BIN-6} in the non-selected muscle ($p < 0.05$). Note that there was also a trend for RTs at MEP_{BIN-1} to be shorter than those in MEP_{BIN-4} ($p = 0.070$) and MEP_{BIN-5} ($p = 0.066$).

4. Discussion

Training accelerated RTs while errors remained low. CSE became larger at rest and preparatory suppression of CSE was stronger after training. Interestingly, subjects who showed the strongest RT improvements at the early Training_{STAGES} were also those displaying the largest initial strengthening in preparatory suppression, especially when probed in the non-selected hand. Such a relationship between RTs and preparatory suppression was also evident at a single-trial level: RTs were generally faster in trials where preparatory suppression was deeper.

Subjects responded faster with training. RTs reflect the sum of the time required for processing the imperative cue, preparing the motor command and initiating the action (Haith et al., 2016; Derosièrè et al., 2019) and, theoretically, training may impact any of these sensory-motor components. Previous studies have shown that RT improvements can result from both faster sensory processing (Clark et al., 2015) and more efficient motor preparation (Mawase et al., 2018). Yet, in an instructed-delay task, the time required for sensory processing and motor preparation is strongly constrained and most of the RT is assumed to reflect the time needed for action initiation (Haith et al., 2016). Hence, the RT gains reported here are likely to reflect a reduction in initiation time. Our findings thus yield an extension of former work, suggesting that, in addition to accelerating sensory processing and motor preparation, training can also boost action initiation.

Resting CSE was higher when assessed in the context of the task (*i.e.*, at TMS_{baseline-in}) than between the blocks (*i.e.*, at TMS_{baseline-out}), consistent with previous data (Labruna et al., 2011; Vassiliadis et al., 2018) and with the observation that task-driven increases in attention amplifies cortical excitability (Kastner et al., 1998, 1999). As expected based on prior observations (*e.g.*, (Pascual-Leone et al., 1995; Butefisch et al., 2000; Duque et al., 2008; Galea and Celnik, 2009; Christiansen et al., 2018), practicing the task led to an increase in resting CSE. Interestingly, this increase was not exclusive to the task and was in fact strongly similar at TMS_{baseline-in} and TMS_{baseline-out}, ruling out the possibility that it resulted from a change in task-related attention over practice (Derosièrè et al., 2015). Rather, our findings support the idea of a plastic reorganization of the motor system, measurable when engaged in the task as well as at rest.

CSE was reduced during action preparation when compared to baseline (during the task), reflecting the well-known preparatory suppression effect (Duque et al., 2017), which was evident in the selected and non-selected hands from the beginning of the training. Contrary to rest, the amplitude of MEPs at TMS_{preparation} did not increase with practice (they remained unchanged), reflecting a strengthening of preparatory suppression with respect to the rising baseline. Notably, although at the group level this strengthening of preparatory suppression appeared late (**Figure 5A**), at the individual level, a majority of subjects already exhibited this effect at early training stages (**Figure 6**).

Based on these findings, one could propose that changes in resting excitability are key to RT improvements, as suggested by the inverse relationship between baseline CSE and RTs described recently (Greenhouse et al., 2017). Yet, we did not find a relationship between training-related changes in baseline excitability and improvements in performance. This is in line with the idea that increased resting CSE is not crucial for immediate performance (Bologna et al., 2015), but may be involved in the long-term retention of the motor behavior (Cantarero et al., 2013). Rather, what was predictive of RT gains in the present study was the change in relative CSE, as measured during action preparation: subjects showing the greater strengthening of preparatory suppression at the early Training_{STAGES} were those who became fastest. These results are consistent with animal studies showing that behavioral improvements in motor learning tasks are associated with changes in relative preparatory activity (Mandelblat-Cerf et al. 2009; Paz et al. 2003; Vyas et al. 2018). Similarly, a recent study using paired-pulse TMS showed that changes in preparatory activity of M1 intra-cortical circuits are correlated to training-related behavioral gains, contrary to changes observed at rest (Dupont-Hadwen et al., 2018). More generally, our findings agree with the idea that efficient action preparation relies on dynamical shifts of neural activity from a baseline state to a preparatory state (Churchland et al., 2012). From this point of view, training may allow tuning the dynamics of preparatory activity, bringing it closer to an optimal state for action initiation (Vyas et al., 2018). In this line, strengthening of preparatory suppression would facilitate action initiation by allowing excitatory inputs targeting the selected motor representation to better stand out against a quiescent background (mostly reflected in the excitability of non-selected effector), ultimately speeding up RTs (Hasbroucq et al., 1997; Greenhouse et al., 2015; Hasegawa et al., 2017).

This interpretation was reinforced by our single-trial analysis showing that RTs depended on the foregoing amount of preparatory suppression. That is, stronger levels of suppression were related to faster initiation times in the very same trials, in agreement with previous results (Hasegawa et al., 2017; Hannah et al., 2018). Interestingly, we found such relationship when considering the non-selected prime-mover but not the selected one. This was also the case for training-related effects, with preparatory suppression in the non-selected effector appearing as the best predictor of RT changes. Hence, despite the fact that preparatory suppression was broad, affecting selected and non-selected muscles, suppression in the non-selected effector turned out to be a better predictor of RTs than suppression in the selected muscle. This difference may be due to the fact that MEP amplitudes in the selected effector are influenced by many overlapping inputs, potentially reflecting the operation of response preparation processes (Duque and Ivry, 2009), while MEP amplitudes in the non-selected effector reflect a purer form of suppression. For instance, Duque and Ivry (2009) found that action preparation entails a release of M1-intracortical inhibition in parallel of the suppression of CSE, specifically in the selected effector. The multiplicity of inputs targeting the selected effector during action preparation may therefore have weakened the link between our measure of preparatory suppression and RTs. Overall, we propose that motor initiation is eased by a form of preparatory suppression encompassing selected and non-selected effectors (Greenhouse et al., 2015) but that the link between suppression and initiation is best evidenced when considering the excitability of non-selected effectors.

5. Conclusion

This study shows that a simple training paradigm can lead to improvements in action initiation that are accompanied by an increase in resting CSE and a strengthening of corticospinal suppression from the rising baseline state. Moreover, contrary to changes in resting CSE, such strengthening of preparatory suppression was linked to RT improvements, especially in the non-selected effector. These findings could have implications for the rehabilitation of patients suffering from impaired action initiation such as in cerebellar ataxia (Battaglia et al., 2006) or Parkinson's disease (Mure et al., 2012).

References

- Algoet M, Duque J, Iannetti GD, Mouraux A. Temporal Profile and Limb-specificity of Phasic Pain-Evoked Changes in Motor Excitability. *Neuroscience* 386: 240–255, 2018.
- Battaglia F, Quartarone A, Ghilardi MF, Dattola R, Bagnato S, Rizzo V, Morgante L, Girlanda P. Unilateral cerebellar stroke disrupts movement preparation and motor imagery. *Clin Neurophysiol* 117: 1009–1016, 2006.
- Bologna M, Rocchi L, Paparella G, Nardella A, Li P, Conte A, Kojovic M, Rothwell JC, Berardelli A. Brain Stimulation Reversal of Practice-related Effects on Corticospinal Excitability has no Immediate Effect on Behavioral Outcome. *Brain Stimul* 8: 603–612, 2015.
- Butefisch CM, Davis BC, Wise SP, Sawaki L, Kopylev L, Classen J, Cohen LG. Mechanisms of use-dependent plasticity in the human motor cortex. *Proc Natl Acad Sci* 97: 3661–3665, 2000.
- Cantarero G, Lloyd A, Celnik P. Reversal of Long-Term Potentiation-Like Plasticity Processes after Motor Learning Disrupts Skill Retention. *J Neurosci* 33: 12862–12869, 2013.
- Christiansen L, Madsen MJ, Bojsen-møller E, Thomas R, Nielsen JB. Brain Stimulation Progressive practice promotes motor learning and repeated transient increases in corticospinal excitability across multiple days. *Brain Stimul* 11: 346–357, 2018.
- Churchland MM, Cunningham JP, Kaufman MT, Foster JD, Nuyujukian P, Ryu SI, Shenoy K V., Shenoy K V. Neural population dynamics during reaching. *Nature* 487: 51–56, 2012.
- Cincotta M, Borgheresi A, Jung P, Balestrieri F, Giovannelli F, Zaccara G, Ziemann U. Physical interactions between induced electrical fields can have substantial effects on neuronal excitation during simultaneous TMS of two brain areas. *Clin Neurophysiol* 116: 1733–1742, 2005.
- Clark K, Gregory Appelbaum L, van den Berg B, Mitroff SR, Woldorff MG. Improvement in visual search with practice: Mapping learning-related changes in neurocognitive stages of processing. *J Neurosci* 35: 5351–5359, 2015.
- Derosiere G. A Dynamical System Framework for Theorizing Preparatory Inhibition. *J Neurosci* 38: 3391–3393, 2018.
- Derosière G, Billot M, Ward ET, Perrey S. Adaptations of motor neural structures' activity to lapses in attention. *Cereb Cortex* 25: 66–74, 2015.
- Derosiere G, Duque J. Tuning the Corticospinal System: How Distributed Brain Circuits Shape Human Actions. *Neurosci* 107385841989675, 2020.
- Derosiere G, Thura D, Cisek P, Duque J. Motor cortex disruption delays motor processes but not deliberation about action choices. *J Neurophysiol* 122: 1566–1577, 2019.
- Derosiere G, Vassiliadis P, Demaret S, Zénon A, Duque J. Learning stage-dependent effect of M1 disruption on value-based motor decisions. *Neuroimage* , 2017a. doi:10.1016/j.neuroimage.2017.08.075.
- Derosiere G, Vassiliadis P, Duque J. Advanced TMS approaches to probe corticospinal excitability during action preparation. *Neuroimage* 213: 116746, 2020.

