The effects of reward on upper limb coarticulation

by

Sebastian Sporn

A thesis submitted
to the University of Birmingham
for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology
University of Birmingham
March 2021

Abstract

Coarticulation has been studied in speech production for over 100 years and more recently has been observed in upper limb movement sequences. Coarticulation of upper limb movements has been shown to underlie our remarkable ability to produce movement sequences such as drinking a cup of coffee or driving a car reliably, smoothly and fast. Crucially, coarticulation breaks down in Parkinson's disease and stroke patients which significantly affects their daily lives. However, how humans learn to coarticulate and specifically how coarticulation can be facilitated is unknown. Reward is a powerful tool to influence human behaviour and over the past decade research on how reward shapes learning has widely increased. Importantly, reward has been shown to enhance execution of simple movements while converging evidence found that reward can improve learning of complex motor skills. The work in this thesis explores whether reward can enhance coarticulation. To address this question a novel motor task is introduced which involves a complex sequence of reaching movements (chapter 2). Across a series of behavioural experiments (chapter 3) we demonstrate that reward facilitates coarticulation which also led to improvements in movements smoothness and speed. Importantly, these improvements were retained over a long period without reward. Extending these findings, we used a neuropharmacological modulation (chapter 4) to investigate whether dopamine plays a role in coarticulation. The results suggest that a dopamine antagonist impairs coarticulation which can be compensated for with reward-based feedback. Finally, we show that coarticulation is guided by the principal of maximum smoothness which explains why coarticulation is constrained by the transition angle between reaching movements (chapter 5).

Acknowledgements

Three and a half years in and one never-ending pandemic still ahead, I have arrived at that once so-far-away-believed stage where I write the acknowledgements section. Exhausted but grateful I now have the chance to thank the people that have accompanied me – unwilling or not –on this journey.

First and foremost, I want to thank my supervisor, Dr Joseph Galea, who allowed me to explore while also keeping me on track. I remember him during my first year, when I told him (for potentially the 100th time) that I will just try this or that again and bemusement was soon replaced by a suspicious fear that here is someone who needs deadlines (and right he was!). His insight, knowledge and especially his help during difficult submission times was invaluable throughout this PhD. So, thank you Joe for your guidance and support – I am not sure this PhD would have been written otherwise.

Thanks to my colleagues Pete, Oli, and Xiuli with whom I shared an office and a big part of my daily life before we all reverted to Zoom. You have inspired this work and have answered so many of my questions with patience and insight. Most importantly, you have made me feel less like an imposter and accepted me as a member in this amazing lab. I am looking forward to catch-ups soon on the green. Oh wait you've all gone! Well has to be a Swedish, Canadian, UK zoom meet then (preferably my time in the morning, ha).

Who would have thought that you can make such good friends in Birmingham?! I did and I can't thank you enough for being there and living this PhD drama series with me (all publishing rights are mine by the way!). To my fellow PhD friends AL (slightly bigger you see), Consu, Lea, Sean and Tom– thanks for your emotional support and for turning grey days into gold. Not only have you embraced me, but you have given me confidence that indeed this can work out. To Sean who has bought a first-class ticket (you probably think for too many years by now) for this one-man show: you've been my flatmate, PhD fellow and close friend. I cannot thank you enough for this time. It made this time very special and yes you will hear from me soon (what's for breakfast?).

To my friends in London! Not only have you been with me during the times when I created latte art for customers that tipped too little, but you have also been there during those last days when I submitted my master's thesis. You inspire me greatly and your friendship is invaluable to me. Naya, Chiara and Lori remember the time in Greece? This is where our future lies – let those tomatoes grow while we forge collaborations and create good science. I adore you and I am blessed with your friendship. To Jacqui, Nels, Yi and Jordan – the froleagues. Meeting you when I was unsure was one of the most affirming times in my life. Not only have you given me the confidence to be but you have also guided, informed and expanded my world and I cannot thank you more. Let's pray the the Fat Walrus will reopen soon so we can have some pints (standard park walks otherwise).

An meine Familie, nun ist das Ende dieses Doktors erreicht und um ehrlich zu sein, ohne Euch waere das nicht moeglich gewesen (Ich weiss ich hab hier keine Umlauttaste von daher sieht alles ein bisschen bloed aus). Es gibt so viele Momente im Leben, in denen ihr mir Vertrauen haettet versagen koennen (oh ist das korrekt? Ich glaub das wird nichts ich schreibs um)… Euer Vertrauen und eure Liebe hat mich hierher gebracht und ich danke Euch aus ganzem Herzen. Ich freu mich jetzt schon wenn wir alle mal wieder zusammen sind – palavern!

To you, Christopher, this PhD is for you. You have shown me a world that I thought was reserved for others. Still overwhelmed by how fortunate I am that you are there, always there, rooting for me despite me thinking it's pointless to go on. I love you.

I love you.

To you.

Content

List of Figures

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Tables

Chapter 5

Chapter 1

Introduction

1.1 Motivation for this thesis

Movements are central to how an agent (i.e., humans and animals) successfully interacts with the environment and can be used to achieve daily life (e.g., making a cup of tea) and other complex (e.g., executing a tennis serve) goals. However, controlling goal-directed movements in an ever-changing environment is tremendously difficult. Additionally, often these goals (e.g., making a cup of tea) require the agent to perform a sequence of movements in order to realise them (e.g., fill the kettle with water, source the tea bag etc). Consequently, even simple goals can pose a challenge to the motor system and even sophisticated robots still fail to complete a variety of human daily life tasks. Yet, humans have the remarkable ability to complete such movement sequences with reliability, smoothness and speed; an ability which breaks down in clinical populations such as Parkinson's disease and stroke patients which severely impacts their daily life.

The present work focuses on our ability to perform movement sequences skilfully and aims at identifying strategies that aid efficient execution. Specifically, how reward affects both the learning and execution of movement sequences represents a core part of the present work. Reward has been a major focus of research across multiple disciplines. For instance, episodic memories that are paired with a rewarding stimulus (i.e., money or enjoyable food) rather than a neutral stimulus are encoded more strongly and remembered better[1]. Similarly, it has been demonstrated that reward serves as a central source in shaping optimal behaviour during decision-making[2]. Considering its pivotal role in shaping both animal and human behaviour, research on the effects of reward on both motor learning and adaptation has been a focus within the field of motor control and motor neuroscience. Particularly, its potentially beneficial effects within rehabilitation settings has been of great interests in recent years[3–5]. Consequently, here the effects of reward on movement sequences are investigated.

1.2 Motor skill learning

Motor skill learning refers to the process of learning and refining new skills such as mastering a tennis serve through the act of repeated practise[6–8]. Within this context, *skill* is defined as learning to execute complex movements (i.e., movement sequences) to a level of proficiency that allows for achieving the task goal reliably, efficiently and fast[9]. Consequently, motor skill learning can be distinguished from other forms of motor learning such as motor adaptation. Motor adaptation primarily describes the adjustment of already existing motor commands to perturbations in the environment[10,11]. In contrast, during motor skill learning motor commands have to be formed and refined de novo (i.e., from 'scratch')[12]. Consequently, motor skill learning is characterised by a slower learning process that typically involves several training sessions. However, once a skill has been mastered it is retained with little decay for a long period of time[13–17].

1.3 Motor sequence learning and coarticulation

Many motor skills require an agent to execute a series of movements with examples ranging from playing the piano to touch typing, to tasks of daily living such as making a cup of coffee. In each case, skill learning refers to the process of executing a series of movements with increasing spatial and temporal accuracy as well as overall movement quality (i.e., smoothness)[12]. Many clinical populations show impairments in sequence learning, which has detrimental effects on their daily living. Stroke survivors decompose movements into jerky sub-movements with available data suggesting that stroke severity and sequence learning ability are correlated[18–20]. Thus, there is an urgent need to understand how movement sequences are learnt in order to inform the development of clinical rehabilitation programmes. In recent years, research has provided valuable insight into different aspects of motor sequence learning with coarticulation of movement sequences being the focus of this thesis.

Ubiquitous in our daily life but often impaired in clinical populations [21–24], coarticulation describes the process of blending together a series of distinct movements into a single continuous action. For example, with learning compound movements such as drinking a cup of coffee are gradually blended together to form a continuous action that is executed with increased speed, smoothness and energetic-efficiency [25–28]. Therefore, coarticulation represents a hallmark of motor sequence learning. Crucially, coarticulation breaks down in stroke[23,24] and PD patients[29], which severely affects their daily life. However, despite its prevalence and importance to characterising impairments in clinical population, how humans learn to coarticulate and specifically how coarticulation can be facilitated is unknown. The work in this thesis explores whether reward can enhance coarticulation (see Chapter 2 for more information on coarticulation).

1.3.1 Implicit and explicit sequence learning

First observed in amnesic patients such as HM[30], it has been shown that motor sequences can be learnt both implicitly and explicitly[31]. While implicit motor learning occurs without knowledge of the task structure and the learned motor sequence, explicit learning requires higher levels of active attention[32–34]. Consequently, these fundamental concepts have been differentiated mainly based on the level of awareness an agent has of the learned skill[35]. Typical daily life examples of implicit learning include learning how to ride a bicycle and to walk as a child; motor skills that can often not be verbalised (procedural). In contrast, learning a dance routine or how to cook a new dish are examples of explicit learning, which requires knowledge of the action sequence (declarative) and attentional resources. On a neural level, explicit learning is believed to be integrated by the hippocampus and its surrounding brain structures, whereas implicit learning is thought be controlled by brain structures such as the basal ganglia[36]. However, delineating precisely what systems underlie both implicit and explicit learning still represents a major research challenge[37]. Additionally, in most real-world situations both implicit and explicit sequence learning contributes to skill learning[38], which further complicates dissociation. For example, when learning how to make a cup of coffee, one must learn the sequence of movements (knowledge) as well as the motor commands to execute the actions skilfully (practise). In line with this approach, research has shown that both explicit knowledge of task elements in combination with implicit features of movement constitute for skilfully executing a movement sequence[39]. Consequently, sequence learning can also be compartmentalised as learning *what* to do (explicit learning) and *how* to do it (implicit learning), with the latter being mainly driven by extensive practise[38].

1.3.2 Components of learning

Within motor sequence learning, spatial-sequential components (what to do) have been dissociated from its motor control components (how to do it)[38–41] allowing for a functional extension of the implicit-explicit framework of sequence learning. The spatial-sequential component refers to the order of the movement sequence in space and time (e.g., when and where to press the next button when entering your PIN). In contrast, the motor control component encompasses the movement dynamics and more specifically the sensorimotor integration of the individual movements (e.g., how fast and with how much pressure you press the button)[38–41]. Consequently, in most cases learning of motor sequences comprises the explicit identification of the movement sequence and the implicit acquisition of skilfully executing the sequence[39]. There is growing evidence that these components are learnt in parallel but have been shown to follow different time courses with different underlying brain networks mediating learning[6,39,42]. Further evidence comes from research showing that motor sequences can be learned implicitly[43–45], which highlights that different mechanisms support explicit and implicit learning. The typical time-course of motor skill learning will be reviewed in the following section.

1.3.3 Stages of sequence learning

There is a growing body of evidence that suggest that motor sequence learning can be segmented into 3 distinct stages, which was first proposed by Fitts and Posner (1967)[34]. These stages differ with regards to the amount of learning taking place and which brain areas and networks contribute to the learning process. Albeit the nomenclature differing across authors and time, current theoretical models of sequence learning propose a segmentation into an early, intermediate and autonomous phase[6–8,38,39,42,47–49].

Phase I: Early stage

The early phase (or sometimes referred to as cognitive or initial stage) is characterised by rapid improvements in the spatial-sequential component of motor sequence learning. Specifically, during this phase the temporal order of the sequence as well as the spatial location of each movement element is learnt. This allows for successful completion of the given movement sequence, albeit slowly, inefficiently and variable[6]. Despite these rapid improvements in successfully completing the given motor sequence[40,50], performance is still heavily reliant on attentional resources and sensory feedback to guide execution[6]. The corticostriatal loop (also called the 'associative' circuit) is particularly engaged during early sequence learning. Both the dorsolateral prefrontal cortex (DLPFC) and the posterior parietal cortex (PPC) are part of the corticostriatal loop and are mainly engaged during explicit learning which is attentionally demanding and relies on working memory[51,52]. Lesions to the PFC in monkeys have shown to impair sequence learning, highlighting the role of frontal brain regions to early sequence learning[53]. Additionally, there is growing evidence that the cerebellum plays a pivotal role in fine-tuning performance through extensive feedback loops with cortical structures such as prefrontal, parietal and premotor brain areas[54,55]. Here, feedforward mechanisms allow for predictive control via integration of current sensory inputs and future motor commands (for a comprehensive review on the role of the cerebellum during motor learning see [56,57]). Predictive control is crucial for navigation and forecasting future body states, especially during early learning when movements tend to be more erroneous[58]. Therefore, both the corticostriatal and cortico-cerebellar loop allow for rapid learning and successful sequence completion during early learning when the task space is relatively unexplored.

Phase II: Intermediate stage

However, with learning, performance stabilises and only incremental improvements in performance are observed across time. During this second phase, which is also called the consolidation period, the sensorimotor map becomes stronger and is gradually stored in the long-term memory[6]. This allows for a faster retrieval while also having to rely less on sensory feedback, which leads to performance becoming less susceptible to interference[6,8,59]. On a neural basis, a decrease in activity in the cerebellar-cortical network can be observed. In contrast increased activity can be found in the striatal-cortical network when sequence learning transitions from the early to the consolidation phase[8,48]. Specifically, the BG appears to update behaviour via links to premotor and motor areas and has been shown to play a central role in concatenating movement elements either through chunking or coarticulation[60]. This results in improvements in performance which are learnt slowly over time.

Phase III: Autonomous stage

The learning curve eventually reaches a plateau, once performance is near optimal. In this final stage, which has also been called retention or advanced stage, performance is almost automatic requiring little active attention (implicit)[49]. There has been considerable research into the concept of automaticity to describe skilled behaviour. Particularly, whether it represents a habit[31,61,62]. Habitual responses can be distinguished from goal-directed action selection which are mostly observed during early motor learning. Given prior knowledge and the task environment at hand, a goal-directed action is chosen which yields the best possible outcome. In contrast, habitual choices of actions are purely based on what has been successful in the past (i.e., what always yielded the best outcome). While goal-directed action choices are computationally heavy and therefore slow, habitual responses require little computation and are significantly faster[31,61,62]. However, this comes at the cost of inflexibility manifesting itself in choices that no longer yield the best outcome. This feature of habitual choices becoming maladaptive has been investigated extensively in relation to addiction[63]. Yet, habitual actions can be overridden by goal-directed choices and it has been assumed that these two systems act in parallel competing for action selection control[62]. Recent work could show that habitual responses are expressed when participants were forced to respond rapidly. In contrast, with more preparation time goal-directed actions were chosen which led to improvements in task success[62]. Therefore, it appears that the likelihood of expressing habits is increased if there is too little time to evaluate better options.

Traditionally, it has been assumed that motor areas such as the primarily motor cortex (M1), premotor cortex (PMC) and the PPC are of central importance during this final stage. While the PMC is involved in the integration of sensorimotor and visual information, M1 appears to store use-dependent motor maps to further optimise performance [38,47,57]. However, recent work found no evidence of M1 storing a motor map of a learnt discrete button-pressing task[64]. Instead, activity in M1 was related to the execution of sequence components (i.e., button presses), while the PMC appeared to store a representation of the sequence[64].

1.3.4 Skill in motor sequence learning

Speed, accuracy and reaction times

Motor skills have been traditionally assessed at the level of changes in temporal and spatial accuracy, which can be traced back to the well-established Fitts' law[65]. This universal principle states that the time to execute a movement is a function of the distance to the target divided by the target size. Consequently, longer distances and smaller target sizes lead to longer movement times. The rationale behind smaller target sizes leading to longer movement times is that, given a constant reaching speed and target distance, the probability of hitting smaller targets accurately decreases due to the inherent noise within the motor system[65]. Therefore, the reaching speed will have to be adjusted to ensure hitting the target accurately and avoiding errors, which led to the formulation of the speed-accuracy trade-off[65,66]. In many sports[67], daily life tasks and experimental settings, the speed-accuracy trade-off naturally limits performance[12]. Skill learning then is related to breaking this natural limit via improving the speed-accuracy trade-off[68,69]. Consistently, it has been demonstrated that repeated practise leads to improvements in the speed-accuracy function and that these improvements can continue for weeks and even years[70,71]. Similarly, it has been shown that reward can shift and even break this trade-off by simultaneously improving both speed and accuracy[72–75]. Thus, there has been a pronounced emphasis on measuring changes in speed and accuracy (or precision) to capture motor skill learning. In line with this approach, reaction times have similarly been employed as a proxy measure for skill[76,77]. In contrast to speed or total movement time, reaction times capture the response time between a go signal and movement initiation. As such reaction times are a valuable tool to measure automaticity, interference and habit formation using task such as the serial reaction time task[76]. Recent work could show that reaction times are driven by both movement preparation and movement initiation[78]. This suggest that apart from the neural delays in initiating an action, reaction times also represent a preparation process during which an appropriate action is selected[78]. Consistently it has been demonstrated that reductions in reaction times are mainly driven by decreases in preparation time through repeated practise[79] and explicit knowledge[80], with chached or habitual responses requiring the least amount of preparation time[81].