- Derosiere G, Zénon A, Alamia A, Duque J. Primary motor cortex contributes to the implementation of implicit value-based rules during motor decisions. *Neuroimage* 146: 1115–1127, 2017b.
- Dupont-Hadwen J, Bestmann S, Stagg CJ. Motor training modulates intracortical inhibitory dynamics in motor cortex during movement preparation. *Brain Stimul* , 2018. doi:10.1016/j.brs.2018.11.002.
- Duque J, Greenhouse I, Labruna L, Ivry RB. Physiological Markers of Motor Inhibition during Human Behavior. *Trends Neurosci* 40: 219–236, 2017.
- Duque J, Ivry RB. Role of corticospinal suppression during motor preparation. *Cereb Cortex* 19: 2013–2024, 2009.
- Duque J, Labruna L, Cazaes C, Ivry RB. Dissociating the influence of response selection and task anticipation on corticospinal suppression during response preparation. *Neuropsychologia* 65: 287–296, 2014.
- Duque J, Lew D, Mazzocchio R, Olivier E, Richard B. Evidence for two concurrent inhibitory mechanisms during response preparation. *J Neurosci* 30: 3793–3802, 2010.
- Duque J, Mazzocchio R, Stefan K, Hummel F, Olivier E, Cohen LG. Memory formation in the motor cortex ipsilateral to a training hand. *Cereb Cortex* 18: 1395–1406, 2008.
- Duque J, Petitjean C, Swinnen SP. Effect of aging on motor inhibition during action preparation under sensory conflict. *Front Aging Neurosci* 8: 1–14, 2016.
- Efron B. Bootstrap Methods: Another Look at the Jackknife. *Ann Stat* 7: 1–26, 1979.
- Ferbert A, Priori A, Rothwell JC, Day BL, Colebatch JG, Marsden CD. Interhemispheric inhibition of the human motor cortex. *J Physiol* 453: 525–546, 1992.
- Galea JM, Celnik P. Brain Polarization Enhances the Formation and Retention of Motor Memories. *J Neurophysiol* 102: 294–301, 2009.
- Grandjean J, Derosiere G, Vassiliadis P, Quemener L, Wilde Y de, Duque J. Towards assessing corticospinal excitability bilaterally: Validation of a double-coil TMS method. *J Neurosci Methods* 293: 162–168, 2018.
- Grandjean J, Duque J. A TMS Study of Preparatory Suppression in Binge Drinkers. *NeuroImage Clin* 28: 102383, 2020.
- Grandjean J, Quoilin C, Duque J. Investigating the effect of anticipating a startling acoustic stimulus on preparatory inhibition. *Neurophysiol Clin* 49: 137–147, 2019.
- Greenhouse I, Sias a., Labruna L, Ivry RB. Nonspecific Inhibition of the Motor System during Response Preparation. *J Neurosci* 35: 10675–10684, 2015.
- Greenhouse XI, King XM, Noah XS, Maddock XRJ, Ivry RB. Individual Differences in Resting Corticospinal Excitability Are Correlated with Reaction Time and GABA Content in Motor Cortex. *J Neurosci* 37: 2686–2696, 2017.
- Groppa S, Oliviero A, Eisen A, Quartarone A, Cohen LG, Mall V, Kaelin-Lang A, Mima T, Rossi S, Thickbroom GW, Rossini PM, Ziemann U, Valls-Solé J, Siebner HR. A practical guide to diagnostic transcranial magnetic stimulation: Report of an IFCN committee. *Clin Neurophysiol* 123: 858–882, 2012.

- Haith AM, Pakpoor J, Krakauer JW. Independence of Movement Preparation and Movement Initiation. *J Neurosci* 36: 3007–3015, 2016.
- Hanajima R, Ugawa Y, Machii K, Mochizuki H, Terao Y, Enomoto H, Furubayashi T, Shiio Y, Uesugi H, Kanazawa I. Interhemispheric facilitation of the hand motor area in humans. *J Physiol* 531: 849–859, 2001.
- Hannah R, Cavanagh SE, Tremblay S, Simeoni S, Rothwell JC. Selective suppression of local interneuron circuits in human motor cortex contributes to movement preparation. *J Neurosci* 38: 2869–17, 2018.
- Hasbroucq T, Kaneko H, Akamatsu M, Possamaï CA. Preparatory inhibition of cortico-spinal excitability: A transcranial magnetic stimulation study in man. *Cogn Brain Res* 5: 185–192, 1997.
- Hasegawa M, Majima K, Itokazu T, Maki T, Albrecht UR, Castner N, Izumo M, Sohya K, Sato TK, Kamitani Y, Sato TR. Selective Suppression of Local Circuits during Movement Preparation in the Mouse Motor Cortex. *Cell Rep* 18: 2676–2686, 2017.
- Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG. Increased Activity in Human Visual Cortex during Directed Attention in the Absence of Visual Stimulation. *Neuron* 22: 751–761, 1999.
- Kastner S, De Weerd P, Desimone R, Ungerleider LG. Mechanisms of Directed Attention in the Human Extrastriate Cortex as Revealed by Functional MRI. *Science* (80-) 282: 108–111, 1998.
- Klein PA, Petitjean C, Olivier E, Duque J. Top-down suppression of incompatible motor activations during response selection under conflict. *Neuroimage* 86: 138–149, 2014.
- Krakauer JW, Ghilardi MF, Mentis M, Barnes A, Veytsman M, Eidelberg D, Ghez C. Differential Cortical and Subcortical Activations in Learning Rotations and Gains for Reaching: A PET Study. *J Neurophysiol* 91: 924–933, 2004.
- Krakauer JW, Hadjiosif AM, Xu J, Wong AL, Haith AM. *Motor Learning*. 9: 613–663, 2019.
- Labruna L, Fernández-Del-Olmo M, Ivry RB. Comparison of different baseline conditions in evaluating factors that influence motor cortex excitability. *Brain Stimul* 4: 152–155, 2011.
- Makino H, Ren C, Liu H, Kim AN, Kondapaneni N, Liu X, Kuzum D, Komiyama T. Transformation of Cortex-wide Emergent Properties during Motor Learning. *Neuron* 94: 880-890.e8, 2017.
- Mandelblat-cerf Y, Paz R, Vaadia E. Trial-to-Trial Variability of Single Cells in Motor Cortices Is Dynamically Modified during Visuomotor Adaptation. *J Neurosci* 29: 15053–15062, 2009.
- Mawase F, Lopez D, Celnik PA, Haith AM. Movement Repetition Facilitates Response Preparation. *Cell Rep* 24: 801–808, 2018.
- Mawase F, Uehara S, Bastian AJ, Celnik P. Motor Learning Enhances Use-Dependent Plasticity. *J Neurosci* 37: 2673–2685, 2017.
- Mure H, Tang CC, Argyelan M, Ghilardi MF, Kaplitt MG, Dhawan V, Eidelberg D. Improved sequence learning with subthalamic nucleus deep brain stimulation: Evidence for treatment-specific network modulation. *J Neurosci* 32: 2804–2813, 2012.
- Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9: 97–113, 1971.

- Pascual-Leone A, Nguyet D, Cohen LG, Brasil-Neto JP, Cammarota A, Hallett M, Nguyen KT, Cohen a D, Brasil-Neto JP, Cammarota A, Hallett M, Nguyet D, Cohen LG, Brasil-Neto JP, Cammarota A, Hallett M, Cohen G, Cammarota A, Pascual-Leone A, Nguyet D, Cohen LG, Brasil-Neto JP, Hallett M, Cammarota A, Hallett M, Cohen G, Cammarota A, Pascual-Leone A, Nguyet D, Cohen LG, Brasil-Neto JP, Hallett M. Modulation of muscle responses evoked by transcranial magnetic stimulation during the acquisition of new fine motor skills. [Online]. *J Neurophysiol* 74: 1037–1045, 1995<http://www.ncbi.nlm.nih.gov/pubmed/7500130>.
- Paz R, Boraud T, Natan C, Bergman H, Vaadia E. Preparatory activity in motor cortex reflects learning of local visuomotor skills. *6*: 882–890, 2003.
- Pearson K, Filon L. VII. Mathematical contributions to the theory of evolution.— IV. On the probable errors of frequency constants and on the influence of random selection on variation and correlation. *Philos Trans R Soc London* 191: 229–311, 1898.
- Quoilin C, Fievez F, Duque J. Preparatory inhibition: Impact of choice in reaction time tasks. *Neuropsychologia* 129: 212–222, 2019.
- Quoilin C, Grandjean J, Duque J. Considering Motor Excitability During Action Preparation in Gambling Disorder: A Transcranial Magnetic Stimulation Study. *Front Psychiatry* 11: 1–13, 2020.
- Quoilin C, Lambert J, Jacob B, Klein PA, Duque J. Comparison of motor inhibition in variants of the instructed-delay choice reaction time task. *PLoS One* 11: 1–16, 2016.
- Quoilin C, Wilhelm E, Maurage P, Timary P De, Duque J. Deficient inhibition in alcohol-dependence : let's consider the role of the motor system ! *Neuropsychopharmacology* 0–8, 2018.
- Reis J, Swayne OB, Vandermeeren Y, Camus M, Dimyan MA, Harris-Love M, Perez MA, Ragert P, Rothwell JC, Cohen LG. Contribution of transcranial magnetic stimulation to the understanding of cortical mechanisms involved in motor control. *J Physiol* 586: 325–351, 2008.
- Rosenkranz K, Kacar A, Rothwell JC. Differential modulation of motor cortical plasticity and excitability in early and late phases of human motor learning. *J Neurosci* 27: 12058–12066, 2007.
- Shmuelof L, Yang J, Caffo B, Mazzoni P, Krakauer JW. The neural correlates of learned motor acuity. 971–980, 2014.
- Steele CJ, Penhune VB. Specific increases within global decreases: A functional magnetic resonance imaging investigation of five days of motor sequence learning. *J Neurosci* 30: 8332–8341, 2010.
- Vassiliadis P. Training_Motor_Preparation_TMS [Online]. 2020.<https://osf.io/8p5wm/>.
- Vassiliadis P, Grandjean J, Derosiere G, Wilde Y De. Using a Double-Coil TMS Protocol to Assess Preparatory Inhibition Bilaterally. *Front Neurosci* 12: 1–14, 2018.
- Vyas S, Even-Chen N, Stavisky SD, Ryu SI, Nuyujukian P, Shenoy K V. Neural Population Dynamics Underlying Motor Learning Transfer. *Neuron* 97: 1177-1186.e3, 2018.
- Vyas S, O'Shea DJ, Ryu SI, Shenoy K. Causal role of motor preparation during error-driven learning. *Neuron* 1–11, 2020.

Wenger E, Kühn S, Verrel J, Mårtensson J, Bodammer NC, Lindenberger U, Lövdén M. Repeated Structural Imaging Reveals Nonlinear Progression of Experience-Dependent Volume Changes in Human Motor Cortex. *Cereb Cortex* 27: 2911–2925, 2017.

Wessel MJ, Draaisma LR, de Boer AFW, Park C hyun, Maceira-Elvira P, Durand-Ruel M, Koch PJ, Morishita T, Hummel FC. Cerebellar transcranial alternating current stimulation in the gamma range applied during the acquisition of a novel motor skill. *Sci Rep* 10: 1–12, 2020.

Wiestler T, Diedrichsen J. Skill learning strengthens cortical representations of motor sequences. *Elife* 2013: 1–20, 2013.

Yokoi A, Diedrichsen J. Neural Organization of Hierarchical Motor Sequence Representations in the Human Neocortex. *Neuron* 103: 1178-1190.e7, 2019.

Annex 2: Selecting and executing actions for rewards

Pierre Vassiliadis^{1,2*}, Gerard Derosiere¹

¹ Institute of Neuroscience, Université Catholique de Louvain, 1200, Brussels, Belgium.

² Defitech Chair for Clinical Neuroengineering, Center for Neuroprosthetics (CNP) and Brain Mind Institute (BMI), Swiss Federal Institute of Technology (EPFL), 1202, Geneva, Switzerland

Reference: Vassiliadis P, Derosiere G (2020). Selecting and executing actions for rewards. The Journal of Neuroscience, 40 (34) 6474-6476.

Review of: Codol O, Holland PJ, Manohar SG, Galea JM (2020) Reward-based improvements in motor control are driven by multiple error-reducing mechanisms. J Neurosci 40:3604-3620.

Personal contribution: Review of literature, Writing – Original Draft, Writing – Review & Editing, Funding acquisition.

* Correspondence

Pierre Vassiliadis
Institute of Neuroscience
Université catholique de Louvain
53, Avenue Mounier
COSY- B1.53.04
B-1200 BRUSSELS
Belgium
Email: pierre.vassiliadis@uclouvain.be

Rewards shape human actions. The mere possibility of earning a reward induces substantial improvements in the way we choose and execute actions (Chen et al., 2017). This observation has raised hope for rehabilitation: reward is regarded as a promising means to magnify the positive effects of practice on motor control (Quattrocchi et al., 2017). Yet, this branch of research is only burgeoning, and neuroscientists have yet to identify the mechanisms through which reward improves movements.

At present, two distinct spheres of study have provided insights into how reward improves motor control. First, studies on action selection show that reward can speed up reaction times (RTs; the time elapsed between stimulus presentation and action initiation; Klein et al., 2012) and enhance selection accuracy (subjects select the “right action” more often when reward is at stake; (Derosiere et al., 2017b, 2017c). Second, studies on action execution reveal a beneficial effect of reward on movement times (MTs; the time elapsed between action initiation and completion; Reppert et al., 2015) and execution accuracy (e.g., subjects execute faster and more precise movements when reaching to a rewarding target; Manohar et al., 2019). Strikingly, most work investigating the effects of reward on action selection and execution have examined these effects in separate studies (Chen et al., 2017), impeding the genesis of an integrative understanding of how reward shapes the two processes in more natural settings, where the effects are likely to co-exist.

Even when considered in isolation, the precise mechanisms underlying the effects of reward on selection and execution processes have remained obscure. An important gap in our knowledge concerns how reward improves execution accuracy. One possibility is that the presence of reward increases limb stiffness, enhancing the resistance of the moving effectors to internal and external perturbations (Gribble et al., 2003) and ultimately reducing movement variability (so-called “motor noise”; Manohar et al., 2015). Yet, the contribution of stiffness to reward-driven improvements in execution accuracy has been speculative.

In a recent article published in *The Journal of Neuroscience*, Codol et al. (2020) addressed the two issues mentioned above. In a series of experiments, the authors asked human subjects to use reaching movements to displace a manipulandum from a starting position to one of four target locations. Before starting each movement, subjects were informed of the maximum reward they could obtain in the trial (0, 10 or 50 pence). In 10p and 50p trials, the magnitude of the reward ultimately obtained by the subject depended on her/his performance in the trial (see below).

The first aim of the study was to test the impact of reward on the speed and accuracy of action selection and execution in a single setting. To do so, on 10p and 50p trials, Codol et al. (2020) provided rewards that were inversely proportional to the RT and the MT combined together (reflecting the speed of selection and execution processes, respectively). Reward magnitude also depended on the accuracy of both selection and execution. Importantly, some trials required subjects to ignore distractor cues; initiating a movement towards these cues was classified as a selection error and thus unrewarded. Furthermore, trials on which the final position of the manipulandum fell more than 4 cm away from the target center were classified as an execution error and thus unrewarded. Hence, to maximize reward in 10p and 50p trials, participants had to select and execute reaching movements as quickly as possible while keeping both selection and execution accuracy high.

A second objective of the study was to test the contribution of limb stiffness to reward-driven improvements in execution accuracy. To investigate this, the authors had subjects perform the same task as described above with the addition that some trials involved a displacement of the manipulandum after movement completion, pushing subjects' arm away from the target. Arm stiffness was evaluated by measuring the amount of force exerted by the subject during this perturbation. The authors were able to assess the impact of reward on stiffness by comparing this measure of force in 50p versus in 0p trials. In a control experiment, the authors also tested the effect of reward on arm stiffness before movement initiation (*i.e.*, the displacement of the manipulandum pushed subjects' arm away from the starting position).

The results indicated that, when considered in a single task, reward can have a dissociable impact on action selection and execution. Indeed, 10p and 50p trials were not associated with any change in selection speed (*i.e.*, no significant effect on RTs, compared to 0p trials), but entailed a boost of execution speed (*i.e.*, a reduction in MTs). Conversely, selection accuracy was enhanced in rewarded trials (*i.e.*, a smaller proportion of movements were initiated towards distractor cues than on 0p trials) while execution accuracy remained unchanged (*i.e.*, the deviation between the manipulandum final position and the target center was stable). Interestingly, computational analyses revealed that the maintenance of high execution accuracy in rewarded trials (despite faster MTs) could be in part attributed to a reduction in motor noise. Most importantly, this reduction in motor noise was associated with

a substantial increase in arm stiffness in 50p compared to 0p trials that was observed specifically at the end of the reaching movement (and not before movement initiation), thus confirming the contribution of end-point stiffness to reward-driven improvements in execution accuracy.