Going beyond: movement variability and smoothness

Over the last few decades, there has been a pronounced interest in capturing motor skill based on the notion of movement efficiency, which relates to concepts such as movement variability[82–84] and smoothness[85–87]. Traditionally, movement variability has been deemed an unwanted by-product of the inherent noise in the motor system[88]. This motor noise has been shown to permeate from every level of the sensorimotor system (i.e., integration and perception of sensory input, generation and initiation of motor responses etc.)[89]. Considering that many daily life tasks have precision requirements, there has been a focus on how the sensorimotor system minimises this inherent noise[90]. Consequently, successful skill learning, and motor expertise has been linked to decreases in movement variability. However, recently research has shown that variability, especially during early learning, provides a source to explore the task space to potentially find better solutions to complete a given task[82,84,88,91,92]. Indeed, research has shown that increases in variability in task-relevant dimensions during early learning led to steeper learning curves and a subsequent decrease in variability (exploitation) across training[82,84,92]. Hence, research has aimed to disentangle motor noise from variability that is driven by exploration[93,94]. Recent findings were able to demonstrate that motor learning can be formalised as a process of sequential decisions that are adjusted for the given motor noise in the system[93]. Furthermore, once a skill is mastered, variability is constrained within goal-relevant dimensions only[94]. This suggests that movement variability is controlled differently depending on the stage of motor learning and the structure of the task. In addition to movement variability, smoothness has been identified as a marker of skilled performance[85–87]. Movement smoothness has been shown to improve during infancy[95], motor learning[27,96] and over the course of stroke recovery[23,24] and is believed to represent the result of effort minimisation[90,97]. As such, movement smoothness might indicate other important aspects of sensorimotor control such as spatio-temporal coordination[85]. However, despite its prevalence and importance to efficiency via effort minimisation, there has been considerable debate how to best measure or quantify motion smoothness[85]. Generally, a movement is considered as smooth if it is executed without prominent interruptions. Consequently, the derivatives of the positional data (x, y) are most commonly used to assess motion smoothness. Jerk measures (3rd derivative) to quantify smoothness are an established approach[98], which is based on the seminal finding that minimum jerk trajectories capture and correspond well to reaching movements in healthy participants[99,100]. However, jerk metrics are very sensitive to noise and may therefore, be less informative in clinical settings[85]. Hence, aspects of the velocity profile have been used to capture smoothness in clinical populations including metrics such as number of peaks, mean velocity and number of intervals with velocities close to 0 cm/s[23]. Both increases in smoothness and decreases in goal-related variability have been related to skill learning, and provide evidence that motor skill learning leads to improvements in efficiency and overall movement quality.

1.4 Motor control and motor sequence learning

It has been long established that repeated practise leads to improvements in speed, accuracy and movement quality. However, how the motor system achieves this feat is less understood. For skill learning to occur, the motor system must deal with several complex problems ranging from motor redundancy and noise to uncertainty due to the volatility of the environment. The following section will provide a brief overview of some these problems. Subsequently two computational models will be presented that were successful at providing a mechanistic framework for how the motor system allows for skill learning.

Motor redundancy and noise

The human body consists of over 600 muscles that control more than 200 joints making the motor system very flexible yet difficult to control considering the very large number of degrees of freedom that come with it[101]. This renders the motor system redundant because there are multiple, and in some cases infinite, ways to execute the same action[101,102]. For example, when executing a simple reaching movement to reach a specific target, an infinite number of trajectories with altogether different speed profiles can lead to task completion. Similarly, a single trajectory can be executed using various joint angles and therefore muscle configurations. Consequently, the motor system is faced with a decision-making problem in terms of having to choose one solution from an abundance of possible solutions. Despite this sheer abundance, human reaching movements appear to follow highly stereotyped patterns (i.e., even single reaching movements are slightly curved). Within sensorimotor control, understanding how and why the motor system chooses one particular solution over all other has been a major focus.

Similarly, how the motor system deals with the inherent noise that is present at all stages of the control hierarchy has been a longstanding question within motor neuroscience[89]. As described in the previous section, motor noise affects the planning and execution of movements and appears to scale with the magnitude of the movement, which has been termed signal-dependent noise[103,104]. How the motor system accounts for this signal-dependent noise is a crucial question and evidence suggests that the nature of this variability might be based on the size principle of the given motor unit[104]. Dealing with noise is further complicated by the fact that we effectively live in the past which is driven by delays in the sensorimotor system[101]. Depending on the sensory modality and the complexity of the sensory input, sensory processing can be delayed by up to 150 ms[101].

Uncertainty and nonstationary

Uncertainty is present both in the environment and due to noise and delays within the motor system itself. Particularly, if the environment is unstable[105] or not fully predictable[106], control of movements can be difficult. Imagine you are asked to execute a series of reaching movements to complete a task, in which force-fields are applied throughout that differ in direction and intensity. If these force-fields were stable across time, initially an agent would have to respond to the force-fields on the fly. However, with training the agent would be able to execute the sequence more skilfully and with less error corrections. In contrast, if the properties of these force-fields randomly change from trial to trial (i.e., the environment is unstable and difficult to predict), the agent can rely less on prior experience and has to adapt to the new situation on each trial. How the motor system learns to respond to this uncertainty and what computations drive possible solutions to this problem has been a fundamental field of research in recent years. Similarly, imagine you were to take part in the same experiment in your 20s and again in your 60s. Your body will have changed over this time manifesting itself in longer delays[107], decreases in muscle strength[108] and sensory acuity[109]. Hence, the motor system is nonstationary, which is also present on shorter timescales through muscle fatigue[101].

Over the past few decades research on sensorimotor control and motor skill learning has made substantial progress at identifying computations that allow the motor system to alleviate these problems.

1.4.1 Kinematic models

Initially, research focused on the problem of redundancy and how agents choose one motor command over an infinite number of other possible motor commands. This is particularly striking considering that despite this sheer abundance, human reaching movements appear to follow highly stereotyped patterns[110]. Based on the observation that such over-trained reaching movements tend to be highly smooth and fluid[100], the criterion of maximum smoothness was introduced[99,100]. This allows to differentiate distinct yet equally available motor commands according to how smooth they are. As such, a scalar value can be associated with each way of executing a particular reaching movement depending on how smooth the solution is. And consequently, the most optimal way of completing the task is to choose the smoothest motor command. Therefore, here the aim of the motor system is to optimise the cost of jerkiness. Indeed, the minimum jerk model[99,100] has been shown to account for a variety of reaching movements[27,99,100,111,112]. However, despite its accuracy in predicting simple and sequential reaching movements across a variety of 2D and 3D reaching movement tasks, its biological feasibility has been challenged[101]. It remains an open question whether the motor system actually cares about movement smoothness and even if it does how it would then compute it. Jerk is the third derivative of positional data which has to be integrated over the whole movement in order to evaluate smoothness of a given movement, which is by no means a trivial computation. Additionally, the minimum jerk model is limited in its generalisability and application to other motor tasks such as a tennis serve.

1.4.2 Optimal feedback control

Extending on the idea of optimality, computational models have been developed in which optimisation is driven by relevant task demands[113,114]. Hence, in contrast to minimising jerk exclusively, the cost function in these models can include other task relevant dimensions, such as accuracy[101,113,114]. In contrast to minimising jerk, increasing accuracy is a natural variable that the motor system should care about considering that it will ultimately ensure task success[101]. Additionally, compared to jerk this cost can easily be measured by estimating the difference between end position and target location and, therefore, might be biologically more feasible. Furthermore, research has shown that such optimal control models can deal with the redundancy in the motor system, because the noise stemming from the muscular system affects accuracy which the model aims to minimise. Thus, the inherent level of noise is accounted for by an optimal control model[115]. Lastly, optimal control models are substantially more generalisable than kinematic models such as the minimum-jerk and minimum-torque model, because the cost function can be a mix of multiple task relevant dimensions[101,113,114].

Importantly, optimal control models were recently combined with feedback control which represents a crucial extension of the optimal control framework. Specifically, by integrating sensory feedback in the framework, the optimal solution to a given task is now represented by a control policy[113]. Depending on a set of rules, the control policy aims to complete the task while minimising a mixed cost function (i.e., overall motor effort). Indeed, research has shown that optimal feedback control models (OFC) can accurately predict behaviour across a broad range of motor tasks including eye[116,117], arm[118] and wrist[119] movements, posture and has even been applied to jumping[120]. While both optimal control models and OFC models can predict human behaviour, only OFC models can account for changes in behaviour due to perturbations (via sensory feedback)[121–123]. Hence, optimal control and particularly, OFC represents an important and highly influential framework to explore human motor behaviour. However, despite its success in modelling human movements OFC models rely on a number of pre-specifications. Its versatility and generalisability via minimising a mix of costs (e.g. accuracy, energy) comes with the problem of having to determine the relative weighting of these costs[101]. This free parameter is often fitted to the empirical data. Furthermore, in its current form OFC models cannot predict execution of long movement sequences.

1.4.3 Reinforcement learning

Reinforcement learning (RL) is an important and widely established model in the field of decision-making and in contrast to OFC models requires substantially less pre-specifications[101]. It centres on the idea that an agent learns from rewarding feedback[124]. Specifically, it describes a learning process during which actions that lead to reward are valued more and subsequently expressed more likely. In contrast, non-rewarded or punished actions are valued less and are executed less frequently. Consequently, over the course of learning an agent will update its belief of how rewarding each action from the set of available options is based on the received feedback[124]. In theory, this learning process will increase the likelihood of maximising future rewards, which is mediated by the exploration-exploitation trade-off[1,2]. Two classes of algorithms have been successful at modelling learning from rewards and have been covered extensively within the RL literature[127]. Firstly, model-free algorithms learn the value of an action when it is chosen and executed. However, here this choice does not affect the value of other options in the set. Thus, in order to determine the value of actions they have to be chosen and cannot be inferred[2,128]. Additionally, model-free algorithms only store the net reward of an action. This means that in multi-step decision-making tasks the agent does not store information about how rewarding the individual decisions were. Therefore, there is no structural understanding of the task and the agent simply tracks which action sequence was rewarded the most[127,128]. This family of algorithms is computationally light, because no structural understanding of the task or model is needed (hence model-free) to complete the task. Despite this reduced need to store complex transition histories, which significantly reduces the amount of information to be computed, this computationally light architecture comes at the cost of inflexibility[127,129,130]. If the reward probabilities change (i.e., a formerly rewarded option is now not being rewarded anymore), the algorithm will continue choosing the formerly rewarding option even though it is not optimal anymore. Only after the option has been devalued, the agent will choose other options and find a new optimal. Considering that choosing one option does not update the value of the other options in the set, model-free algorithms tend to be slow, susceptible to environmental change and inflexible. However, they might be biologically more plausible considering the lightness of their computational architecture[127,129,130].

In contrast, model-based algorithms store a model of the environment, which means that here an agent updates the value of options even when not executed. Therefore, executing one action leads to changes in the probability of choosing other options. In addition, model-based algorithms do not only store the net value of actions but build value relationships between all states. This makes them very resilient to environmental changes because changes in the reward probabilities can be accounted for. However, this also means that here the computational architecture is very heavy or expensive, which will further increase with task difficulty. As such it has been argued that model-based algorithms despite being more flexible and resilient to change are potentially biologically implausible[127,129,130].

This dichotomy between model-free and model-based algorithms in RL has been well established in decision-making, whereas RL in motor skill and motor sequence learning is a more recent development. Furthermore, a growing body of work suggests that model-free RL relies on implicit learning processes, whereas model-based RL requires working memory and is thus considered to be more explicit[131–133].

1.4.4 Reinforcement learning and dopamine

RL has been successful in modelling and predicting animal and human behaviour across a broad range of tasks (for reviews see [128,134,135]). Consequently, there has been much interest and research into the neural underpinnings of RL. The neuromodulator dopamine (DA) is believed to be central to RL, which has been demonstrated in both animal[135–138] and humans[1,2]. Specifically, in RL learning is driven by prediction errors and the phasic bursting of DA in the midbrain has been shown to be a putative neurobiological substrate of this signal[1]. Mechanistic accounts posit that reward prediction errors drive reward-based learning and represent the difference between expected reward and received reward[124]. This difference is used to update the value of actions. Consequently, actions that lead to positive prediction errors were more rewarding than expected and hence should be valued more. Similarly, actions that lead to negative prediction errors should be valued less. In a now seminal study Schultz (1986) could show that when primates found food (reward) a strong phasic burst of DA neuronal firing occurred[142]. This increase in bursting was not observed when the monkey found non-food items[142]. Moreover, research has demonstrated that when highly probable rewards did not occur, DA neurons in the midbrain show a phasic decrease in firing[143]. This is consistent with a negative prediction error. Consequently, phasic firing of DA neurons in response to reward appear to explain well how action values are updated during RL[128].

In addition, dopamine projections are found from the midbrain to the striatum, which is on of the main input areas of the BG[144]. The striatum is part of the cortico-basal ganglia-thalamo-cortical loop, which consists of modulatory dopaminergic projections to the cortex[145]. It is assumed that dopaminergic reward prediction error signals regulate synaptic plasticity in the cortex[146,147]. Hence, positive reward prediction errors are believed to increase synaptic plasticity through and increase in phasic DA bursting and this is thought to contribute to a strengthening of associations leading to reward[138].

Therefore, there is a mounting evidence that DA subserves the computations in RL, with clinical studies lending further support. Parkinson's disease (PD) is a degenerative disorder that is characterised by a severe depletion of dopaminergic neurons in the striatum. PD patients suffer from motor impairments which involve symptoms such as bradykinesia, rigidity and tremor. Hence, dopamine appears to play a role not only in supporting RL but also in motor control[128]. Assessing RL in PD patients provides valuable insights into its neural underpinnings and specifically the role of DA, by comparing task performance ON and OFF medication. An influential study assessed how well PD patients learn from positive and negative reinforcement in a probabilistic learning task. They found that patients ON medication learnt better from positive reinforcement, while negative reinforcement led to poorer learning. In contrast, patients OFF medication learnt better from negative reinforcement and less from positive. The authors embedded their findings in a Go-NoGO model based on the assumed working mechanisms of the basal ganglia[148]. In their Go-NoGO model, the direct and indirect pathways from the striatum are proposed to underlie positive and negative RL, respectively. Specifically, D1 receptors in the direct (Go) pathway are activated when a reward prediction error is positive, whereas D2 receptors In the indirect (NoGo) pathway are active with a negative reward prediction error. The authors concluded that higher dopamine levels during ON medication led to more D1 binding biasing activation towards the direct pathway and as a result improved positive RL. Conversely, lower levels of dopamine during OFF medication led to less binding and biased activation towards the indirect pathway, which improved negative RL.

Despite indicating a strong support for DA underpinning RL, recent studies were unable to replicate these findings[149–151]. These studies could not find any effect of medication during the learning phase but found impairments in performance during consolidation[149,150]. Specifically, patients OFF medication showed impaired performance during retrieval, which has been explained by the lack of dopamine (OFF medication) leading to decreases in synaptic plasticity thereby impairing memory consolidation[149,150].

1.4.5 Reward, dopamine and motor sequence learning

RL has been influential within fields of decision-making, however its role in motor skill learning and more specifically motor sequence learning has been a more recent development. It is important to bear in mind that motor sequence learning includes learning the order of a motor sequence (i.e., what to do when and where) and the kinematic dynamics of the individual movements (i.e., how to do it). In the literature these two components are often referred to as action selection and action execution and reward has been shown to act on both.