The dissociable effect of reward on selection and execution speeds is striking. In fact, one major framework in motor neuroscience views selection and execution processes as part of a continuum with a shared neural basis centered on the motor system (Cisek, 2007). In this view, altered activity in specific neural structures (e.g., in the case of reward processing, midbrain dopaminergic neurons; Schultz, 2015) could produce changes in both selection and execution processes at the behavioral level. The roots of this idea lie so deep within the field that researchers often consider RTs and MTs together as a single measure, thought to reflect action vigor (Shadmehr et al., 2019). The findings of Codol et al. (2020) ask us to reconsider carefully this vision, suggesting that, in some conditions, the speed of action selection and execution can be regulated by independent (yet likely interacting) neural structures. Consistent with this hypothesis, a recent study revealed the existence of distinct subpopulations of midbrain dopaminergic neurons, with some cells encoding behavioral choice and others sensitive to movement features (Engelhard et al., 2019).

An alternative explanation for the lack of effect of reward on RTs may however arise if one concedes that this measure not only reflects the speed of action selection but also the rapidity of sensory processing (Haith et al., 2016; Vassiliadis et al., 2020b), and that reward could have affected these two processes in opposite ways. In fact, the task described above put a considerable demand on sensory processing, as it required participants to discriminate between four target locations and, in some trials, to avoid distractor cues. This time-consuming process relies on attentional mechanisms that amplify and suppress neural responses in visual neurons encoding target and distractor cues, respectively (Itthipuripat et al., 2019). The prospect of reward may have strengthened the emphasis on such attentional mechanisms, slowing them down to take more time to sharpen visual activity. Importantly, this interpretation offers a potential mechanistic explanation for how subjects may have improved selection accuracy in rewarded trials. Notably, if such a scenario holds true, the lack of effect of reward on RTs may have emerged from a concomitant, antagonistic hastening of action selection. In this case, the increase in selection speed would be concurrent to the

boost of execution speed, and would be therefore in accordance with the continuum framework mentioned above. This hypothesis suggests new avenues of research, aiming to disentangle the effects of reward on the different processes occurring between sensation and action.

Another important finding of the study is that reward reduced motor noise through increased limb stiffness, limiting the potential negative consequence of high execution speed on accuracy. Interestingly, the movement pattern reported by Codol et al. (2020) — a parallel increase in movement speed and stiffness during rewarded trials — is similar to that observed when participants are exposed to unpredictable perturbations of their movements during execution (Crevecoeur et al., 2019). This pattern is thought to reflect the implementation of a specific strategy of the motor system (so-called “robust strategy”), minimizing the impact of perturbations on action execution in uncertain environments (Bian et al., 2020). Critically, the results of Codol et al. (2020) suggest that the presence of reward also influences the reliance on such a robust strategy. More generally, the reliance of the motor system on this strategy may depend on the expected outcome of a movement: it increases both when the risk of execution failure is high (*i.e.*, in uncertain environments) and when adequate execution can lead to a reward.

The finding of a reward-driven increase in stiffness has at least two major implications for the development of rehabilitation protocols. First, high stiffness may induce muscular fatigue, a process that might reduce the magnitude of rehabilitative learning (Branscheidt et al., 2019). Therefore, therapists should track patients’ fatigue systematically when training involves reward. Second, the ability to regulate limb stiffness could be a relevant marker of whether a patient may or may not benefit from reward-based rehabilitation. For instance, patients with excessive stiffness (*e.g.*, due to post-stroke spasticity) may not display the reward-driven improvements in execution reported by Codol et al. (2020), at least not without appropriate anti-spastic treatment.

To conclude, the study by Codol et al. (2020) builds on timely questions regarding the mechanisms underlying the impact of reward on motor control. In a series of experiments, the authors show that the presence of reward can have dissociable impacts on action selection and execution, with effects on the latter process associated with increased arm stiffness. As we discussed, these findings provide mechanistic insights and have implications for future clinical translation.

References

- Bian T, Wolpert DM, Jiang ZP (2020) Model-free robust optimal feedback mechanisms of biological motor control. *Neural Comput* 32:562–595.
- Branscheidt M, Kassavetis P, Anaya M, Rogers D, Huang HD, Lindquist MA, Celnik P (2019) Fatigue induces long-lasting detrimental changes in motor-skill learning. *Elife* 8 Available at: <https://elifesciences.org/articles/40578>.
- Chen X, Holland P, Galea JM (2017) The effects of reward and punishment on motor skill learning. *Curr Opin Behav Sci* 20:83–88 Available at: <http://dx.doi.org/10.1016/j.cobeha.2017.11.011>.
- Cisek P (2007) Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos Trans R Soc Lond B Biol Sci* 362:1585–1599.
- Codol O, Holland PJ, Manohar SG, Galea JM (2020) Reward-Based Improvements in Motor Control Are Driven by Multiple Error-Reducing Mechanisms. *J Neurosci* 40:3604–3620 Available at: <http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.2646-19.2020>.
- Crevecoeur F, Scott SH, Cluff T (2019) Robust Control in Human Reaching Movements: A Model-Free Strategy to Compensate for Unpredictable Disturbances. *J Neurosci* 39:8135–8148.
- Derosiere G, Vassiliadis P, Demaret S, Zénon A, Duque J (2017a) Learning stage-dependent effect of M1 disruption on value-based motor decisions. *Neuroimage* 162:173–185.
- Derosiere G, Zénon A, Alamia A, Duque J (2017b) Primary motor cortex contributes to the implementation of implicit value-based rules during motor decisions. *Neuroimage* 146:1115–1127 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2016.10.010>.
- Engelhard B, Finkelstein J, Cox J, Fleming W, Jang HJ, Ornelas S, Koay SA, Thiberge SY, Daw ND, Tank DW, Witten IB (2019) Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature* 570:509–513 Available at: <http://dx.doi.org/10.1038/s41586-019-1261-9>.
- Gribble PL, Mullin LI, Cothros N, Mattar A (2003) Role of cocontraction in arm movement accuracy. *J Neurophysiol* 89:2396–2405.
- Haith AM, Pakpoor J, Krakauer JW (2016) Independence of Movement Preparation and Movement Initiation. *J Neurosci* 36:3007–3015 Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.3245-15.2016> <http://www.ncbi.nlm.nih.gov/pubmed/26961954>.
- Itthipuripat S, Vo VA, Sprague TC, Serences JT (2019) Value-driven attentional capture enhances distractor representations in early visual cortex. *PLoS Biol* 17.
- Klein P-A, Olivier E, Duque J (2012) Influence of Reward on Corticospinal Excitability during Movement Preparation. *J Neurosci* 32:18124–18136.
- Manohar SG, Chong TT, Apps MAJ, Jarman PR, Bhatia KP, Husain M, Manohar SG, Chong TT, Apps MAJ, Batla A, Stamelou M (2015) Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control Article

Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control. *Curr Biol* 25:1707–1716 Available at: <http://dx.doi.org/10.1016/j.cub.2015.05.038>.

Manohar SG, Muhammed K, Fallon SJ, Husain M (2019) Neuropsychologia Motivation dynamically increases noise resistance by internal feedback during movement. *Neuropsychologia* 123:19–29 Available at: <https://doi.org/10.1016/j.neuropsychologia.2018.07.011>.

Quattrocchi G, Greenwood R, Rothwell JC, Galea JM, Bestmann S (2017) Reward and punishment enhance motor adaptation in stroke. *J Neurol Neurosurg Psychiatry:jnnp-2016-314728* Available at: <http://jnnp.bmj.com/lookup/doi/10.1136/jnnp-2016-314728>.

Reppert TR, Lempert KM, Glimcher PW, Shadmehr R (2015) Modulation of saccade vigor during value-based decision making. *J Neurosci* 35:15369–15378.

Schultz W (2015) Neuronal Reward and Decision Signals: From Theories to Data. *Physiol Rev* 95:853–951 Available at: <http://physrev.physiology.org/lookup/doi/10.1152/physrev.00023.2014>.

Shadmehr R, Reppert TR, Summerside EM, Yoon T, Ahmed AA (2019) Movement Vigor as a Reflection of Subjective Economic Utility. *Trends Neurosci* xx:1–14 Available at: <https://doi.org/10.1016/j.tins.2019.02.003>.

Vassiliadis P, Derosièrè G, Grandjean J, Duque J (2020) Motor training strengthens corticospinal suppression during movement preparation. *bioRxiv*:2020.02.14.948877.

Annex 3: Beyond motor noise: considering other causes of impaired reinforcement learning in cerebellar patients

Pierre Vassiliadis^{1,2*}, Gerard Derosiere¹, Julie Duque¹

¹ Institute of Neuroscience, Université Catholique de Louvain, 1200, Brussels, Belgium.

² Defitech Chair for Clinical Neuroengineering, Center for Neuroprosthetics (CNP) and Brain Mind Institute (BMI), Swiss Federal Institute of Technology (EPFL), 1202, Geneva, Switzerland

Reference: Vassiliadis P, Derosiere G, Duque J (2019). Beyond motor noise: considering other causes of impaired reinforcement learning in cerebellar patients. eNeuro, 6, 1.

Review of: Therrien A, Wolpert D and Bastian AJ (2018). Increasing Motor Noise Impairs Reinforcement Learning in Healthy Individuals. eNeuro. 2018 Aug 13; 5(3).