The effect of reward on action selection and execution was first shown in simple motor tasks. Early results come from studies employing saccadic eye movement tasks. Capitalising on the speed-accuracy trade-off and its implications for human and animal behaviour, these studies operationalise changes in reaction time (RT) or movement time (MT) as a proxy to assess reward-based effects on behaviour. Takikawa and colleagues (2001) could show that monkeys make faster saccadic eye movements to rewarded targets (juice) than to non-rewarded targets, while preserving similar levels of accuracy. Similar results were obtained using human participants[152–154] with Manohar et al. (2015) demonstrating that humans can perform rewarded saccades faster as well as more accurately[73]. These results indicate that reward can not only shift the speed-accuracy trade-off curve but break it by simultaneously invigorating speed and precision. They were able to partially replicate these results by comparing the effects of contingent and non-contingent rewards on the performance of saccadic eye movements. Their findings imply that both contingency and reward rate increased movement vigour as expressed in higher peak velocities of rewarded saccades[155]. More recently, several studies attempted to translate the effects of reward on saccadic movements to simple arm movements. Assessing the RTs of wrist movements in trials with varying amounts of reward available, Opris and colleges (2011) demonstrated that in monkeys, RTs were the fastest when the upcoming trial was cued as a high reward trial. In addition, they could show that reward predictability modulated changes in movement vigour[156], a result that complements the findings of reward contingency and saccadic peak velocity in humans (see also[157]). A potential explanation could be that and agent is more willing to produce an action with a higher metabolic cost (i.e., effort) associated with it, if the reward is high and can be reliably anticipated. Expanding on the saccadic one-direction-rewarded task by Takikawa et al. (2001), Summerside and colleagues (2018) adapted this task to include arm reaching movements. They asked participants to produce centre-out reaching movements to four different targets while one of them was rewarded. In line with the original findings, participants reduced RTs and MTs alongside exhibiting higher peak velocities in reaching movements that were paired with reward[74]. In summary, reward has been consistently shown to improve performance on simple tasks. However, the literature is far more limited assessing the effects of reward on sequential actions.

Using a serial reaction time task (SRTT), Wächter et al. (2009) could show that retention of reward-based procedural learning was increased through reward. However, they failed to observe any improvements in the execution of the individual button presses. Consequently, they suggested that reward is beneficial to implicit motor sequence learning, but has no effect on the execution of movements[158]. In line with these results an influential study by Abe et al. (2011) could demonstrate that monetary reward led to improved retention and offline gains in a force tracking task (FTT). Similarly to Wächter et al. (2009), they did not observe improvements in performance during training, but found that reward led to performance being maintained which was found even 30 days after the initial training. In contrast, work by Steel et al. (2016) did not find any reward-based improvements during learning or retention using both a SRTT and a FTT. However, they found that punishment led to improvements in the execution of the individual button presses in the SRTT[159]. Furthermore, a study by Dayan et al. (2014) demonstrated that highly stochastic reward feedback improved both motor skill learning, enhancing both online and offline gains using a sequential visual isometric task. Consequently, the results on the effect of reward on motor sequence learning and execution is far more heterogeneous than in simple motor tasks. Differences in task designs and reward feedback have been attributed to this lack of consensus.

However, considering the strong link between reward and dopamine, which has been shown to underlie learning through reward prediction errors, recent research has sought to determine the role of dopamine during motor sequence learning. Using a SRTT task, Boyd et al. (2009) showed that stroke patients with a damage to the basal ganglia exhibited severe impairment in motor sequence learning. Across a series of blocks, performance only marginally improved suggesting that the basal ganglia is involved in motor sequence learning[20]. Similar results come from a study investigating explicit and implicit sequence learning in PD patients using a probabilistic SRT task. Their findings demonstrate that PD patients were impaired in both implicit and explicit sequence learning[43]. Converging with these results, Jin et al. (2014) found that in rodents, neural activity in the basal ganglia encoded an entire action sequences as a single motor command. Specifically, they demonstrated that both the indirect and direct pathway was involved in action initiation but showed differential patterns of activity during sequence execution. These findings suggest that the basal ganglia, which is connected to the cortex via dopaminergic projections is involved in both sequence learning and its execution[60,160]. Further support comes from a series of studies by Shin et al. (2018, 2020) in which they could show that DA neurons in both the direct and indirect pathway of the striatum were involved during both reward-based learning and motor control[161,162]. These findings highlight that DA is involved in both RL and motor sequence learning, whilst further challenging the antagonistic dichotomy between the direct and indirect pathway within RL as proposed by Frank et al. (2006)[148]. In an effort to assess the role of dopamine during motor sequence learning, recent neuropharmacological studies have used DA agonists and antagonists to selectively block or enhance D1 and D2 receptor binding. Trembley et al. (2009) used raclopride, a selective D2 antagonist, to investigate the role of D2 receptor binding to motor sequence learning in monkeys. They found that raclopride impaired consolidation of a new motor sequence. Importantly, when raclopride was administered after a motor sequence was already learnt, no impairments in consolidation were observed[163]. These findings align with the behavioural results of Wächter et al. (2009) and Abe et al. (2010) and suggest that D2 receptor binding play a role in consolidating motor skills. In a subsequent study, Trembley et al. (2010) investigated the role of dopamine in movement chunking by assessing PD patients ON and OFF medication. In short, chunking refers to discrete actions that are temporally aligned with very short reaction times in between them. Hence, chunking is considered to represent efficient skill learning within discrete action sequence tasks (further information on chunking will be provided in the next section). The authors found that patients OFF medication showed severe impairment in chunking even for well-trained motor sequences. In comparison, patients ON medication did not differ significantly from healthy controls[164]. Converging results come from Seo et al. (2010) who also observed that patients OFF medication showed impairments in a motor sequence learning task using stochastic reward feedback[165]. Furthermore, Lissek et al. (2014)showed that learning on a complex motor sequence task was impaired following administration of tiapride, a highly selective D2 antagonist, hence mirroring previous work on the importance of D2 receptor binding during motor sequence learning[166]. Interestingly, a recent study by Augustin et al. (2020) showed the selective depletion of D2 receptors in the indirect pathway was also associated with impaired initiation and vigour of self-paced actions[167]. Therefore, D2 receptors appear to be involved in both motor sequence learning and response vigour. Further evidence from studies assessing response vigour found that PD patients OFF medication show a reduction in response vigour in both a saccadic eye task[168] and an effort-based decision-making task[169]. Recently, Niv et al. (2007) have suggested that changes in tonic dopamine underpins the determination of response vigour [170]. By extending the RL framework, they could show that on a computational level vigour represents the net rate of rewards and as such acts as the opportunity cost of time[170,171]. Hence, how much effort an agent should exert to obtain a future reward depends on the reward history and more specifically on the reward average.

To illustrate this, let us consider two rats that engage in a lever pressing task in which food pellets drop after four presses. One rat is hungry, while the other one is sated. This difference in satiety will affect the utility of receiving a food pellet which in turn will affect the reward magnitude. Hence, for a hungry rat receiving food is more rewarding which will increase the net reward rate. This will lead to increases in the opportunity cost of time thereby favouring faster responses. Niv et al. (2007) argue that this corresponds to the energising role of motivation on behaviour and suggest that tonic dopamine underpins this computation[170,171]. Indeed, tonic dopamine is suitable to represent the net reward rate because it reflects phasic DA firing across time[172]. Consequently, there is ample evidence that both tonic and phasic DA influence human behaviour and contribute to both skill learning and response vigour. However, despite these advances in understanding the role of reward, and more specifically DA in motor skill learning, there is currently no consensus whether reward enhances learning of complex sequential actions. Consequently, this thesis aims to investigate the effects of reward on movement vigour and particularly motor skill learning in complex sequential tasks.

1.5 Structure of the thesis

In chapter 2, I present the development of the continuous sequential reaching task (CSRT) which is then used in chapters 3 and 4. Chapter 3, investigates the effect of explicit reward on sequential reaching. While chapter 4 uses a neuropharmacological manipulation to assess the role of dopamine in underlying the relationship between reward and sequential reaching. Chapter 5 investigates how environmental (task) constraints influences how reward affects sequential reaching. Finally, chapter 6 discusses the impact of this work in relation to previous literature and introduces possible future directions.

Chapter 2

Task development

2.1 Task design: Background

The aim of this thesis was to investigate the effect of reward on complex sequential movements. In nature, sequential actions can be discrete or continuous. Discrete movements are characterised by a period without motion (i.e., zero velocity) before and after its execution, such as a single finger flexion to press a button[173]. In contrast, continuous movements lack such marked stop periods, and are often described as rhythmic in the literature to refer to periodically repeated actions such as walking or swimming[174]. The previously described serial reaction time task (SRTT) is often used to assess motor sequence learning on a discrete serial task. Here participants are asked to a press a button in response to a specific cue[76]. This task has been successfully used to assess both implicit and explicit motor sequence learning[175], while improvements are mainly seen at the level of action selection (i.e., reductions in response times). This means that over the course of practise, individual motor primitives (i.e., button presses) can be aligned temporally resulting in a reduction in response times. This will lead to in part temporally overlapping movements, which over time will be executed as a holistic unit[176] (i.e., chunk). However, the motor primitives underlying the formed chunk will remain stable and are not kinematically modified[176]. A prominent example of a continuous motor sequence task is the force tracking task (FTT)[177]. Here participants are asked to pinch a force transducer between thumb and index finger to move a cursor upward (increase of grip force) or downward (release). The aim is to follow a predetermined path as accurately as possible[177]. Hence, in this task individual movement elements do not have a marked end or start point and the underlying movement primitives can be modified to improve execution[177]. Recently, Codol et al. (2020) showed that reward had dissociable effects on action selection and execution in a simple reaching task[75]. The authors showed that movement time (i.e., the time from initiating the reaching movement until reaching the target) was decreased with reward at no cost to movement accuracy through an increase in arm stiffness. Additionally, during distractor trials, action selection was improved when the trial was rewarded. These findings align with previous research which has shown that reward improves execution via decreases in movement time across a broad range of simple motor tasks[73–75,155,178]. However, the effects of reward on more complex movements such as reaching movements is less explored. There is currently a lack in consensus in the literature[71,158,159,177]. For example, a recent study could not find that reward improved movement execution or retention on both a SRTT and FFT[159]. In line with these results, two studies using a SRTT and FFT respectively, could also show that reward did not improve execution via a reduction in MTs[158] or improve accuracy[177], respectively. However, both studies observed an increase in retention, suggesting that reward improved motor retention[158,177]. Consequently, in these studies reward appears to lead to enhanced retention during complex continuous tasks which is often not observed during simple discrete tasks[74,75,179]. Codol et al. (2020) offers an explanation suggesting that increases in metabolic cost through increases in arm stiffness are only viable as long as reward is available to pay for it. Consequently, once reward is removed performance reverts back to near baseline levels[75]. However, it is unclear why this does not seem to be the case during more complex continuous tasks. To address these inconsistencies, a novel task was developed in which participants were asked to execute a series of reaching movements.

2.1 Pilot studies

In two experiments, we sought to investigate the effects of reward on the execution of a sequence of reaching movements. Changes in performance were measured based on the time in which participants completed a given trial (movement time) and as such served as a proxy measure for reward-based movement invigoration. The following two sections will provide an overview of the design and the results from these experiments.

2.1.1 Experiment 1

Participants: 14 participants (6 males; age range 18 - 26) were recruited to participate in this experiment, which had been approved by the local research ethics committee of the University of Birmingham. All participants were novices to the task paradigm and were free of motor, visual and cognitive impairment. Most participants were self-reportedly right-handed (N = 3 left-handed participants) and gave written informed consent prior to the start of the experiment. For their participation, participants were remunerated with either course credits or money (£7.5/hour) and were able to earn additional money during the task depending on their performance. Depending on the experiment, participants were pseudo-randomly allocated to one of the available groups.

Experimental apparatus: All experiments were performed using a Polhemus 3SPACE Fastrak tracking device (Colchester, Vermont U.S.A; with a sampling rate of 110Hz). Participants were seated in front of the experimental apparatus which included a table, a horizontally placed mirror 25cm above the table and a screen. A low-latency Apple Cinema screen was placed 25cm above the mirror and displayed the workspace and participants' hand position (represented by a green cursor – diameter 1cm). On the table, participants were asked to perform 2-D reaching movements. Looking into the mirror, they were able to see the representation of their hand position reflected from the screen above. This setup effectively blocked their hand from sight. The experiment was run using MATLAB (The Mathworks, Natwick, MA), with Psychophysics Toolbox 3.

Task design: Participants were asked to produce 4 sequential reaching movements to targets displayed on the screen to complete a trial (Figure 1a). To investigate whether reward can invigorate movements selectively in a within-participant design, 4 movement sequences were created (Figure 1b). They were unique in their target arrangement but were of equal in length (15 cm) and total amount of curvature (540°) and therefore, had similar task demands (hereafter called 'shapes').

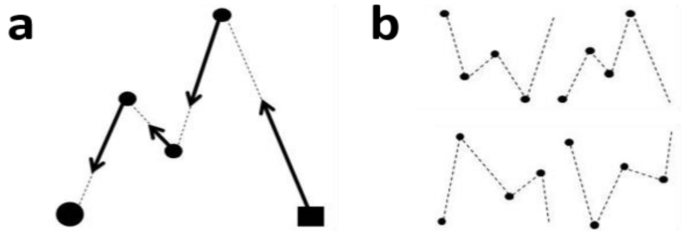


Figure 1 | Experimental design. (a) Schematic representation of a target arrangement. Participants were asked to perform reaching movements to pass through 4 targets to complete a trial. **(b)** In total, four target arrangements ('shapes') were presented. Note that the dotted lines in-between targets were not presented.

This was confirmed in a pilot study in which 20 participants were asked to complete 30 trials on each shape without receiving reward. Subsequent performance analysis revealed no differences in movement time between the four shapes.

A trial consisted of three elements: a preparation box (width 2cm) on the right side of the workspace, a start box (width 2cm) positioned to its left, and 4 targets (diameter 2cm, Figure 2a). Participants had to pass the cursor through the preparation box for the start box to appear and were instructed that once they moved into the start box, target appearance would be triggered. The aim of the task was to pass the cursor through all 4 targets without missing any of them. Participants had to repeat a trial if they exceeded 3 seconds in total movement time (MT) or missed a target, which enforced accuracy in the reaching movements. In this experiment participants were able to earn money (reward) depending on their performance, with faster MTs leading to higher rewards. By enforcing accuracy throughout the experiment, any shifts in the speed-accuracy trade-off could be explained by changes in MTs. Reward trials were cued using an auditory and visual stimulus prior to the start of the trial (Figure 2b). Once participants moved into the preparation box just to the right of the start box, they heard a sound (auditory stimulus) and the start box appeared in yellow (visual stimulus) rather than in black. In the no reward trials participants were instructed to complete the trial 'as fast and accurately as possible' (Figure 2c).

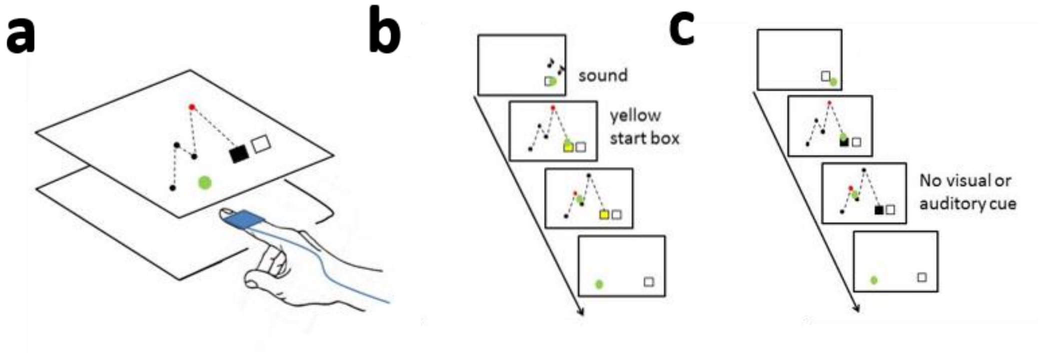


Figure 2 | Trial design. (a) Illustration of the task workspace including preparation box, start box and targets. **(b)** Illustration of a reward **(c)** and no reward trial. In reward trials once participants passed though the preparation box, an auditory and visual cue was triggered to indicate an upcoming reward trial. Note that the dotted lines in-between targets in **(a-c)** were not presented.

Experimental procedure: The experiment consisted out of 3 parts: baseline, training and post assessment and lasted for approximately 1 hour (Figure 3). Additionally, it included an initial learning phase prior to the start of the experiment for participants to be able to memorise the reaching sequences.

Baseline: Participants completed 2 trials on each of the four shapes (8 trials in total) during Baseline, which were presented in a random order during a block (i.e., 2 blocks of 4 trials with each block containing all shapes). These trials were used to assess whether there were any pre-training differences between groups. Participants were instructed to 'move as fast and accurately as possible', while no performance-based feedback was given at the end of each trial.