Personal contribution: Review of literature, Writing – Original Draft, Writing – Review & Editing, Funding acquisition.

* Correspondence

Pierre Vassiliadis

Institute of Neuroscience

Université catholique de Louvain

53, Avenue Mounier

COSY- B1.53.04

B-1200 BRUSSELS

Belgium

Email: pierre.vassiliadis@uclouvain.be

Significance statement:

Motor and reinforcement learning have been classically linked to functionally independent brain networks centered on the cerebellum and the basal ganglia respectively. In a recent study published in *eNeuro*, Therrien et al. (2018) showed that increasing motor noise in healthy subjects disrupts reinforcement learning. However, this impairment remained well below that detected in cerebellar patients even when motor noise in healthy subjects was adjusted to match that observed in the patients. This suggests that impaired reinforcement learning following cerebellar damage cannot be solely accounted for by altered motor noise in these patients. Based on recent anatomical and functional evidence, we argue that the cerebellum may directly contribute to reinforcement learning, consistent with its tight connections with the basal ganglia.

The ability to adapt to changes occurring in the environment is a fundamental feature of human behavior, which relies on both sensory and reward feedback. On the one hand, the role of sensory feedback has been largely considered by studying how motor commands adapt to visual perturbations (e.g., a visuomotor rotation), a process called error-based learning (Shadmehr et al., 2010; Wolpert et al., 2011; Kim et al., 2018; Roemmich and Bastian, 2018). This type of motor learning involves the computation of sensory prediction errors (SPE), namely the difference between predicted and actual sensory outcome (Tseng et al., 2007; Schlerf and Ivry, 2012; Shadmehr, 2017, 2018). On the other hand, the role of reward feedback has been mostly investigated in tasks that require learning what action to select or not, by updating reward predictions based on previous experience, a process named reinforcement learning (Lee et al., 2012; Derosiere et al., 2017b, 2017a; Gershman and Daw, 2017; O'Doherty et al., 2017). A central aspect here is the computation of reward prediction errors (RPE), namely the difference between predicted and actual rewards (Schultz, 2015).

For a long time, motor learning and reinforcement learning have been studied apart and have been linked to functionally independent brain networks (Doya, 2000), mostly centered on either the cerebellum (Tseng et al., 2007; Schlerf and Ivry, 2012; Taylor and Ivry, 2014; Herzfeld et al., 2018) or on dopaminergic-basal ganglia circuits (Lee et al., 2012; O'Doherty et al., 2017), respectively. However, this view has changed in the past few years, with recent works indicating that rewards can strongly impact motor learning (Abe et al., 2011; Dayan et al., 2011; Izawa and Shadmehr, 2011; Galea et al., 2015; Nikooyan et al., 2015; Quattrocchi et al., 2017; Song and Smiley-Oyen, 2017). Hence, efforts are now made to understand how motor and reinforcement learning may interact at the neural level (Wilkinson et al., 2015; Mawase et al., 2017; Uehara et al., 2017). Consistently, Therrien et al. (2016) recently reported data pointing towards an implication of the cerebellum in reinforcement-based motor learning. As such, cerebellar patients exhibited a reduced ability to learn from reinforcement in a visuomotor adaptation task, compared to healthy subjects (Therrien et al., 2016). However, because the patients also exhibited increased motor noise (i.e., defined as an uncontrollable source of motor variability), this deficit could have occurred indirectly, due to an impaired ability to precisely relate an action to the received reward. The work reported by the same authors in *eNeuro* aimed at tackling this issue by increasing motor noise artificially in healthy individuals (Therrien et al., 2018).

Therrien et al. (2018) used the same visuomotor adaptation task as in their previous 2016 paper. Specifically, the healthy subjects were required to make reaching movements towards a visual target with no visual information on the position of their hand and received binary reward feedback. In order to get a reward, subjects had to learn to alter their reach angle to counteract a visuomotor rotation between the location of the visual target and their hand position: the reward feedback was based on this modified reach angle. Critically, the authors added motor noise by deviating the subjects' reach in a way that was proportional to baseline motor variability. The experimental design involved both a low and a high noise condition, with the latter set to approximate the level of motor noise observed in cerebellar patients (Therrien et al., 2016). Hence, such approach allowed comparing reinforcement learning abilities between patients and healthy controls with comparable motor noise. The authors report an impaired reinforcement learning in healthy individuals in the high-noise compared to the control condition (i.e., when no noise was added). Yet, one critical result is that this impairment remained well below that observed in cerebellar patients. This finding indicates that motor noise does not entirely account for the reinforcement learning deficits observed following cerebellar damage.

A main line of argumentation in the paper focuses on the reduced proprioceptive acuity of cerebellar patients. As such, even with added motor noise, healthy subjects can still relate the rewards they receive to their reach angle, based on proprioception, while this ability is known to be altered in cerebellar patients (Miall and King, 2008; Bhanpuri et al., 2013; Weeks et al., 2017, 2018). Hence, the reduction in proprioceptive precision might have indirectly altered reinforcement learning in the patients. Note though that the clinical tests run by Therrien et al. (2016) failed to reveal any reduction in proprioceptive precision in the patients. Hence, even if a discrete reduction in proprioceptive precision could have gone unnoticed based on clinical tests (Rinderknecht et al., 2018), we would like to propose that alterations in proprioceptive precision may not completely explain reinforcement learning deficits observed in the patients. Rather, cerebellar damage may directly alter reinforcement learning, as already suggested by others (Swain et al., 2011; McDougle et al., 2016; Miall and Galea, 2016).

Our viewpoint is supported by recent anatomical studies showing bidirectional connections between the cerebellum and dopaminergic-basal ganglia routes (Bostan et al., 2010; Chen et al., 2014; Bostan and Strick, 2018). Specifically, the dentate nucleus of the

cerebellum sends disynaptic projections to the striatum (Hoshi et al., 2005), and to midbrain dopaminergic structures (Watabe-Uchida et al., 2012). Conversely, the cerebellar cortex receives disynaptic projections from the subthalamic nucleus (Bostan et al., 2010). Functionally, recent works in rodents provide evidence that reward expectation modulates the firing rate of cerebellar cells (Ohmae and Medina, 2015; Wagner et al., 2017; Heffley et al., 2018). In the same vein, neuroimaging studies in humans have reported activity related to RPEs in the cerebellum (O'Doherty, 2004; Ramnani et al., 2004; Seymour et al., 2004; Tanaka et al., 2004; Tobler et al., 2006; Garrison et al., 2013), suggesting that this structure is functionally involved in processing reward feedback. These works are in agreement with the result of a previous study showing that cerebellar patients exhibit altered reinforcement learning in a decision-making task requiring very simple movements (Thoma et al., 2008). In line with these considerations, structural and functional alterations of the cerebellum were found in individuals suffering from an addiction, such as alcohol or cocaine dependence, a condition characterized by abnormal reward processing (Moulton et al., 2014; Miquel et al., 2016; Moreno-Rius and Miquel, 2017). Furthermore, an important feature of reinforcement-based compared to error-based adaptation is that the former increases trial-to-trial movement variability, reflecting an exploration process of the environment (Izawa and Shadmehr, 2011; Taylor and Ivry, 2014; Dhawale and Smith, 2017). Following this idea, modelling work in the present study showed that healthy subjects increased motor exploration following unrewarded compared to rewarded trials. This effect was absent in the patients reflecting an inability to modulate behavior optimally according to reward feedback. In this view, a recent study showed that poor performance in a visuomotor adaptation task in cerebellar patients is not only due to impaired error-based learning but also to a difficulty in using feedback information to develop and maintain an explicit aiming strategy (Butcher et al., 2017). Hence, it seems that cerebellar dysfunction could have impaired the ability to learn both from errors and reward feedbacks.

Nevertheless, an important point that needs to be raised here is the age difference between the healthy subjects tested in the commented paper (Therrien et al., 2018) and the cerebellar patients to which they are compared but that were originally tested in (Therrien et al., 2016). As such, the healthy subjects (25.0 +/- 4.8 years old) were much younger than the patients (61.5 +/- 10.0 years old) and in fact, when the groups were matched for age in Therrien et al. (2016), the healthy (older) controls also exhibited impaired motor exploration,

to a comparable extent as the patients. This suggests that ageing could also have contributed to the reduced reinforcement learning abilities of the patients (Chowdhury et al., 2013). Further studies are therefore required to examine the respective contribution of ageing and cerebellar dysfunction to reinforcement learning.

In conclusion, the work by Therrien and colleagues provides new insights into the influence of motor noise on reinforcement learning in healthy subjects and in patients suffering from cerebellar impairment. Moreover, the data are also consistent with the view that the cerebellum may be directly involved in reinforcement learning and more precisely in reinforcement-based motor learning. Future studies could directly test this hypothesis by relating the reinforcement learning impairment of patients to their score at the International Cooperative Ataxia Rating Scale, reflecting the severity of the cerebellar impairment. This line of research opens very interesting perspectives to design innovative multi-approach neurorehabilitation strategies.