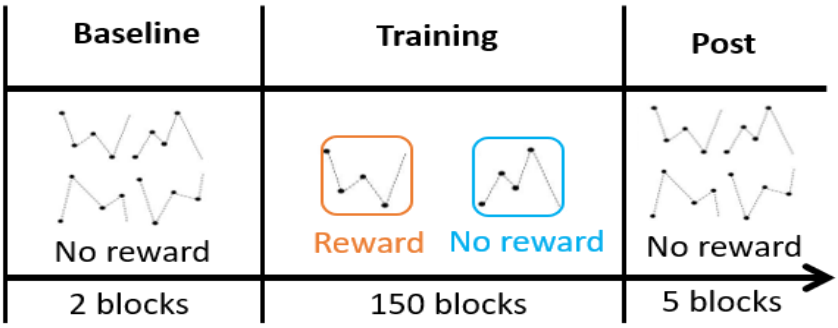


Figure 3| Design of Experiment. In each block all shapes were presented in a random order. During Training participants received reward depending on their movement time on one randomly chosen target arrangement. Another was only trained but was not reward.

Training: During training participants practised two shapes which were counterbalanced across participants (2 shapes x 150 trials = 150 blocks). Performance on one shape was rewarded (RS) and participants received monetary reward depending on how fast they completed the trial (see Reward structure and feedback for more information). No reward was available on the other shape (NS). The remaining two shapes were not trained at all and were only presented during a post assessment as a control (CS) (Figure 3).

Post assessments: Participants were asked to complete a post assessment (4 shapes x 5 trials = 5 blocks) during which no reward was available on any shape. Similarly, to the no reward trials during training, participants were asked to complete each trial as 'fast and accurately as possible'. Importantly, trials had to be repeated when participants missed a target, hence high levels of accuracy were enforced throughout.

Reward structure and feedback: (1) Reward trials were cued using a visual stimulus prior to the start of the trial. Once participants moved into the preparation box, the start box appeared in yellow (visual stimulus) rather than in black (Figure 2b). Participants were informed that faster MTs would earn them more money, with a maximum amount of 5p available in each trial. While participants moved from the start box to the preparation box to initiate a new trial, the amount earned in the previous trial was displayed on the top of the screen (i.e., 'You have earned 2p out of 5p'). We used a closed-loop design to calculate the amount of reward earned in each trial. To calculate this, we included the MT values of the last 20 trials and organised them from fastest to slowest to determine the rank of the current trial within the given array. A rank in the top three (<= 90%) returned a value of 5p, ranks >= 80% and <90% were valued at 4p; ranks >=60% and <80% were awarded 3p; ranks >=40% and < 60% earned 2p while 1p was awarded for ranks >=20% and < 40%. A rank in the bottom three (<20%) returned a value of 0p. When participants started a new experimental block, performance in the first trial was compared to the last 20 trials of the previously completed block. (2) No reward trials were not cued, and no reward was available for participants. However, participants were instructed to 'move as fast and accurately as possible'.

Data analysis: Analysis code is available on the Open Science Framework website, alongside the experimental datasets at: https://osf.io/62wcz/. The analyses were performed in Matlab (Mathworks, Natick, MA). To measure and compare performance between conditions Kruskal-Wallis and Wilcoxon single rank tests were employed, since most of the obtained data was non-parametric. For this analysis, median MTs of the first (early Training) and last (late training) 20 trials of training for each participant in both the reward and no reward condition were used. For Post-hoc analysis, Tukey's Test for single-step multiple comparisons were used.

Movement Time (MT): MT was the main outcome measure of this experiment and was measured as the time between exiting the start box and reaching the end target.

Results. MT performance is illustrated in Figure 4. Baseline performance did not differ significantly between shapes (Kruskal-Wallis test; p = 0.156), which suggests that all four shapes exhibit similar task demands; a result which is in accordance with the findings of the pilot study assessing performance on the same. With the start of Training, performance on the rewarded shape (RS) and non-rewarded shape (NS) improved significantly with respect to their Baseline MT (Wilcoxon test; $p < 0.0001$ for RS Baseline vs early Training and $p < 0.0063$ for NS Baseline vs early Training). Importantly, performance on RS was significantly faster during late training than on NS (Wilcoxon test; RS vs NS: $p = 0.0176$). These results align with previous findings on the effect of reward on simple reaching movements[74,75,179]. However, no differences in performance between RS and NS could be observed during post (Kruskal-Wallis test; $p = 0.4082$ for RS vs NS), which suggests that in this task reward did not improve retention of performance.

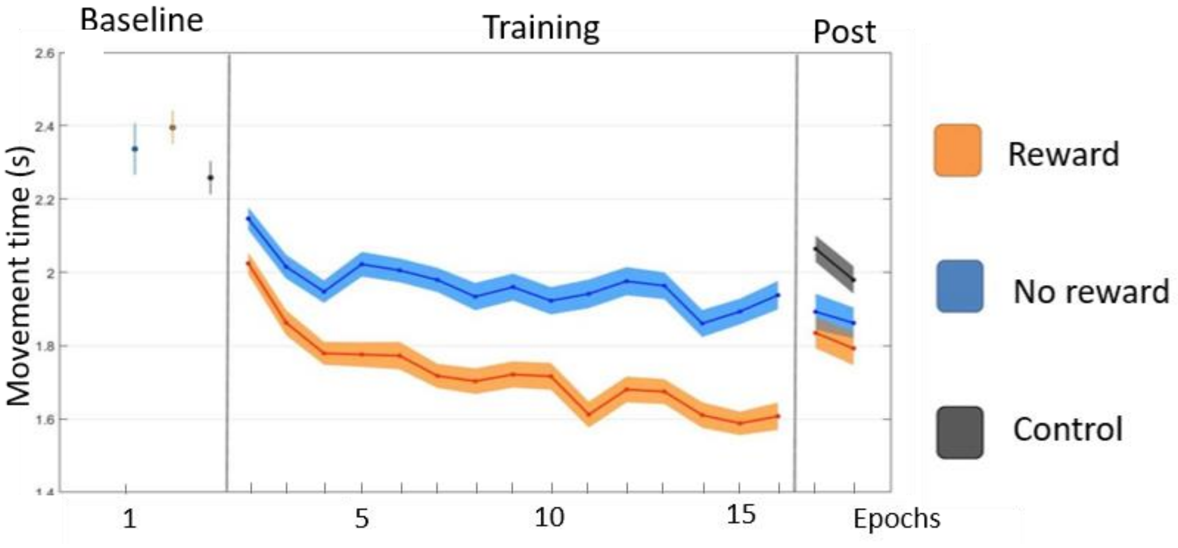


Figure 4 | Results Experiment 1. Movement times (MT) across participants for baseline, training and post assessment. Values are averaged across 2 trials in baseline and 10 trials in both training and post. Coloured lines represent group mean and shaded areas represent s.e.m.

Nevertheless, MT performance on both trained target arrangements (RS and NS) was faster than performance on the control shapes (Kruskal-Wallis test; $p < 0.0001$, Tukey's test $p < 0.0001$ for RS vs CS; $p = 0.0032$ for NS vs CS and $p = 0.4082$ for RS vs NS). Hence, these results are not consistent with previous findings that have shown reward-based improvements in retention for complex sequence tasks[158,177], while also indicating that task exposure via training improves performance.

2.1.2 Experiment 2

In a subsequent experiment, an additional testing day was added to the study design, while replicating experiment 1 on the first testing day. Previous work could show that reward led to improvements in retention in complex sequence tasks[158,177]. To investigate whether reward similar to these previous results improves retention of reward-based improvements in MTs in this task a second testing day was added to the design.

Participants: 14 participants (5 males; age range 18 - 24) were recruited to participate in the second experiment, which had been approved by the local research ethics committee of the University of Birmingham. All participants were novices to the task paradigm and were free of motor, visual and cognitive impairment. Most participants were self-reportedly right-handed (N = 4 left-handed participants) and gave written informed consent prior to the start of the experiment. For their participation, participants were remunerated with either course credits or money (£7.5/hour) and were able to earn additional money during the task depending on their performance. Depending on the experiment, participants were pseudo-randomly allocated to one of the available groups.

Experimental apparatus and task design: Both the experimental apparatus and task design were the same as in Experiment 1.

Experimental procedure: To assess possible retention effects, we split Experiment 2 over two consecutive assessment days (Figure 5). During Day 1 participants underwent the same experimental protocol as participants in Experiment 1. During Training 2 shapes were chosen to be trained while one was rewarded (RS) and the other was not (NS), which was counterbalanced across participants. As mentioned before, the shapes not chosen for RS and

NS, served as a control condition throughout the experiment (CS). Assessment on Day 2 was always scheduled 24 hours after testing on Day 1 and included a familiarisation phase prior to the start of the main experiment. During the second testing day, participants engaged in an additional post assessment (4 shapes x 5 trials = 5 blocks) and furthermore in three tests:

(1) shape that was previously rewarded during Training (Day 1) is again rewarded (Test RS)

(2) shape that was previously trained but not rewarded during Training is rewarded (Test NS)

(3) one shape of the control set is pseudo-randomly chosen to be rewarded (Test CS).

In each sub assessment participants completed 25 blocks with each block containing all 4 shapes (25 x 4 trials) presented in a random order. Participants were informed that during the Test, only one shape would be rewarded and that upcoming reward trials would be indicated with the auditory and visual cues already encountered during Day 1. The order of sub assessments was counterbalanced to control for order effects, and participants were unaware which shapes would be rewarded in each assessment.

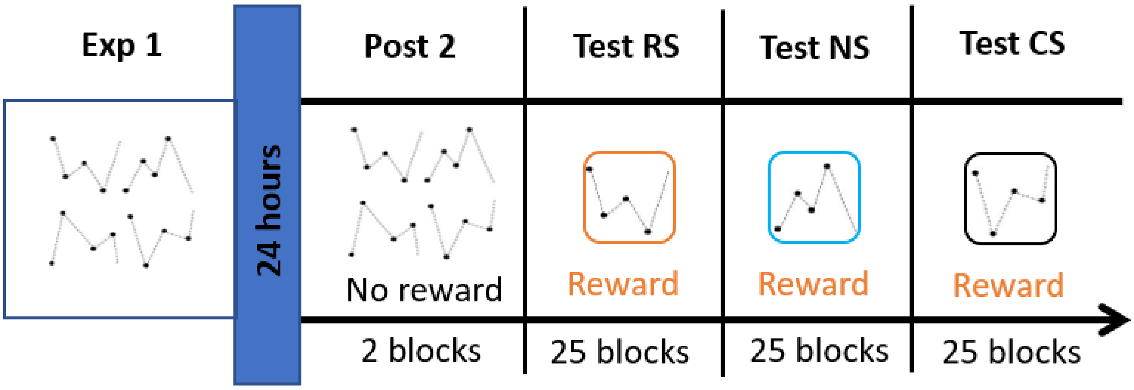


Figure 5 | Experimental design (Experiment 2). Design of Experiment. In this experiment, participants completed experiment 1 on the first testing day. On the second testing day, participants engaged in a second post assessment in which all shapes were presented but no reward was available. Over the course of three tests, reward was available for one shape while all shapes were presented in a random order.

Reward structure and feedback and data analysis: The same Reward structure and feedback as used in Experiment 1 was used in Experiment 2. Similarly, Kruskal-Wallis and Wilcoxon single rank tests were employed again since most of the obtained data was non-parametric. For this analysis, median MTs of the first (early Training) and last (late training) 20 trials of training for each participant in both the reward and no reward condition were used. To assess differences in performance across the Test phase median MTs were calculated.

Results. Day 1 – Replication of results obtained in Experiment 1

MT performance for Day 1 and Day2 are shown in Figure 6. Baseline performance did not differ significantly between conditions (Kruskal-Wallis test; p = 0.1569) and was comparable to MTs produced during Baseline in Experiment 1. Reward expectancy had a similar effect on performance in RS and NS as compared to Experiment 1 with significant decreases in MT during early Training (Kruskal-Wallis test; p < 0.0001 for RS Baseline vs early Training and p = 0.0063 for NS Baseline vs early Training). This effect was again significantly more pronounced in RS when compared to NS during early and late Training (Kruskal-Wallis test; p = 0.01629 for early Training RS vs NS; p < 0.0001 for late Training RS vs NS. During Post assessment performance on RS and NS was significantly faster than CS (Kruskal-Wallis test; p < 0.0001, Tukey's test p < 0.0001 for RS vs CS and p = 0.0001 for NS vs CS). However, performance on RS in comparison to NS did not yield significant results (Tukey's test p = 0.3491). This result is mainly driven by a sharp increase in MT in RS during Post with MTs comparable to MTs produced during late Training in NS. These findings underline that once reward is removed, its effect on movement vigour and subsequent performance decreases. Overall, we were able to replicate the findings from Experiment 1 in Day 1 of Experiment 2.

Day 2 – Retention and generalisation effects

Performance in Post 2 improved significantly compared to Post (Day 1) in all conditions (Kruskal-Wallis test; p = 0.0044 for RS Post vs Post 2; p = 0.0033 for NS Post vs Post 2 and p < 0.0001 for CS Post vs Post 2) indicating a retention effect for all shapes. Importantly, no differences in MT between conditions could be observed (Kruskal-Wallis test; p = 0.3932) which suggests that when not under a reward regime of task ability generalised and is similar across all shapes after 24 hours. During Test, participants completed three sub assessments in which RS, NS and one shape of CS (pseudo-randomly chosen) was rewarded independently in an expected reward design. Across all Tests (Test RS, Test NS, Test CS) we found a significant performance effect using three separate Kruskal-Wallis tests (p < 0.0001 for Test RS and Test NS, respectively and p = 0.0005 for Test CS). Post-hoc analysis revealed that the rewarded shape was performed significantly faster compared to the non-rewarded shapes in each sub assessment (p < 0.0001 for all Tukey's Tests in all 3 Tests). Using a Kruskal-Wallis test we then compared performance on the rewarded shape in each sub assessment with each other (Kruskal-Wallis test; p = 0.0196). Post-hoc analysis revealed that RS was performed significantly faster than CS (Kruskal-Wallis test;p = 0.0196, Tukey's test p = 0.0221 for RS vs CS). However, despite faster MTs in RS than in NS this result failed to reach significance CS (Kruskal-Wallis test; p = 0.0196, Tukey's test p = 0.0811 for RS vs NS). In addition, no significant differences could be observed for NS and CS (Kruskal-Wallis test; p = 0.0196, Tukey's test p = 0.8586 for NS vs CS).

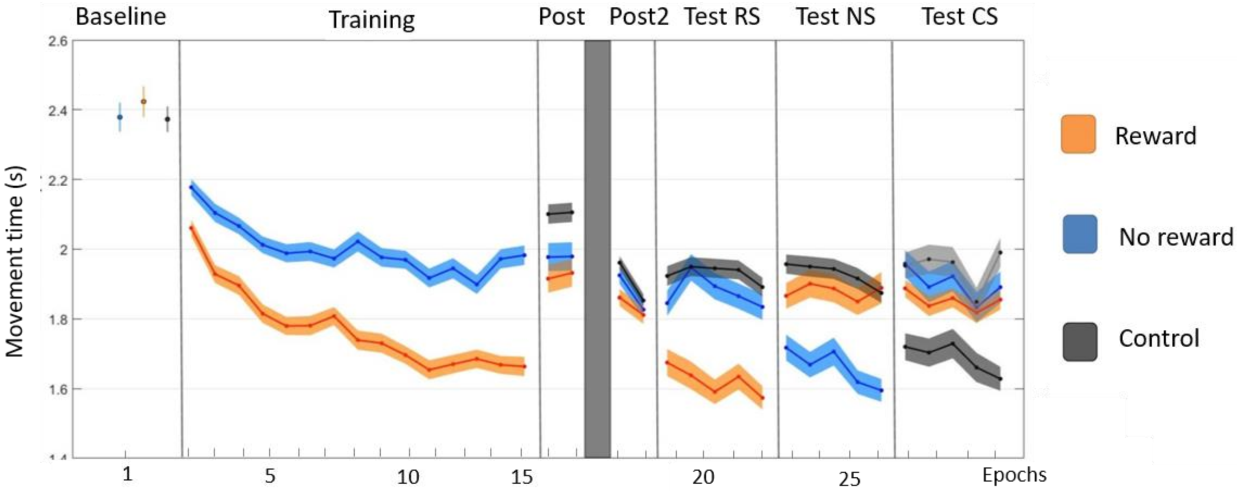


Figure 6 | Results Experiment 2. Movement times (MT) across participants. Values are averaged across 2 trials in baseline and 10 trials in training, both posts and all three tests. Coloured lines represent group mean and shaded areas represent s.e.m.

The aim of these experiments was to determine if reward can invigorate performance of sequential reaching movements. Reaching movements that were paired with reward were consistently performed faster (MT) during Training and Test in both Experiment 1 and 2. These results complement the current literature on motivational vigour and its effects on performance[73–75,155]. However, the results also highlight that rewarded training during the first experimental day does not lead to improvements in retention on the second day. Hence previous findings on the beneficial effects of reward on retention could not be replicated [71,158,177].