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. *Curr Biol* 21:557–562 Available at: <http://dx.doi.org/10.1016/j.cub.2011.02.030>.
- Bhanpuri NH, Okamura AM, Bastian AJ (2013) Predictive Modeling by the Cerebellum Improves Proprioception. 33:14301–14306.
- Bostan AC, Dum RP, Strick PL (2010) The basal ganglia communicate with the cerebellum.
- Bostan AC, Strick PL (2018) The basal ganglia and the cerebellum: Nodes in an integrated network. *Nat Rev Neurosci* 19:338–350 Available at: <http://dx.doi.org/10.1038/s41583-018-0002-7>.
- Butcher PA, Ivry RB, Kuo S, Rydz D, Krakauer JW, Taylor JA (2017) Higher Neural Functions and Behavior The cerebellum does more than sensory prediction error-based learning in sensorimotor adaptation tasks. :1622–1636.
- Chen CH, Fremont R, Arteaga-Bracho EE, Khodakhah K (2014) Short latency cerebellar modulation of the basal ganglia. *Nat Neurosci* 17:1767–1775 Available at: <http://dx.doi.org/10.1038/nn.3868>.
- Chowdhury R, Guitart-Masip M, Lambert C, Dayan P, Huys Q, Düzel E, Dolan RJ (2013) Dopamine restores reward prediction errors in old age. *Nat Neurosci* 16:648–653.
- Dayan E, Averbach BB, Richmond BJ, Cohen LG (2011) Stochastic reinforcement benefits skill acquisition. :140–143.
- Derosiere G, Vassiliadis P, Demaret S, Zénon A, Duque J (2017a) Learning stage-dependent effect of M1 disruption on value-based motor decisions. *Neuroimage* Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1053811917307164>.
- Derosiere G, Zénon A, Alamia A, Duque J (2017b) Primary motor cortex contributes to the implementation of implicit value-based rules during motor decisions. *Neuroimage* 146:1115–1127 Available at: <http://dx.doi.org/10.1016/j.neuroimage.2016.10.010>.
- Dhawale AK, Smith MA (2017) The Role of Variability in Motor Learning.
- Doya K (2000) Complementary roles of basal ganglia and cerebellum in learning and motor control Kenji Doya. :732–739.
- Galea JM, Mallia E, Rothwell J, Diedrichsen J (2015) The dissociable effects of punishment and reward on motor learning. *Nat Neurosci* 18:597–602 Available at: <http://dx.doi.org/10.1038/nn.3956%5Cn10.1038/nn.3956%5Cnhttp://www.nature.com/neuro/journal/v18/n4/abs/nn.3956.html#supplementary-information>.

Garrison J, Erdeniz B, Done J (2013) Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neurosci Biobehav Rev* 37:1297–1310 Available at: <http://dx.doi.org/10.1016/j.neubiorev.2013.03.023>.

Gershman SJ, Daw ND (2017) Reinforcement Learning and Episodic Memory in Humans and Animals : An Integrative Framework.

Heffley W, Song EY, Xu Z, Taylor BN, Hughes MA, McKinney A, Joshua M, Hull C (2018) Coordinated cerebellar climbing fiber activity signals learned sensorimotor predictions. *Nat Neurosci* 21:1431–1441 Available at: <http://dx.doi.org/10.1038/s41593-018-0228-8>.

Herzfeld DJ, Kojima Y, Soetedjo R, Shadmehr R (2018) Encoding of error and learning to correct that error by the Purkinje cells of the cerebellum. *Nat Neurosci* 21:736–743 Available at: <http://dx.doi.org/10.1038/s41593-018-0136-y>.

Hoshi E, Tremblay L, Féger J, Carras PL, Strick PL (2005) The cerebellum communicates with the basal ganglia. *Nat Neurosci* 8:1491–1493.

Izawa J, Shadmehr R (2011) Learning from sensory and reward prediction errors during motor adaptation. *PLoS Comput Biol* 7:1–12.

Kim HE, Morehead JR, Parvin DE, Moazzezi R, Ivry RB (2018) Invariant errors reveal limitations in motor correction rather than constraints on error sensitivity. *Commun Biol* 1:19 Available at: <http://www.nature.com/articles/s42003-018-0021-y>.

Lee D, Seo H, Jung M (2012) Neural Basis of Reinforcement Learning and Decision Making. *Annu Rev Neurosci* 35:287–308.

Mawase F, Uehara S, Bastian AJ, Celnik P (2017) Motor Learning Enhances Use-Dependent Plasticity. *J Neurosci* 37:2673–2685 Available at: <http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.3303-16.2017>.

McDougle SD, Boggess MJ, Crossley MJ, Parvin D, Ivry RB, Taylor JA (2016) Credit assignment in movement-dependent reinforcement learning. *Proc Natl Acad Sci* 113:6797–6802 Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1523669113>.

Miall RC, Galea J (2016) Cerebellar damage limits reinforcement learning. :4–7.

Miall RC, King D (2008) State estimation in the cerebellum. *Cerebellum* 7:572–576.

Miquel M, Vazquez-Sanroman D, Carbo-Gas M, Gil-Miravet I, Sanchis-Segura C, Carulli D, Manzo J, Coria-Avila GA (2016) Have we been ignoring the elephant in the room? Seven arguments for considering the cerebellum as part of addiction circuitry. *Neurosci Biobehav Rev* 60:1–11 Available at: <http://dx.doi.org/10.1016/j.neubiorev.2015.11.005>.

Moreno-Rius J, Miquel M (2017) The cerebellum in drug craving. *Drug Alcohol Depend* 173:151–158
Available at: <http://dx.doi.org/10.1016/j.drugalcdep.2016.12.028>.

Moulton EA, Elman I, Becerra LR, Goldstein RZ, Borsook D (2014) The cerebellum and addiction: Insights gained from neuroimaging research. *Addict Biol* 19:317–331.

Nikooyan AA, Ahmed AA, Nikooyan AA, Ahmed AA (2015) Reward feedback accelerates motor learning
Reward feedback accelerates motor learning. :633–646.

O'Doherty JP (2004) Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Curr Opin Neurobiol* 14:769–776.

O'Doherty JP, Cockburn J, Pauli WM (2017) Learning, Reward, and Decision Making. *Annu Rev Psychol* 68:73–100 Available at: <http://www.annualreviews.org/doi/10.1146/annurev-psych-010416-044216>.

Ohmae S, Medina JF (2015) Climbing fibers encode a temporal-difference prediction error during cerebellar learning in mice. *Nat Neurosci* 29:1–7.

Quattrocchi G, Greenwood R, Rothwell JC, Galea JM, Bestmann S (2017) Reward and punishment enhance motor adaptation in stroke. *J Neurol Neurosurg Psychiatry:jnnp-2016-314728* Available at:
<http://jnnp.bmj.com/lookup/doi/10.1136/jnnp-2016-314728>.

Ramnani N, Elliott R, Athwal BS, Passingham RE (2004) Prediction error for free monetary reward in the human prefrontal cortex. *Neuroimage* 23:777–786.

Rinderknecht MD, Lambercy O, Raible V, Büsching I, Sehle A, Liepert J, Gassert R (2018) Reliability, validity, and clinical feasibility of a rapid and objective assessment of post-stroke deficits in hand proprioception. *J Neuroeng Rehabil* 15:1–15.

Roemmich RT, Bastian AJ (2018) Closing the Loop: From Motor Neuroscience to Neurorehabilitation. *Annu Rev Neurosci* 41:415–429 Available at: <https://www.annualreviews.org/doi/10.1146/annurev-neuro-080317-062245>.

Schlerf J, Ivry RB (2012) Encoding of Sensory Prediction Errors in the Human Cerebellum. 32:4913–4922.

Schultz W (2015) Neuronal Reward and Decision Signals: From Theories to Data. *Physiol Rev* 95:853–951
Available at: <http://physrev.physiology.org/lookup/doi/10.1152/physrev.00023.2014>.

Seymour B, O'Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS (2004) ID 1627 letters to nature Temporal difference models describe higher-order learning in humans. *Nature* 429:664–667.

Shadmehr R (2017) Learning to Predict and Control the Physics of Our Movements. *J Neurosci* 37:1663–1671 Available at: <http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.1675-16.2016>.

Shadmehr R (2018) Motor Learning: A Cortical System for Adaptive Motor Control. *Curr Biol* 28:R793–R795 Available at: <https://doi.org/10.1016/j.cub.2018.05.071>.

Shadmehr R, Smith MA, Krakauer JW (2010) Error Correction , Sensory Prediction , and Adaptation in Motor Control.

Song Y, Smiley-Oyen AL (2017) Probability differently modulating the effects of reward and punishment on visuomotor adaptation. *Exp Brain Res* 235:3605–3618.

Swain RA, Kerr AL, Thompson RF (2011) The Cerebellum: A Neural System for the Study of Reinforcement Learning. *Front Behav Neurosci* 5:1–6 Available at: <http://journal.frontiersin.org/article/10.3389/fnbeh.2011.00008/abstract>.

Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S (2004) Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. 7:1–2.