2.2 Strategies to reduce MTs during sequential reaching

Results from experiment 1 and 2 suggest that reward-based improvements in MT during sequential reaching follows a previously observed 'on-off' pattern during single-target reaching tasks[179]. Specifically, reductions in MT were seen when participants received reward-based feedback but disappeared once reward was removed.

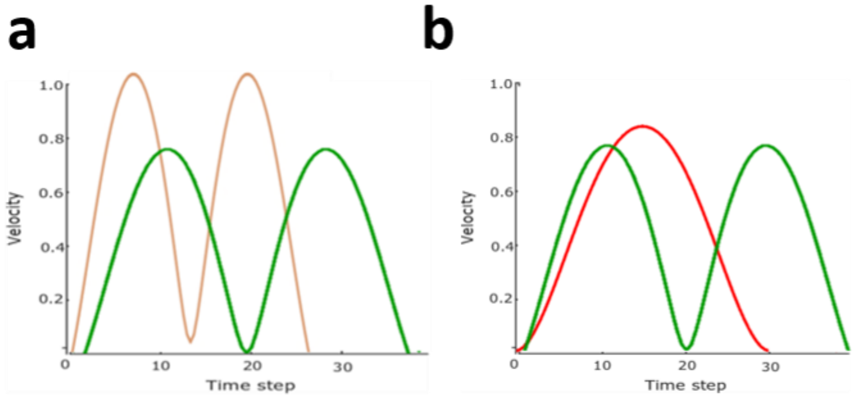


Figure 7| Movement invigoration of sequential reaching movements. (a) Illustration of invigoration of reaching performance. Reward will reduce MTs via increases in peak velocity as seen in the orange line. **(b)** Illustration of coarticulation which leads to reductions in MT via decreases in dwell time around the via point as seen in the red line.

As previously discussed, this 'on-off'-pattern may be explained by reward paying the cost of the increased metabolic effort to move fast while remaining accurate (via increases in arm stiffness)[75]. However, in sequential reaching tasks participants can reduce MTs through two strategies (Figure 7a,b): 1) increasing peak velocities of each individual reaching movement and 2) reduce the dwell time in each target (i.e., periods of stopping with velocities close to 0 cm/s). The first strategy has been shown to underlie reward-based MT improvements during single-target reaching movements[74,75]. However, this strategy comes with a marked escalation in metabolic costs and as such might not represent a sustainable option once reward is removed[75]. The second strategy centres around the idea of increasing movement velocities when passing through targets instead of stopping when reaching them. This strategy will reduce overall MT via a reduction in dwell time around these targets effectively making them via points. Crucial to this strategy is that individual reaching movements are no longer treated as discrete actions but are performed continuously.

2.2.1 Coarticulation

The idea of integrating sequential movements into a continuous action lies at the core of the concept of coarticulation. Coarticulation has been an influential concept in the language production literature for more than 100 years and describes how consecutive phonemes are altered to allow for fluid speech production[180–182]. Specifically, how a preceding phoneme is adjusted to allow for a fluid transition to the next[180–182] as illustrated by the example of articulating ‘H-U’ and ‘H-A’. Hence, depending on the following phoneme and particularly on its phonetic gesture, the coordination of the preceding H is modified which results in temporally overlapping gestures[181]. In such a case of anticipatory coarticulation, there is a lack of parsing between the phonemes which results in a fluid articulation of each speech segment. Importantly, coarticulation is not only the result of physiological constraints of the mouth since vowel coarticulation differs across languages[183]. Consequently, coarticulation in language production is a learned behaviour which has to planned to be executed[183]. However, coarticulation is not restricted to speech production and has been observed in many other behaviours such as finger spelling in sign language[184]. Similarly, coarticulation has been investigated during piano playing in expert pianists[185], while recent research found signatures of coarticulation in typing[186–188]. Additionally, coarticulation has also been observed during various 2D and 3D upper limb actions[111,189–193]. Consequently, coarticulation appears to be a central mechanism to human movement that is ubiquitous in our daily life. Imagine, you want to drink a sip of your coffee. Rather than reaching for your cup with your hand still closed, you start opening your hand while reaching for it. Similarly, you will tilt your cup while moving it to your already opened mouth. This simple example of drinking a sip of coffee highlights the importance of coarticulation to daily activities and illustrates its main feature: it allows for discrete sub-movements to be merged into a single continuous action. Thus, upper limb coarticulation represents a hallmark of skilled sequential performance, which alongside other markers can be used to assess the quality of movements. However, despite its prevalence, previous work using a sequential reaching task has shown that upper limb coarticulation takes humans up to 8 days (1200-2000 trials)[25–27] and monkeys up to 30,000-40,000 trials[194]. This highlights that coarticulation is characterised by a very slow and difficult learning process which is not simply the logical consequence of training. Furthermore, research has demonstrated that coarticulation of sub-movements leads to the gradual development of a new motor primitive that is globally planned and, once initiated, must run to completion[26]. As such, coarticulation is thought to represent a strategy for the evolution of sequential behaviour towards increased efficiency[194,195]. Crucially, coarticulation breaks down in stroke[23,24] and PD patients[29], which severely affects their daily life. In patients exhibiting upper limb impairment following a stroke, behaviour is often characterised by a decomposition of movements into jerky and inefficient sub-movements, with successful recovery being associated with a return to continuous, smooth, and efficient movements[23,24]. Furthermore, work on coarticulation in PD patients suggests that dopamine (DA) might play a role in coarticulation. Research has shown that PD patients exhibit impairments in coarticulation during cursive handwriting[29]. Specifically, compared to healthy controls, PD patients wrote words as a series of discrete movements which was independent of movement speed[29].

In summary, upper limb coarticulation represents a successful strategy to produce smooth and energetically efficient movements; a skill that appears to break down in clinical populations such as stroke and PD patients. Additionally, coarticulating sequential reaching movements leads to a reduction in dwell times around the via points, which effectively reduces overall MT. In contrast, invigoration of peak velocities to decrease MTs has been shown to be energetically expensive and might only represent a viable option as long as reward negates this motor cost. To investigate whether participants pursue either or both strategies, and whether coarticulation leads to improvements in retention, a novel sequential reaching task was developed.

2.3 Continuous sequential reaching task (CSRT)

Previous work by Sosnik et al. (2004, 2007, 2014) could show coarticulation of sequential reaching movements using a task which consisted of a sequence of point-to-point movements (Figure 8a). Participants were asked to connect these four target points 'as rapidly and as accurately as possible'[25]. To track coarticulation, changes in participants' velocity profile were assessed[25–27]. Completing the point-to-point task by executing four discrete reaching movements would be reflected in the velocity profile by four consecutive bell-shapes (each representing one reaching movement). Consequently, the velocity profile should be characterised by four peak velocities and five minimum velocities (including start and end) corresponding to the stop periods when reaching a target (Figure 8b).

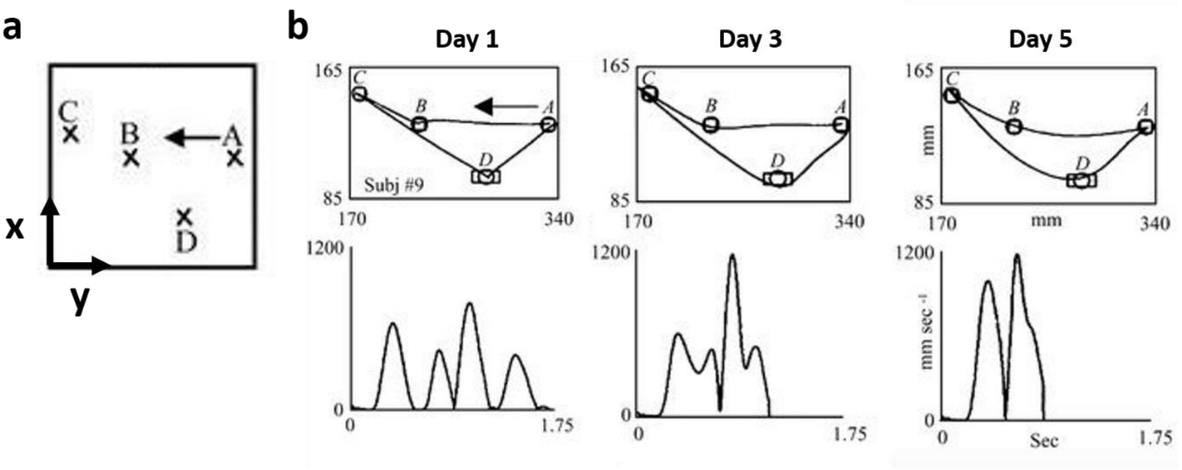


Figure 8 | Extensive training of sequential reaching movements results in coarticulation. Figure adapted from Sosnik et al. (2004). **(a)** Task design. Participants were asked to execute four point-to-point reaching movements (ABCD). **(b)** Upper and lower plots denote example trajectories and corresponding profiles, respectively. Over training, performance changes both in the trajectory (via increases in curvature) and velocity profile (via a reduction in the number of peaks) suggesting that reaching movements have been coarticulated.

Indeed, Sosnik et al. (2004) found that, during the initial testing sessions, participants executed a series of discrete reaching movements[25]. However, with training, movement paths became more curvilinear and the velocity profile began to change. Instead of displaying a four bell-shape velocity profile, performance was now characterised by a double-peaked velocity profile (Figure 8b)[25]. Hence, over the course of training, participants learnt to not stop in two targets and effectively treat them as via points. Therefore, instead of executing four discrete reaching movements, participants ended up producing two coarticulated reaching movements[25–27], which also led to reductions in MT (Figure 8b). Furthermore, they were able to show that, over time, participants' performance aligned with the predictions of a minimum-jerk model[25–27]. The authors argued that these results suggest that maximisation of movement smoothness is an inherent goal of the motor system that allows for a more efficient execution[25]. However, as previously discussed, it remains an open question whether the motor system cares about smoothness and if it does how it were to compute it[101]. Similarly, whether a coarticulated movement sequence represents a new motor primitive or whether it is still represented as a sequence of discrete movements that are dynamically integrated is still an open question[194]. Despite these open questions, Sosnik et al. (2004, 2007, 2014) consistently showed that through coarticulation performance became faster (faster movement time) and more efficient (less jerk)[25–27,196].

The novel task design (CSRT), which is used in most experiments presented in this thesis, was based on this work and involves participants executing a series of 8 reaching movements to hit visual targets that were arranged in a semi-circle around the start box (Figure 9). Participant behaviour could range from executing 8 discrete movements (i.e., stopping at each via point) to a series of 5 movements that would reflect the outcome of minimising jerk across the trial via coarticulation (Figure 9).

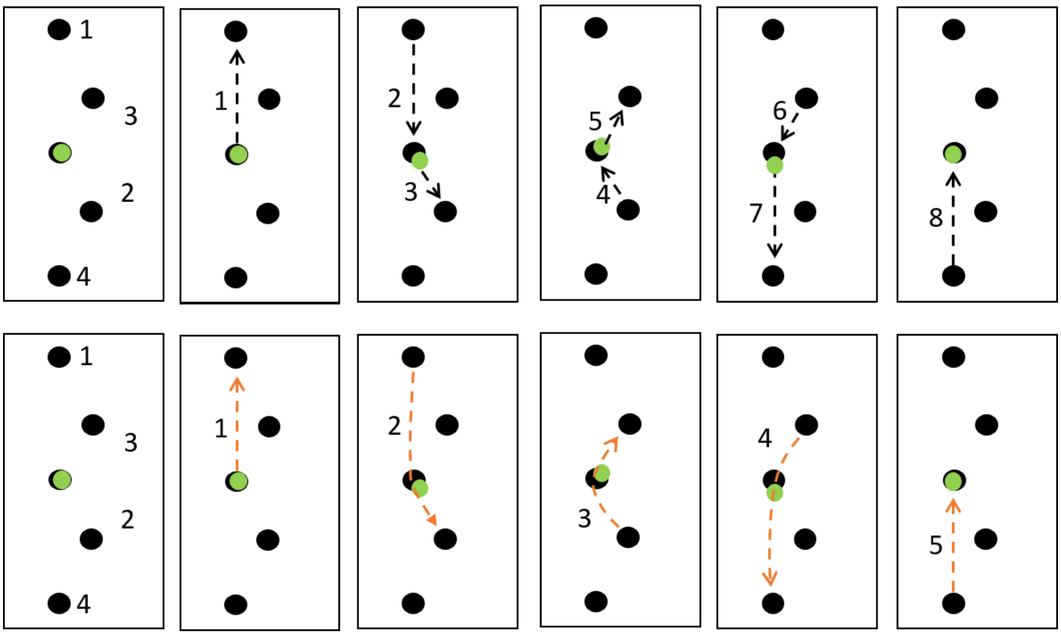


Figure 9 | Task design CSRT. Participants were asked to execute a series of reaching movements to targets in a prespecified order (top panel). Based on the predictions of a minimum-jerk model, it was hypothesised that coarticulation would occur in the in-centre-out reaching movements (reaching movement 2,3 and 4 – lower panel). This would effectively reduce the number of reaching movements from eight to five.

These predictions were based on the result of modelling a trajectory which would minimise jerk across the sequence of movements[28,99]. According to the model[28,99], it was hypothesised that coarticulation would mainly occur in the central set of three in-centre-out reaching movements (Figure 9). Such changes in the number of reaching movements come with marked differences in the velocity profile. Specifically, executing discrete reaches characterised by pronounced stops in-between movements would correspond to an eight-peaked velocity profile. In contrast, coarticulating three consecutive pairs of reaching movements would reduce the number of peaks to five (Figure 10). The aim of this task was to investigate whether reward-based invigoration of performance would be driven by either 1) increases in peak velocities and/or 2) increases in coarticulation which leads to reductions in dwell time. Both strategies have been shown to reduce MTs participants pursue either or both strategies and whether coarticulation leads to improvements in retention a novel sequential reaching task was developed.

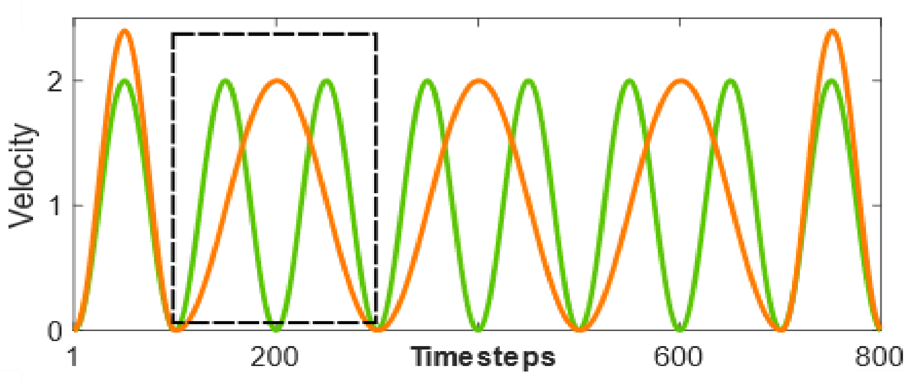


Figure 10 | Changes in the velocity profile though coarticulation. Schematic velocity profiles highlighting the difference between single versus coarticulated reaching movements for a trial. Based on the predictions of a minimum jerk model, we assumed that the three sets of in-centre-out reaching movements could be coarticulated which would reduce the number of velocity peaks from 8 (green) to 5 (orange).

The CSRT design was favoured over the original design by Sosnik et al. (2004) because the design neatly allows for different reaching sequences with regards to the reaching order while preserving the spatial location of the targets. This feature is particularly useful and can be leveraged to conduct future experiments, such as when trying to assess whether coarticulation is generalisable (i.e., are previously coarticulated reaching movements preserved in a different reaching sequence). Additionally, compared to the task design used in the pilot studies, the CSRT was specifically designed to address the overarching research question of this thesis and its main feature is that it enables coarticulation of sequential reaching movements. This has been shown to reduce MTs over the course of training[25].

2.3.1 Performance analysis in CSRT

To analysis performance on the CSRT a range of outcome measures were employed which will be presented briefly in this section.

Movement time (MT): Changes in MT have consistently been used as a proxy to measure the effect of reward on performance invigoration. Hence, to assess whether reward invigorates performance in the CSRT, MT was measured as the time between exiting the start box and reaching the last target. This time excludes reaction time, which describes the time between target appearance and when the participants' start position exceeded 2cm (i.e., when the participant has exited the start box).

Peak and minimum velocities: Peak and minimum velocities (Figure 11) were determined by firstly calculating the derivative of the positional data (*x, y*). The obtained velocity profiles for each trial were subsequently smoothed using a gaussian smoothing kernel (σ = 2).

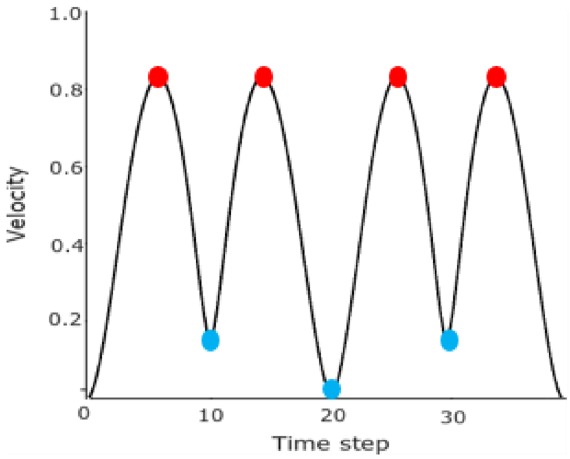


Figure 11 | Peak and minimum velocities. Schematic velocity profile with corresponding peak (red) and minimum (blue) velocities.

The velocity profile was then divided into segments representing movements to each individual target (8 segments) by identifying when the positional data was within 2cm of a target. Peak velocity (v_{peak}) was determined for each segment by finding the maximum velocity:

$$v_{peak} \triangleq max_{\epsilon\,[t_1\,t_2]}\, v(t) \qquad \text{Equation 1}$$

Where *v(t)* is the velocity of segment *t,* and t_1 *and* t_2 represent the start and end of segment *t* respectively. Similarly, minimum velocities (v_{min}) were determined by:

$$v_{min} \triangleq min_{\epsilon\,[t_1\,t_2]}\, v(t) \qquad \text{Equation 2}$$

The individual peak and minimum velocities were then averaged for each trial.

Coarticulation index (CI): To investigate whether participants increase coarticulation of sequential movements to reduce MTs, a novel index was developed. CI can reliably determine how much coarticulation is present in a given sequence. Coarticulation describes the blending together of individual motor elements into a singular smooth action. This is represented in the velocity profile by the stop period between the two movements gradually disappearing and being replaced by a single velocity peak (Figure 12a) [25–27]. Based on the predictions of a model which minimised jerk (Equation 4), it was hypothesised that coarticulation would mainly occur between out-centre centre-out reaching movements (Figure 12b). We therefore excluded the first and last target reach from this analysis. To measure coarticulation, we compared the mean peak velocities of two sequential reaches with the minimum velocity around the via point. The smaller the difference between these values, the greater coarticulation had occurred between the two movements (Figure 12a) [193]. The velocity profile was cut into 3 segments depending on the peak velocities of the out-centre and centre-out reaching movements. The minimum velocity of these segments was calculated and compared to the average of the peak velocities:

$$Coarticulation\ Index \triangleq 1 - \frac{\left(\frac{v_{peak^1} - v_{peak^2}}{2}\right) - v_{min}}{\left(\frac{v_{peak^1} - v_{peak^2}}{2}\right)} \qquad \text{Equation 3}$$

with v_{peak1} and v_{peak2} representing the velocity peak value of the out-centre and centre-in reaching movement of a given segment, respectively, and v_{min} representing the minimum value between these two points. The obtained difference was normalised, ranging from 0 to 1, with 1 indicating a fully coarticulated movement. Given that in this task three pairs of movements were able to be coarticulated, the maximum CI value was 3 in each trial. Note that this measure is MT insensitive since it focuses on the difference in values not on their absolute value. This means slow and fast movements could have a similar CI value.

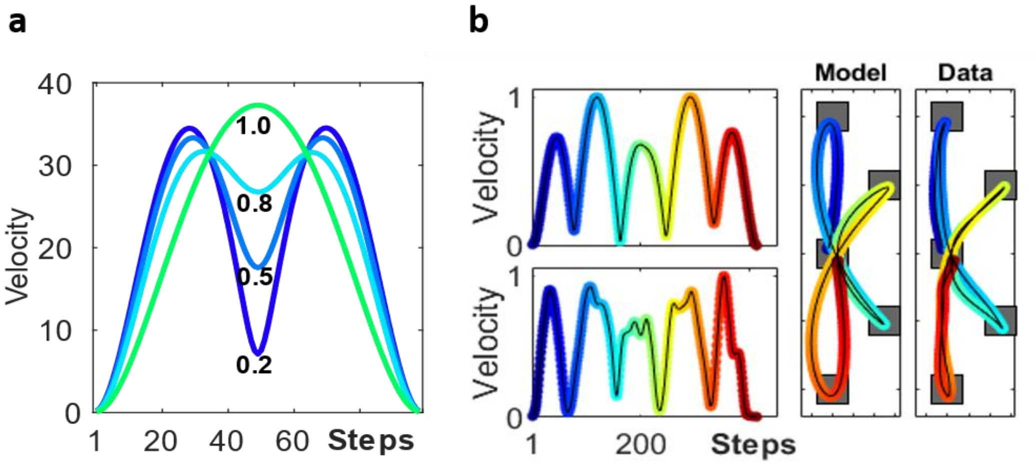


Figure 12 | Coarticulation index (CI) and minimum-jerk model. (a) Schematic illustration of how coarticulation manifests itself in the velocity profile. Increases in the minimum velocity around the target (via point) will reduce the difference to the peak velocities of the two reaching movements. The smaller this difference, the less dwell time around the target. **(b)** Minimum-jerk trajectory prediction ('Model') and the corresponding velocity profile (top panel). These predictions were used to compare the how well the model fits actual data on a trial-by-trial basis.

Minimum-jerk model: Previous work by Sosnik et al. (2004) has shown that, with training, performance on a sequential reaching task progressively aligns with the predictions of a minimum-jerk model[196]. This suggests that performance becomes more optimal with regards to smoothness maximisation. To assess whether this can be replicated in the CSRT, performance was compared to the predictions of a minimum-jerk model (Figure 12b). A traditional minimum-jerk model for motor control is guided by optimisation theory, where a 'cost' is minimised over the trajectory [28,99]. In the case of the minimum-jerk model, the cost is defined as the squared jerk (3rd derivative of position with respect to time):

$$jerk \triangleq \frac{1}{2}\int_{t_1}^{t_2}\left(\left|\frac{d^3x}{dt^3}\right|^2 + \left|\frac{d^3y}{dt^3}\right|^2\right) dt \qquad \text{Equation 4}$$

Here x and y represent the position of the index finger over time (t), while t_1 and t_2 define the start and end of a trial in seconds (t). The Matlab code provided by Todorov and Jordan (1998)[197] was used to compute the minimum-jerk trajectory (a trajectory that minimised Equation 4), and the accompanying velocity profile, given a set of via points, start/end position and movement time [28]. The model fit was then calculated using mean square error (immse function in Matlab) between the predicted and actual velocity profile, to estimate the fit on a trial-by-trial basis. Due to the two-dimensional structure of trajectories, in this analysis velocity profiles rather than the trajectories were used for this comparison.

Spectral arc length: As previously described, motion smoothness has been identified as an important marker of skilled performance and has been widely used to assess recovery following stroke[85–87]. However, how to best measure or quantify motion smoothness has been a centre of debate[85]. Several smoothness metrics have been shown to be sensitive to changes in MT (e.g., log dimensionless jerk[85,87]). Consequently, in addition to the traditional jerk metric in the modelling analysis, spectral arc length was used to measure motion smoothness. Spectral arc length has been shown to be less sensitive to differences in MTs and more sensitive to changes in smoothness [85,87]. The spectral arc length is derived from the arc length of the power spectrum of a Fourier transformation of the velocity profile. An open-source Matlab toolbox was used to calculate this value for each trajectory [198].

Spatial reorganisation: In addition to CI, coarticulation can also be expressed spatially. This approach is based on previous work by Sosnik et al. (2004) who showed that with coarticulation the via point of the reaching movement starts to shift (Figure 13a)[25]. During early training, the location of the via point corresponds to the target location. However, with training the via point shifts in location along the trajectory (Figure 13a). Here an adapted approach was pursued, in which coarticulation was operationalised as the radial distance between the peak velocity (v_{peak}) on the sub-movements and the minimum velocity (v_{min}) around the via point (Figure 13b). This distance becomes smaller with increased coarticulation [25,26] and reflects the merging of two sub-movements into one (Figure 13c).

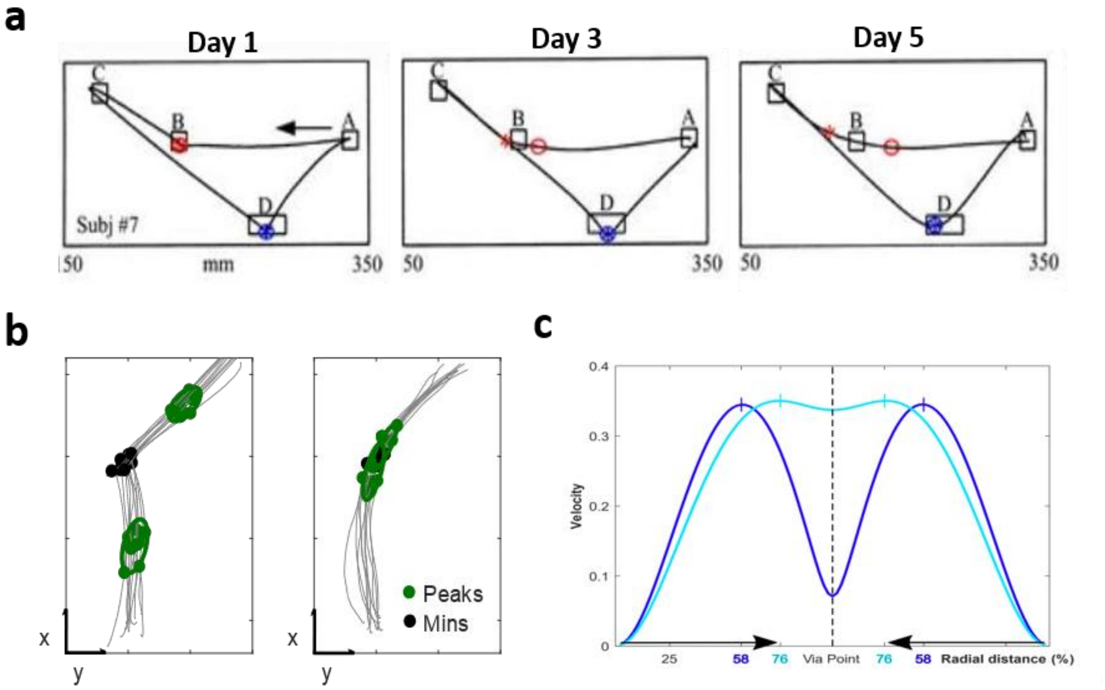


Figure 13 | Radial distance as a proxy measure of coarticulation. (a) Adapted from Sosnik et al. (2004). The red circle representing the via point in this case shifts with coarticulation from the target B to a more central location along the path between target AC. **(b)** Example data for the spatial location of peak velocities when performing two individual (left panel) and one coarticulated (right panel) reaching movement. With coarticulation, the peak velocities (green dots) drift towards the via point (in black min velocities of the reaching movement) **(c)** Illustration of radial distance (%) values depending on different levels of coarticulation. Considering two target reaches (same length), the spatial position (*x,y*) changes depending on the level of coarticulation thereby exhibiting less radial distance to the via point.

To measure these changes in radial distance between peaks and via points, a sliding window approach of 10 trials at a time was used. For each target reach (excluding the first and the last) we fitted a confidence ellipse [199] with a 95% confidence criterion around the scatter of the spatial position (*x, y*) of each peak velocity of the included trials (Figure 13b). The confidence ellipses were obtained using principal component analysis to determine the minimum and maximum dispersion of the included data points in the *x-y* plane. To measure the distance between the scatter and its corresponding via point, we determined the ellipse's centroid (point of intersection of ellipse's axes) and calculated the radial distance to the via point. The obtained distance values were normalised and ranged from 0-100%, with 100% representing 0 cm distance between the centroid and the via point. Considering that individual reaching movements display a bell-shaped velocity profile, with the v_{peak} situated approximately in the centre of the movement, radial distance values between 45-55% can be expected if each movement is executed individually (Figure 13c). To understand the optimal radial distance to the via point, we measured the radial distance for a trajectory which minimised jerk (using the minimum-jerk model). This suggested that values of 82-85% represent the optimal range for coarticulated movements in this task.

2.3.2 Conclusions and outlook

Seeking reward is a powerful tool for shaping behaviour [124,134] and it has been shown that reward leads to saccadic and reaching actions to be executed with greater speed and accuracy [72–75,179,200–202]. However, it has also been shown that these reward-based effects on movement are transient i.e. the effects are lost when reward is removed[73–75]. Codol et al. (2020) offers an explanation suggesting that increases in metabolic cost through increases in arm stiffness are only viable as long as reward is available to pay for it. Consequently, once reward is removed performance reverts back to near baseline levels[75]. However, it is unclear why this does not seem to be the case during more complex continuous tasks[158,177]. To address these inconsistencies, two pilot studies were conducted that assessed the effect of reward on complex movements in a continuous reaching task. The results from Experiment 1 and 2 show that reaching movements that are paired with reward are consistently performed faster (MT). These results complement the current literature on motivational vigour and its effects on performance[73–75,155]. However, the results also highlight that rewarded training during the first experimental day does not lead to improvements in retention on the second day. Hence previous findings on the beneficial effects of reward on retention could not be replicated [71,158,177]. Despite this, a steady improvement in MTs can be observed across training which suggests that deceases in MTs are not solely driven by a tonic invigoration effect but that over the course of training participants learn how to execute the sequence faster. This observation and that fact that movement sequences can be executed faster via two strategies: 1) increases in peak velocities and 2) reduction in dwell times (i.e., coarticulation), inspired the development of a novel task (CSRT). The task design allows for improvements in MT to be driven by both strategies and therefore provides a unique framework to study the effects of reward the execution of complex movement sequences. Specifically, whether reward-based improvements are driven by one or both strategies can be explored using the novel task design (Chapter 3).

Chapter 3

1

Long-term retention of reward-based improvements during complex movements is facilitated by coarticulation

3.1 Introduction

Seeking reward is a powerful tool for shaping behaviour [124,134]. For example, the expectation of reward causes individuals to perform saccadic and reaching actions with greater speed and accuracy [72–75,179,200–202]. As a result, there has been significant interest in the potential of using reward to enhance motor behaviour in both healthy individuals and clinical populations such as stroke patients [4,5]. However, it has been shown that these reward-based effects on movement are cognitively demanding, energetically-inefficient and transient i.e. the effects are lost when reward is removed [73–75]. Importantly, these results stem from simple tasks that involve singular discrete actions e.g., movement towards a single static target. In contrast, during more complex sequential or continuous tasks, the beneficial effects of reward appear long lasting and persist even after the removal of reward [158,159,177]. The mechanism, however,

---

This chapter has been published as a pre-print on bioRxiv at the URL https:\\www.biorxiv.org/content/10.1101/2020.06.15.152876v1.full. I designed and implemented the task, acquired, analysed and interpreted the data, created the figures and wrote the manuscript. X. Chen helped with the data analysis (minimum-jerk model) which was implemented by me. J.M Galea helped design the task, interpret the data and gave feedback during the writing of the manuscript. J.M. Galea provided the funding, facilities and materials to conduct this work, by which reward induces these long-lasting effects and the reason why sequential or continuous behaviour seems critical is unknown. An interesting possibility is that reward-driven performance gains are only maintained if they are accompanied by improvements in kinematic efficiency. We propose that this is possible during sequential actions through coarticulation.

Ubiquitous in our daily life but often impaired in clinical populations [21–24], coarticulation describes the process of blending together a series of distinct movements into a single continuous action. For example, compound actions such as drinking a cup of coffee or cleaning your teeth are often executed discretely with pronounced stop periods between them when first encountered [60,83,181,194,195,203–205]. However, with learning these sub-movements are gradually blended together to form a continuous action that is executed with increased speed, smoothness and energetic-efficiency [25–28]. Therefore, coarticulation represents a hallmark of skilled sequential behaviour performance as it reflects the evolution of behaviour, both temporally and spatially, towards increased efficiency [25–28,83,194]. Coarticulation is observed across many motor behaviours such as speech production [180,182], sign language [184], piano playing [185], typing [186–188] and various other upper limb actions [111,190–193,206]. Within sequential reaching, the coarticulation of sub-movements leads to the gradual development of a new motor primitive that is globally planned and, once initiated, must run to completion [26]. However, despite its prevalence, previous work has shown that coarticulation takes humans up to 8 days (1200-2000 trials) [25–27] and monkeys up to 30,000-40,000 trials [194]. Additionally, it has been observed that not all participants learn to coarticulate [25]. This highlights that coarticulation is characterised by a very slow and difficult learning process which is not simply the logical consequence of training.

Here we tested the hypothesis that reward can facilitate coarticulation during sequential reaching and thereby promote energetically efficient performance gains that persist even in the absence of reward. To this end, our main experiment assessed the effect of reward on sequential reaching performance and the evolution of coarticulation over the course of two testing days. We then carried out two further experiments to assess the robustness of these performance gains during an additional testing day without reward availability, and to investigate whether reward or performance-based feedback drove these observed improvements.