Taylor JA, Ivry RB (2014) Cerebellar and Prefrontal Cortex Contributions to Adaptation, Strategies, and Reinforcement Learning, 1st ed. Elsevier B.V. Available at: <http://dx.doi.org/10.1016/B978-0-444-63356-9.00009-1>.

Therrien AS, Wolpert DM, Bastian AJ (2016) Effective Reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. *Brain* 139:101–114.

Therrien AS, Wolpert DM, Bastian AJ (2018) Increasing Motor Noise Impairs Reinforcement Learning in Healthy Individuals. *Eneuro* 5:ENEURO.0050-18.2018 Available at: <http://eneuro.sfn.org/lookup/doi/10.1523/ENEURO.0050-18.2018>.

Thoma P, Bellebaum C, Koch B, Schwarz M, Daum I (2008) The cerebellum is involved in reward-based reversal learning. *Cerebellum* 7:433–443.

Tobler PN, Doherty JPO, Dolan RJ, Schultz W, Philippe N, Doherty JPO, Dolan RJ (2006) Human Neural Learning Depends on Reward Prediction Errors in the Blocking Paradigm. :301–310.

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

Uehara S, Mawase F, Celnik P (2017) Learning Similar Actions by Reinforcement or Sensory-Prediction Errors Rely on Distinct Physiological Mechanisms. *Cereb Cortex*:1–13 Available at: <http://academic.oup.com/cercor/article/doi/10.1093/cercor/bhx214/4157545/Learning-Similar-Actions-by-Reinforcement-or>.

Wagner MJ, Kim TH, Savall J, Schnitzer MJ, Luo L (2017) Cerebellar granule cells encode the expectation of reward. *Nature* 544:96–100 Available at: <http://dx.doi.org/10.1038/nature21726>.

Watabe-Uchida M, Zhu L, Ogawa SK, Vamanrao A, Uchida N (2012) Whole-Brain Mapping of Direct Inputs to Midbrain Dopamine Neurons. *Neuron* 74:858–873 Available at: <http://dx.doi.org/10.1016/j.neuron.2012.03.017>.

Weeks HM, Therrien AS, Bastian AJ (2017) Proprioceptive Localization Deficits in People With Cerebellar Damage. *Cerebellum* 16:427–437.

Weeks HM, Therrien AS, Bastian AJ (2018) The cerebellum contributes to proprioception during motion. :693–702.

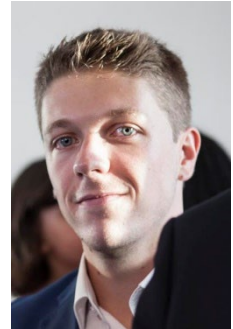
Wilkinson L, Steel A, Mooshagian E, Zimmermann T, Keisler A, Lewis JD, Wassermann EM (2015) Online feedback enhances early consolidation of motor sequence learning and reverses recall deficit from transcranial stimulation of motor cortex. *Cortex* 71:134–147.

Wolpert DM, Diedrichsen J, Flanagan JR (2011) Principles of sensorimotor learning. *Nat Rev Neurosci* 12 Available at: <http://dx.doi.org/10.1038/nrn3112>.

Curriculum Vitae

Pierre Vassiliadis

Born March 28th, 1993 in Brussels
Nationality: French
Contact: Neuro-X Institute, Campus Biotech
Chemin des Mines 9
1202 Geneva, Switzerland
pierre.vassiliadis@gmail.com



Current position

Joint PhD Student in Neuroscience at the **Ecole Polytechnique Fédérale de Lausanne** (*EPFL*, Lausanne, Switzerland ; April 2020-present) and the **Université Catholique de Louvain** (*UCLouvain*, Brussels, Belgium ; October 2018-March 2020).

Supervisors: Prof. Friedhelm Hummel (*EPFL*) and Prof. Julie Duqué (*UCLouvain*)
Co-supervisor: Dr. Gerard Derosiere (*UCLouvain*).

Education

2018: Master Degree in Medicine (MD) with High Honors (*UCLouvain*, Brussels, Belgium).

2018: Accepted for a specialization in Neurology (*UCLouvain*/St Luc University Hospital)

2014-2018: Medical Research Student in Neuroscience at IoNS, (*UCLouvain*, Brussels, Belgium). Supervisor: Prof. Julie Duqué; Co-supervisor: Dr. Gerard Derosiere.

Master thesis title: *Exploring the functional role of M1 in reward-based decision-making and reinforcement learning: a TMS and behavioural project.*

2014: Bachelor Degree in Medicine with Honors (*UCLouvain*, Brussels, Belgium).

2011: French Baccalaureate with Highest Honors (*Lycée Français Jean Monnet de Bruxelles*, Brussels, Belgium).

Languages: French: Mother tongue.
English: Fluent.
Greek: Fluent.
German: Level B1.

Computer Softwares and Languages: Matlab (extensive), R (basic), Statistica (extensive), Microsoft Office (extensive), Inkscape (extensive).

Other experiences

2020: Training certificate in Good Clinical Practice (TRREE training program in research ethics evaluation)

2017-2018: Medical Internships in Neurology (*Cliniques Universitaires Saint-Luc, CHU Mont-Godinne, Grand Hopital de Charleroi*; total of 6 months).

2017: Medical Internships (*UCLouvain*, 12 months).

2017: Medical Exchange Internship at *Centre Hospitalier Universitaire de Montréal*, Montreal, Canada (2 months).

2014: Instructor Certificate in Basic Life Support/Automated External Defibrillator (BLS/AED).

2012: Certificate in BLS/AED.

Publications

Published:

- (1) **Vassiliadis P**, Lete A, Duque J, Derosiere G (2022). Reward timing matters in motor learning. *iScience*, 25: 104290.
- (2) Bigoni C, Cadic-Melchior A, **Vassiliadis P**, Morishita T, Hummel FC (2022). An Automatized Method to Determine Latencies of Motor-Evoked Potentials under physiological and pathophysiological conditions. *Journal of Neural Engineering*, 19 024002.
- (3) **Vassiliadis P**, Derosiere G, Dubuc C, Lete A, Crevecœur C, Hummel FC, Duque J (2021). Reward boosts reinforcement-based motor learning. *iScience*, 24:102821.
- (4) **Vassiliadis P**, Grandjean J, Derosiere G, Duque J (2020). Motor training strengthens corticospinal suppression during movement preparation. *Journal of Neurophysiology*, 124: 1656–1666.
- (5) **Vassiliadis P**, Derosiere G (2020) Selecting and executing actions for rewards. *The Journal of Neuroscience*, 40 (34) 6474-6476.
- (6) Derosiere G, **Vassiliadis P**, Duque J (2020). Advanced TMS approaches to probe corticospinal excitability during action preparation. *NeuroImage*, 213:116746.
- (7) **Vassiliadis P**, Derosiere G, Duque J (2019). Beyond motor noise: considering other causes of impaired reinforcement learning in cerebellar patients. *eNeuro*, 6, 1.

- (8) **Vassiliadis P**, Grandjean J, Derosiere G, de Wilde Y, Quemener L, Duque J (2018). Using a double-coil TMS protocol to assess preparatory inhibition bilaterally. *Frontiers in Neuroscience*, 12, 139.
- (9) Grandjean J, Derosiere G, **Vassiliadis P**, Quemener L, de Wilde Y, Duque J (2018). Towards assessing corticospinal excitability bilaterally: Validation of a double-coil TMS method. *Journal of Neuroscience Methods*, 293:162–168.
- (10) Derosiere G, **Vassiliadis P**, Demaret S, Zénon A, Duque J. (2017). Learning stage-dependent effect of M1 disruption on value-based motor decisions. *Neuroimage*, 162:173-185.

Under review:

- (1) **Vassiliadis P**, Beanato E, Popa T, Windel F, Morishita T, Neufeld E, Duque J, Derosiere G, Wessel MJ, Hummel FC (2022). Non-invasive stimulation of the human striatum disrupts reinforcement learning of motor skills (biorXiv: <https://www.biorxiv.org/content/10.1101/2022.11.07.515477v1>).
- (2) Wessel MJ*, Beanato E*, Popa T, Windel F, **Vassiliadis P**, Menoud P, Beliaeva V, Violante IR, Dzialecka P, Park CH, Maceira-Elvira P, Morishita T, Cassara A, Steiner M, Grossman N, Neufeld E, Hummel FC (2022). LTP-like noninvasive striatal brain stimulation enhances striatal activity and motor skill learning in humans (biorXiv: <https://www.biorxiv.org/content/10.1101/2022.10.28.514204v1>).
- (3) Neige C, Ali Zazou A, **Vassiliadis P**, Lebon F, Brees T, Derosiere G (2022). Probing the influence of SMA and vmPFC on the motor system with dual-site transcranial magnetic stimulation. (biorxiv: <https://doi.org/10.1101/2022.01.18.476729>)

In preparation:

- (1) **Vassiliadis P**, Pinheiro D, Shokur S, Morishita T, Micera S, Hummel FC (2022). Sensory-dependent benefit of reinforcement feedback on motor control and learning.
- (2) **Vassiliadis P**, Fleury L, Morishita T, Hummel FC (2022). Reward sensitivity modulates reinforcement learning of motor skills in stroke patients.
- (3) Lete A*, **Vassiliadis P***, Hummel FC, Derosiere G, Duque J (2022). The effect of reward on motor plasticity during motor learning. *: *equal contribution*
- (4) Beanato E*, Moon HJ*, Wessel MJ, Popa T, Windel F, Menoud P, **Vassiliadis P**, Gauthier B, De Falco E, Alania K, Violante I, Dzialecka P, Neufeld E., Grossman N, Hummel FC*, Blanke O* (2022). Spatial navigation modulation via hippocampal temporal interference stimulation.
- (5) Popa T*, Beanato E*, Wessel MJ, Menoud P, Windel F, **Vassiliadis P**, Alania K, Violante I, Dzialecka P, Neufeld E., Grossman N, Hummel FC (2022). Temporal interference effects on hippocampal activity and associative memory.