3.2 Methods

Participants: 107 participants (17 males; age range 18 - 35) were recruited to participate in three experiments, which had been approved by the local research ethics committee of the University of Birmingham. All participants were novices to the task paradigm and were free of motor, visual and cognitive impairment. Most participants were self-reportedly right-handed (N = 7 left-handed participants) and gave written informed consent prior to the start of the experiment. For their participation, participants were remunerated with either course credits or money (£7.5/hour) and were able to earn additional money during the task depending on their performance. Depending on the experiment, participants were pseudo-randomly allocated to one of the available groups.

Experimental apparatus: All experiments were performed using a Polhemus 3SPACE Fastrak tracking device (Colchester, Vermont U.S.A; with a sampling rate of 110Hz). Participants were seated in front of the experimental apparatus which included a table, a horizontally placed mirror 25cm above the table and a screen (Figure 1a). A low-latency Apple Cinema screen was placed 25cm above the mirror and displayed the workspace and participants' hand position (represented by a green cursor – diameter 1cm). On the table, participants were asked to perform 2-D reaching movements. Looking into the mirror, they were able to see the representation of their hand position reflected from the screen above. This setup effectively blocked their hand from sight. The experiment was run using MATLAB (The Mathworks, Natwick, MA), with Psychophysics Toolbox 3.

Task design: Participants were asked to hit a series of targets displayed on the screen (Figure 1b). Four circular (1cm diameter) targets were arranged around a centre target ('via target'). Starting in the via target, participants had to perform eight continuous reaching movements to complete a trial. Target 1 and 4 were displaced by 10cm on the *y*-axis, whereas Target 2 and 3 were 5cm away from the via target with an angle of 126 degrees between them (Figure 1b). Our task design was based on previous work [25–27,196] in which the authors were able to observe coarticulation using similar angles and reaching distance configurations. Participant behaviour could range from executing 8 individual movements (i.e., stopping at each via point; Figure 1c) to a series of 5 coarticulated movements that would reflect the outcome of minimising jerk across the trial (chapter 2 Equation 4; Figure 1d) [28]. Changes in the number of reaching movements were associated with marked differences in the velocity profile. Specifically, individual reaches were characterised by pronounced stops in-between movements, with velocity dropping close to zero. As reaching movements merge through coarticulation, these dips in velocity gradually disappear (Figure 1c). To start each trial, participants had to pass their cursor though the preparation box (2x2cm) on the left side of the workspace, which triggered the appearance of the start box (2x2cm) in the centre of the screen. After moving the cursor into the start box, participants had to wait for 1.5s for the targets to appear. This ensured that participants were stationary before reaching for the first target. Target appearance served as the go-signal and the start box turned into the via target (circle). Upon reaching the last target (via target), all targets disappeared, and participants had to wait for 1.5s before being allowed to exit the start box to reach for the preparation box to initiate a new trial.

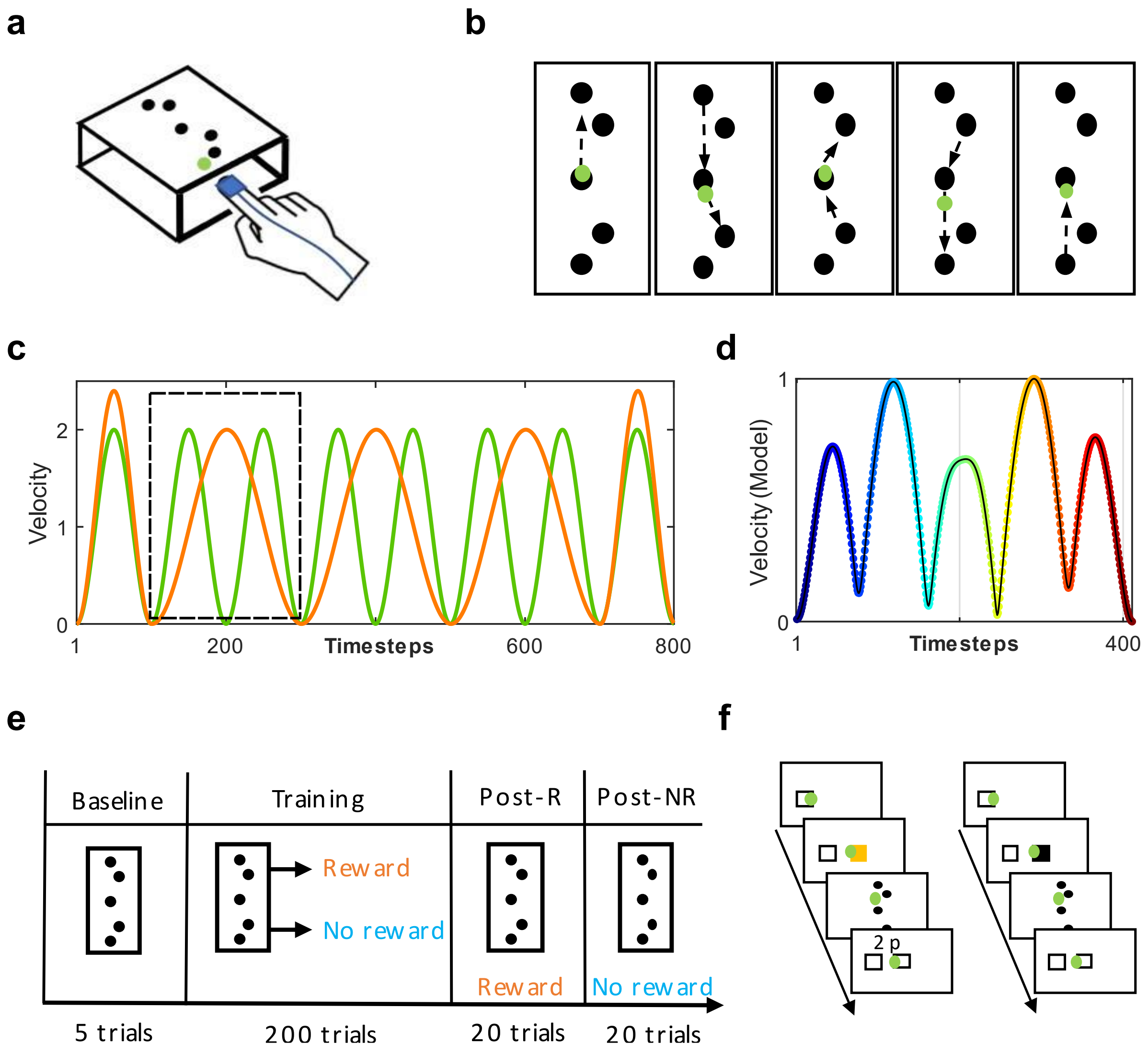


Figure 1 | Experimental setup. a) Participants wore a motion-tracking device on the index finger and the unseen reaching movements were performed across a table whilst a green cursor matching the position of index finger was viewable on a screen. **b)** 8 movement sequential reaching task. The participants started from the centre target. **c)** Schematic velocity profiles highlighting the difference between single versus coarticulated reaching movements for a trial. Based on the predictions of a minimum jerk model, we assumed that the three sets of in-centre-out reaching movements could be coarticulated which would reduce the number of velocity peaks from 8 (green) to 5 (orange). **d)** Velocity profile for a trial predicted by a minimum jerk model. **e)** Study design. After baseline (10 trials) participants were randomly allocated to a reward and no reward group (training - 200 trials) and performed an additional 20 trials in each post assessment; one with reward (post-R) and one without (post-NR) (counterbalanced across participants). This design was repeated 24 hours later (Day 2). **f)** Rewarded trials were cued using a visual stimulus prior to the start of the trial. At the end of the trial, participants received trial-based monetary feedback based on their last 20 trials (closed-loop design). In no reward trials, participants were instructed to be as fast and accurate as possible.

Participants had to repeat a trial if they missed a target or performed the reaching order incorrectly. Similarly, exiting the start box too early either at the beginning or at the end of each trial resulted in a missed trial.

Reward structure and feedback: Participants in experiment 1 and experiment 2 experienced either reward or no reward trials depending on the current experimental phase (Figure 1e): (1) Reward trials were cued using a visual stimulus prior to the start of the trial. Once participants moved into the preparation box, the start box appeared in yellow (visual stimulus) rather than in black (Figure 1f). Participants were informed that faster MTs would earn them more money, with a maximum amount of 5p available in each trial. While participants moved from the start box to the preparation box to initiate a new trial, the amount earned in the previous trial was displayed on the top of the screen (i.e. 'You have earned 2p out of 5p'). We used a closed-loop design to calculate the amount of reward earned in each trial. To calculate this, we included the MT values of the last 20 trials and organised them from fastest to slowest to determine the rank of the current trial within the given array. A rank in the top three (<= 90%) returned a value of 5p, ranks >= 80% and <90% were valued at 4p; ranks >=60% and <80% were awarded 3p; ranks >=40% and < 60% earned 2p while 1p was awarded for ranks >=20% and < 40%. A rank in the bottom three (<20%) returned a value of 0p. When participants started a new experimental block, performance in the first trial was compared to the last 20 trials of the previously completed block. (2) No reward trials were not cued, and no reward was available for participants. However, participants were instructed to 'move as fast and accurately as possible'. In experiment 3, participants were randomly allocated to one of the four groups: (1) no reward, (2) reward without performance-based feedback, (3) reward with random feedback and (4) reward with accurate feedback. Participants in the no reward (1) and reward with accurate feedback (4) groups underwent the same regime as the no reward and reward groups in experiment 1 respectively. To investigate whether reward and/or feedback drove performance changes, we changed the reward and feedback structure for the two remaining groups. Participants in the reward without performance-based feedback group (2) were able to earn money depending on their performance in each trial. Reward trials were cued with a yellow start box prior to the start of the trial. However, participants did not receive any feedback on their performance after completing a given trial. They were asked to initiate a new trial and 'be as fast and accurate as possible to earn more money'. In contrast to this, participants in the reward with random feedback group (3) received feedback after completing a reward trial during training, which were also cued with a yellow start box. However, feedback in this group was not performance-based, but was drawn randomly from feedback given to participants in experiment 1. To this end, we strung together all reward values given to participants in the first experiment and randomly chose a value for feedback in a given trial in experiment 3. Participants, therefore, received feedback which was similar in reward probability without corresponding to actual performance.

Experiment 1 experimental procedure: In this experiment, we investigated whether reward can invigorate performance on a sequential reaching task. The experiment included an initial learning phase prior to the start of the experiment as well as a baseline, training and two post assessments. The same design without the learning phase was repeated 24 hours later (Figure 1b). Participants were pseudo-randomly allocated to either the reward or no reward group (N = 21 each) and were informed that at some point during the experiment they would be able to earn additional money depending on their performance.

Learning: We included a learning phase prior to the start of the experiment for participants to be able to memorise the reaching sequence. This allowed us to attribute any performance gains to improvements in execution rather than memory. Once participants waited 1.5s inside the start box, the targets appeared which were numbered clockwise from 1 to 4 starting with the central top target. Participants were also able to see a number sequence at the top left of the screen displaying the order of target reaches (1 – 3 – 2 - 4). Participants were instructed to hit the targets according to the number sequence while also hitting the via target in between target reaches. They had to repeat a trial if they missed a target or performed the reaching order incorrectly. Similarly, exiting the start box too early either at the beginning or at the end of each trial resulted in a missed trial. After a cued trial, participants were asked to complete a trial from memory without the number sequence or numbers inside the targets. If participants failed a no cue trial more than twice, cues appeared in the following trial as a reminder. After a maximum of 10 cue and 10 no cue trials participants completed this block.

*Baseline***:** Participants in both groups completed 10 baseline trials, which were used to assess whether there were any pre-training differences between groups. Both groups were instructed to ‘move as fast and accurately as possible’, while no performance-based feedback was given at the end of each trial.

Training: Participants in the reward group were informed that during this part they would be able to earn money depending on how fast they complete each trial (200 reward trials). In contrast, participants in the no reward group engaged in 200 no reward trials and were again instructed to move as fast and as accurately as possible.

*Post assessments***:** On both testing days, participants from both groups were asked to complete two post assessments (20 trials each); one with reward trials (post-R) and one with no reward trials (post-NR). The order was counter-balanced across participants.

Experiment 2 experimental procedure: In this experiment, we aimed to test how robust reward-driven performance gains were over an additional testing day without reward availability. Participants (N = 5) underwent the same regime as the reward group in experiment 1 on the first two testing days. On the third testing day after baseline, participants were asked to complete 200 no reward trials.

Experiment 3 experimental procedure: Here, we investigated whether the observed performance gains depend on reward expectation, performance-based feedback or a combination of both. To this end, we allocated participants (N = 60) to one of the four groups:

(1) no reward, (2) reward without performance-based feedback, (3) reward with random feedback and (4) reward with accurate feedback (see Reward Structure and Feedback for more information). Participants underwent the same procedure as participants in experiment 1 over the course of the first testing day. After a learning phase and a baseline part, participants engaged in 200 training trials which differed with regards to their reward and feedback structure. Similarly, to experiment 1 and 2, participants then completed the post assessments, which were counter-balanced across participants.

Data analysis: Analysis code is available on the Open Science Framework website, alongside the experimental datasets at: https://osf.io/62wcz/. The analyses were performed in Matlab (Mathworks, Natick, MA). We used parameters to assess performance which were described in more detail in chapter 2.

Movement Time (MT): MT was measured as the time between exiting the start box and reaching the end target. This excludes reaction time, which describes the time between target appearance and when the participant has exited the start box. Trials with MTs beyond 9.0s were excluded from analysis, which amounted to 0.37% of all trials.

Peak Velocity: Through the derivative of positional data (*x, y*), we obtained velocity profiles and found the peak velocities of each reaching movement. These were then averaged to provide a peak velocity measure for each trial.

Coarticulation Index (CI): To measure coarticulation, we compared the mean peak velocities of the two sequential reaches with the minimum velocity around the via point. The smaller the difference between these values, the greater coarticulation had occurred between the two movements[193]. We normalised the obtained difference, ranging from 0 to 1, with a CI value of 1 indicating a fully coarticulated movement.

Error: We operationalised error as the amount of repetitions necessary to complete a given trial. Trials had to be repeated if participants missed a target or if they exited the start box before the targets appeared or disappeared.

Minimum-jerk model: We used the minimum-jerk model to model trajectories for each angle configuration used in this experiment using the Matlab code provided by Todorov and Jordan (1998)[28]. We then calculated the mean square error (immse function in Matlab) between the predicted and actual velocity profile, which were both normalised and interpolated (N = 500), to estimate the fit on a trial-by-trial basis. Due to the two-dimensional structure of trajectories, we used velocity profiles rather than the trajectories for this comparison.

Spectral Arc Length: To assess movement smoothness, we used spectral arc length as our smoothness metric. Spectral arc length has been shown to be less sensitive to differences in MT and more sensitive to changes in smoothness [85,87]. The spectral arc length is derived from the arc length of the power spectrum of a Fourier transformation of the velocity profile. We used an open-source Matlab toolbox to calculate this value for each trajectory [198].

For both spectral arc length and the minimum-jerk model, we only included non-corrected trials. Trials that were classified as corrected included at least one corrective movement to hit a previously missed target. These additional movements added peaks to the velocity profile which complicated model comparison and increased jerkiness disproportionally. Therefore, 1820 trials were excluded for both analyses (8.68% of all trials).

Spatial reorganisation: To measure changes in radial distance between peaks and via points, a confidence ellipse [199] with a 95% confidence criterion was fitted around the scatter of the spatial position (*x, y*) of each peak velocity of the included trials (sliding window of 10). The distance between the centroid of the ellipse and the via point represents the radial distance.

Variability: To assess changes in variability we measured the area of the peak velocity ellipses using the same approach as for spatial reorganisation. The area of the ellipse represents the total variance of the included data points in the *x*-*y* plane and was calculated by multiplying the axes of the ellipse with pi (π) [199,207]. Data was normalised to the baseline for each group.

Statistical analysis: Wilcoxon tests were used to analyse differences in performance during baseline. Mixed model ANOVAs were used to assess statistical significance of our results in experiment 1. We carried out separate analyses for training with timepoint (early training (first 20 trials), late training (last 20 trials)) and group (reward, no reward) as factors for both days. Similarly, for post assessment condition (post-R, post-NR (all 20 trials in each)) and group (reward, no reward) were used as factors. We used one-sample Kolmogorov-Smirnov tests to test our data for normality and found that all measures were non-parametric. Median values were therefore used as input in all mixed model ANOVAs. Wilcoxon tests were employed when a significant interaction and/or main effects were reported and corrections for multiple comparisons were performed using Bonferroni correction. Linear partial correlations (fitlm function in Matlab) were used to measure the degree of association between the chosen variables, while accounting for the factor group. Piecewise linear spline functions were fitted through the scatter of spatial distance values and CI levels using least square optimisation by means of shape language modelling (SLM) [208]. We used three knots as input for the linear model.