Oral communications

Conferences

- (1) **Vassiliadis P** (2021). Sensory-dependent benefit of reinforcement feedback on motor control. Annual Meeting of the NeuroLeman Network 2021 (NLN'21). Geneva, Switzerland.
- (2) **Vassiliadis P** (2018) Contribution of the primary motor cortex to action value encoding during motor decisions. *Invited symposium* at the Belgian Association for Psychological sciences (BAPS). Gent, Belgium.

Posters

- (1) **Vassiliadis P**, Beanato E, Windel F, Popa T, Morishita T, Neufeld E, Duque J, Derosiere G, Wessel MJ, Hummel FC. Non-invasive temporal interference stimulation of the human striatum at 80 Hz, but not 20 Hz, disrupts reinforcement motor learning. (2022). Neural Control of Movement (NCM). Dublin, Ireland.
- (2) Beanato E*, Wessel MJ*, Popa T, Windel F, **Vassiliadis P**, Menoud P, Beliaeva V, Violante IR, Dzialecka P, Park CH, Maceira-Elvira P, Morishita T, Cassara A, Steiner M, Grossman N, Neufeld E, Hummel FC (2022). Modulatory effects of temporal interference stimulation on striatal activity. 32nd International Congress of Clinical Neurophysiology (ICCN). Geneva, Switzerland.
- (3) **Vassiliadis P**, Lete A, Derosiere G, Hummel FC, Duque J. The effect of reward on motor plasticity during motor learning. 4th International Brain Stimulation Conference 2021.
- (4) Wessel MJ, Beanato E, Popa T, Windel F, Menoud P, Beliaeva V, **Vassiliadis P**, Grossman N, Neufeld E, Hummel FC. Evidence for temporal interference (TI) stimulation effects on motor striatum. 4th International Brain Stimulation Conference 2021.
- (5) **Vassiliadis P**, Derosiere G, Dubuc C, Lete A, Crevecoeur C, Hummel FC, Duque J (2020). The added value of motivation for motor learning. Neural Control of Movement (NCM) - poster accepted; event canceled.
- (6) **Vassiliadis P**, Grandjean J, Derosiere G and Duque J (2019). Training-related modulation of preparatory activity in the motor system. Society for Neuroscience (SfN). Chicago, USA.
- (7) **Vassiliadis P**, Derosiere G, Dubuc C, Crevecoeur C and Duque J (2019). Impact of reinforcement on action selection, initiation and execution during motor skill learning. Neural Control of Movement (NCM). Toyama, Japan.

Also presented at:

Belgian Society for Neuroscience (BSN). Brussels, Belgium.

Symposium on the "Biology of Decision Making" (SBDM). Oxford, United Kingdom.

- (8) **Vassiliadis P**, Grandjean J, Derosiere G, De Wilde Y, Quemener L and Duque J (2018). Probing preparatory inhibition bilaterally with double-coil TMS. Belgian Brain Congress 2018. Liège, Belgium.
- (9) Grandjean J, Derosiere G, **Vassiliadis P**, De Wilde Y, Quemener L and Duque J (2018). Validation of a new double-coil TMS method to assess corticospinal excitability bilaterally Belgian Brain Congress 2018. Liège, Belgium.

- (10) Grandjean J, Derosiere G, **Vassiliadis P**, de Wilde Y, Quemener L, Duque J. (2017) Validation of a double-coil TMS method to assess corticospinal excitability. 2nd International Brain Stimulation Conference. Barcelona, Spain.
- (11) Grandjean J, Derosiere G, **Vassiliadis P**, de Wilde Y, Quemener L, Duque J. (2017) Investigating the reliability of a double-coil TMS method to assess corticospinal excitability bilaterally. 27th Neural Control of Movement (NCM) Meeting. Dublin, Ireland.
- (12) Derosiere G, Demaret S, **Vassiliadis P**, Zenon A, Duque J. (2016) Implicit encoding of action values by the human primary motor cortex. 38th International Symposium of the Groupe de recherche sur le système nerveux central (GRSNC). Montreal, Canada.
- (13) Derosiere G, **Vassiliadis P**, Demaret S, Zenon A, Duque J. (2015) Disentangling the involvement of primary motor cortex to value-based reinforcement learning and value-based decision-making. 45th Annual meeting of the Society for Neuroscience. Chicago, USA.

Supervision and Teaching

- Supervision

Aegryan Lété (Master in Medicine, *UCLouvain*, Sep 2019 – June 2023)
 Chiara Galletti (Master in Physics, *EPFL*, Sep 2021-Dec 2021)
 Cécile Dubuc (Master in Biomedical Sciences, *UCLouvain*, Feb 2019 - Aug 2020)
 Wanda Materne (Master in Physiotherapy, *UCLouvain*, Jul 2019 – June 2020)

- Teaching

Feb 2022-June 2022:

- Fundamentals of Neuroengineering with Prof. Silvestro Micera (45 hours as teaching assistant, *EPFL*, Lausanne, Switzerland)
- **Teaching Assistant Award** of the School of Engineering for the course Fundamentals of Neuroengineering 2022 (*EPFL*)

Feb 2021-June 2021:

- Fundamentals of Neuroengineering with Prof. Silvestro Micera (90 hours as teaching assistant, *EPFL*, Lausanne, Switzerland)

Oct 2018- February 2020:

- Motor rehabilitation and kinesiology seminars, 3rd year, Physiotherapy (2x3 hours; *UCLouvain*, Brussels, Belgium)
- Neuroscience and neuro-related pathologies, 5th year, Physiotherapy (2x4 hours; *UCLouvain*, Brussels, Belgium).
- In-depth questions of Cognitive Neuroscience, 5th year, Biomedical Sciences (2x4 hours; *UCLouvain*, Brussels, Belgium).

2014-2017:

- Instructor in Cardio-Pulmonary resuscitation (BLS/AED), 3rd year, Medicine (30 hours, *UCLouvain*, Brussels, Belgium)

Science communication

June 2022 : « Récompenser pour apprendre : oui mais quand ? » ; UCLouvain website : <https://uclouvain.be/fr/sites/woluwe/actualites/recompenser-pour-apprendre-oui-mais-quand.html>

November 2021 : « La motivation améliore l'apprentissage moteur »; article in *Le Journal du Médecin* (4 novembre 2021 ; N°2691)

November 2021 : « La motivation détermine l'apprentissage moteur »; Daily Science : <https://dailyscience.be/04/11/2021/la-motivation-determine-lapprentissage-moteur/>

October 2021 : « Le rôle de la récompense pendant l'apprentissage moteur » ; FNRS.news: <https://www.frs-fnrs.be/fr/communication/fnrs-news>

July 2021 : « L'impact déterminant de la motivation sur l'apprentissage moteur » ; UCLouvain website : <https://uclouvain.be/fr/chercher/actualites/l-impact-determinant-de-la-motivation-sur-lapprentissage-moteur.html>

July 2020 : Videos of presentation (Gustave Boël - Sofina mobility fellowship):

https://vimeo.com/479436446?embedded=true&source=vimeo_logo&owner=41376553

<https://www.facebook.com/frs.fnrs/videos/fnrsawards-pierre-vassiliadis-boursier-friafnrs-uclouvain/902673103578635/>

Fellowships

2018-2022: Doctoral Research Fellowship, Belgian National Fund for Scientific Research (FNRS/FRIA, 4 years, ± 120 k€)

2022: Belgian FNRS travel grant (0.4 k€)

2022: Wallonia Brussels International World mobility fellowship (WBI.World, 3 months [Oct-Dec], ± 6 k€)

2021-2022: Wallonia Brussels International World mobility fellowship (WBI.World, 1 year, ± 25 k€)

2020-2021: Wallonia Brussels International World mobility fellowship (WBI.World, 1 year, ± 25 k€)

2020-2021: Gustave Boël - Sofina mobility fellowship (1 year, ± 20 k€)

2020: *UCLouvain* doctoral mobility grant (6 months, 4 k€)

2018-2020: Doctoral Research Fellowship, Belgian Special Research Fund (FSR; 2 years, ± 60 k€, declined).

2019: 3 Belgian FNRS travel grants (1.7 k€)

Reviewing for ISI-indexed journals

Journal of Neuroscience (n=2), Journal of Neurophysiology (n=1), Scientific reports (n=1), Experimental Brain Research (n=2).