A repeated-measure ANOVA was used to test for significance of our results in experiment 2. We compared MT and CI performance separately with timepoint (early training, late training over all 3 testing days) as the within factor. Due to our data being non-parametric after using one-sample Kolmogorov-Smirnov tests, we included median values as input for all repeated-measure ANOVAs. Wilcoxon test was used as post-hoc test and multiple comparisons were corrected for using Bonferroni corrections.

We used mixed model ANOVAs to statistically analyse our results from experiment 3. Separate analyses were carried out for training and post assessment for both MT and CI. Timepoint (early training, late training) and group (no reward, reward without feedback, reward with random feedback, reward with correct feedback) were factors to assess performance in training. Post assessment performance was analysed with a mixed model ANOVA that had condition (Post-R, Post-NR) and group (no reward, reward without feedback, reward with random feedback, reward with correct feedback) as factors. We used median values as input, because one-sample Kolmogorov-Smirnov tests confirmed a lack of normality. Wilcoxon tests were employed when a significant interaction and/or main effects were reported and corrections for multiple comparisons were performed using Bonferroni correction.

3.3 Results

Reward invigorates sequential reaching movements

MT reflected the total movement duration from exiting the start box until reaching the last target. Our results highlight that reward instantaneously invigorated sequential reaching behaviour with these performance gains being maintained even in the absence of reward (Figure 2a). Specifically, we found a significant decrease in MT over the course of training on both days for both groups (mixed-effect ANOVA; timepoint (early (1st 20 trials) vs late (last 20 trials)) x group; main effect for timepoint, day 1: $F = 18.29$, $p < 0.0001$; day2: $F = 8.51$, $p = 0.0058$; Figure 2b,d). Importantly, despite no differences in MT between groups during baseline on day 1 (Wilcoxon test; $Z = -1.38$, $p = 0.17$), we found that the reward group produced significantly faster MTs than the no reward group across training on both days (main effect for group, day 1: $F = 18.96$, $p < 0.0001$; day 2: $F = 17.58$, $p < 0.0001$; Figure 2b,d). Across post assessments, a significant interaction was found between phase (post-R (all 20 trials) vs post-NR (all 20 trials)) and group (mixed ANOVA; condition x group; interaction, day 1: $F = 18.07$, $p < 0.0001$; day 2: $F = 19.99$, $p < 0.0001$). Specifically, there was a significant difference in MT between groups during post-NR (Wilcoxon test; day 1: $Z = -2.82$, $p = 0.0192$; day 2: $Z = -3.27$, $p = 0.0044$; Figure 2c,e) but not post-R (Wilcoxon test; day 1: $Z = -1.13$, $p = 1$; day 2: $Z = -1.38$, $p = 1$). This indicates that the no reward group were able to instantaneously invigorate their performance during post-R. However, these performance gains were not maintained during post-NR, suggesting that they remained transient in nature.

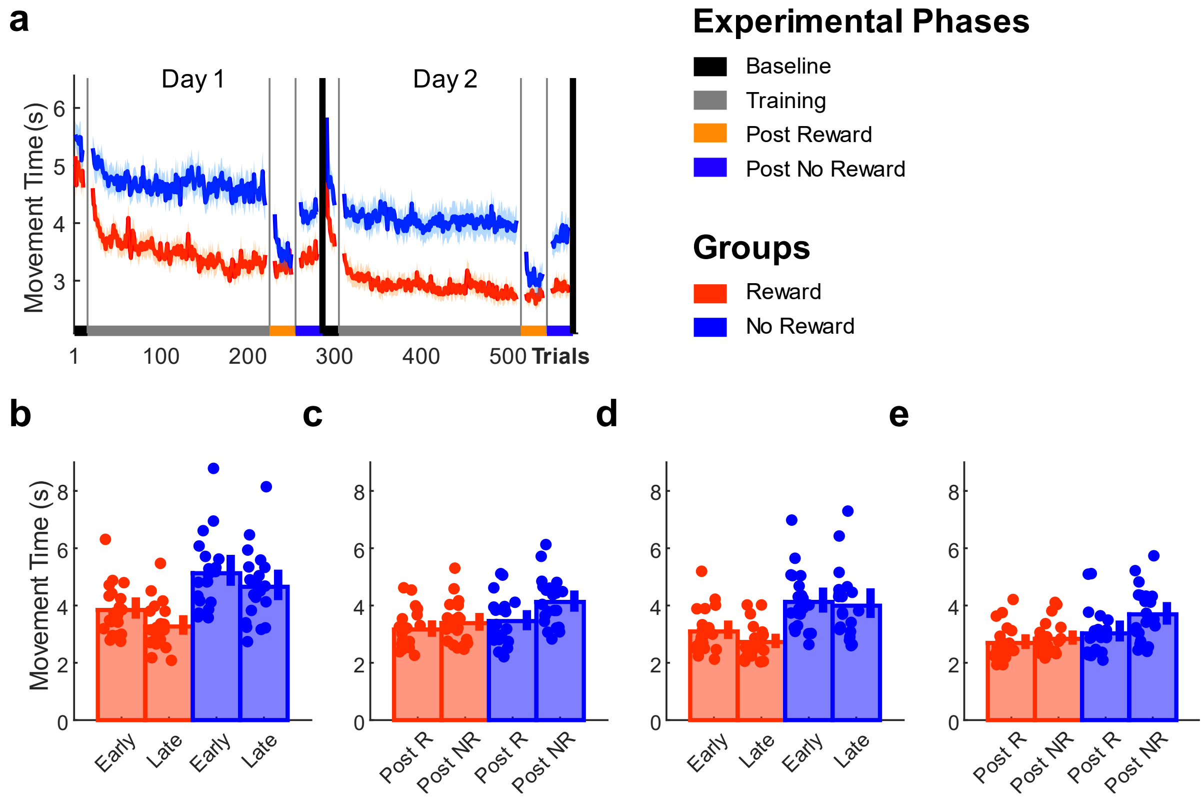


Figure 2 | Reward-based improvements in MT. a) Trial-by-trial changes in MT averaged over participants for both groups. **b-e)** Median MT values for each participant for both groups. **b)** Comparing MT performance during training on day 1 (early (first 20 trials) vs late (last 20 trials)). **c)** Post assessment performance (day 1; post-R vs post-NR). **d)** Training (day 2; early vs late). **e)** Post assessment performance (day 2; post-R vs post-NR). Shaded regions/error bars represent SEM.

In contrast, the enhanced MT performance in the reward group did not change significantly between post assessments, implying that performance gains had become reward-independent (Wilcoxon test; day 1: Z = -1.11, p = 1; day 2: Z = -0.91, p = 1; Figure 2c,e). The counterbalancing of post assessment order did not affect these results (Supplementary Figure 1). These decreases in MT could be driven by two processes: (1) an invigoration of each reaching movement's peak velocity which has been observed to underlie the transient performance gains in singular discrete reaching tasks [73–75] and (2) the coarticulation of sequential reaching movements which decreases MTs via a reduction in dwell time around the central target.

Reward-based invigoration of peak velocities is instantaneous

To assess changes in peak velocities (chapter 2, equation 1), we averaged peak velocity across the 8 reaching movements (Figure 3a). Over the course of training, we found significant increases in peak velocity on both days (mixed ANOVA; timepoint x group; main effect for timepoint, day 1: F = 9.72, p = 0.0034; day 2: F = 8.56, p = 0.0057; Figure 3b,d). Despite there being no differences in peak velocities during baseline on day 1 (Wilcoxon test; Z = 0.70, p = 0.4812), the reward group produced significantly higher peak velocities than the no reward group across training (main effect for group, day 1: F = 19.42, p < 0.0001; day 2: F = 19.28, p < 0.0001; Figure 3b,d).

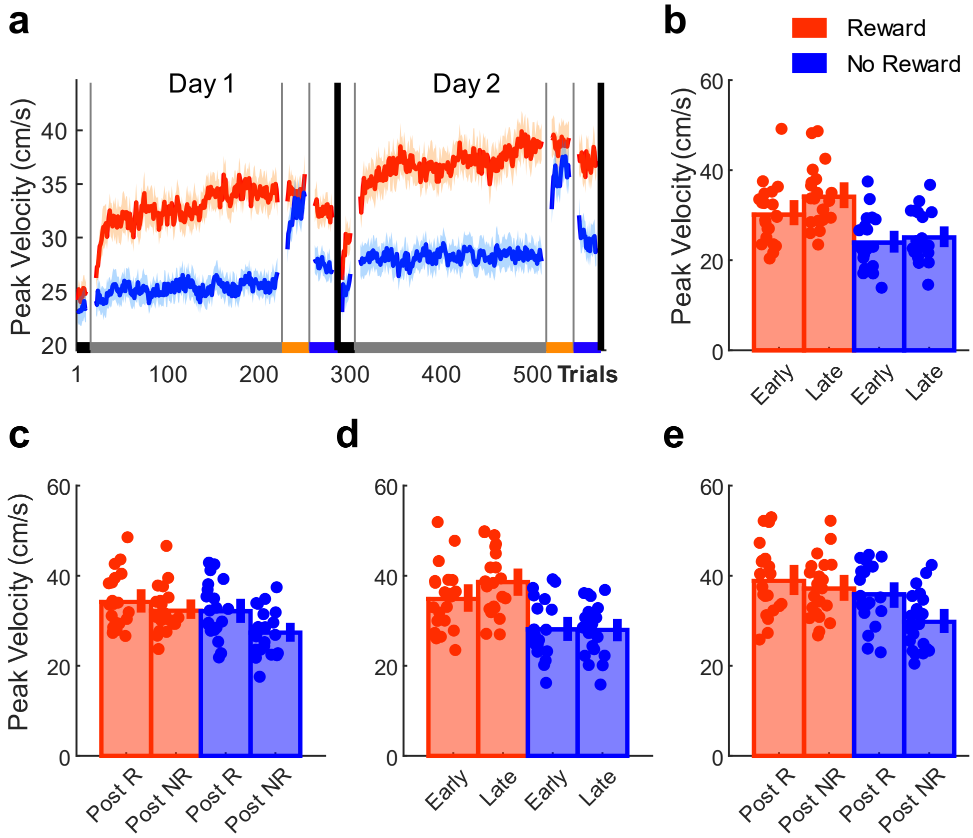


Figure 3 | Reward-based improvements in peak velocity. a) Trial-by-trial changes in peak velocity averaged over participants for both groups. **b-e)** Median peak velocity values for each participant for both groups. **b)** Comparing peak velocity performance during training on day 1 (early (first 20 trials) vs late (last 20 trials)). **c)** Post assessment performance (day 1; post-R vs post-NR). **d)** Training (day 2; early vs late). **e)** Post assessment performance (day 2; post-R vs post-NR). Shaded regions/error bars represent SEM.

This supports existing findings that reward-based invigoration of MT can be driven by increases in peak velocities [73–75]. Yet again, the no reward group exhibited a pronounced 'on-off' effect across post assessments (post-R vs post-NR) for both days (mixed-effect ANOVA; condition x group; interaction, day 1: F = 8.02, p = 0.0072; day 2: F = 24.92, p < 0.0001). Specifically, we found a significant difference between groups during post-NR (Wilcoxon test; day 1: Z = 2.84, p = 0.018; day 2: Z = 3.07, p = 0.0084; Figure 3c,e) but not post-R (Wilcoxon test; day 1: Z = 0.86, p = 1; day 2: Z = 0.93, p = 1; Figure 3c,e). Similarly to MT, these results suggest that the no reward group were able to instantaneously increase peak velocity during post-R, but these performance gains remained transient in nature and were not maintained during post-NR. In contrast, peak velocities in the reward group remained elevated across post assessments irrespective of reward availability (Wilcoxon test; day 1: Z = 0.86, p = 1; day 2: Z = 0.68, p = 1; Figure 3c,e).

Reward-based facilitation of coarticulation is training-dependent

Coarticulation describes the blending of individual motor elements into a combined smooth action. This is represented in the velocity profile by the stop period between two movements gradually disappearing and being replaced by a single velocity peak (Figure 1c). To measure coarticulation, we developed a coarticulation index (chapter 2, Equation 3) that compared the mean peak velocities of two sequential reaches with the minimum velocity around the via point. The smaller the difference between these values, the greater coarticulation had occurred of these two movements as reflected by CI value closer to 1 (Figure 4a). As the central 3 segments of the movement could potentially be coarticulated, the CI ranged from 0-3 for each trial (Figure 4b). Our results show that reward facilitates movement coarticulation and leads to stable changes in behaviour even in the absence of reward.

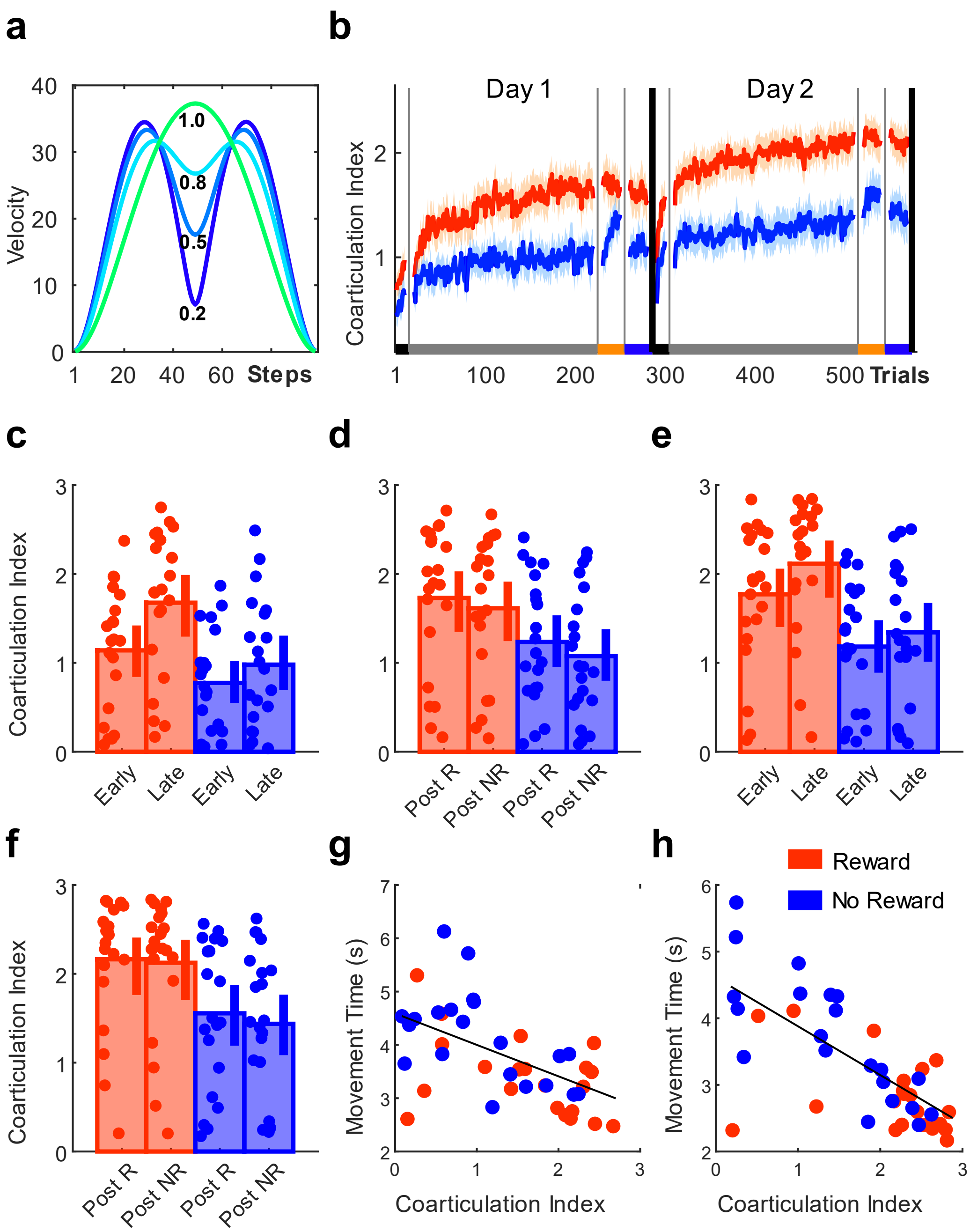


Figure 4 | Reward-based improvements in CI levels. a) Illustration of CI levels. **b)** Trial-by-trial changes in CI levels averaged over participants for both groups. **c-f)** Median CI levels for each participant for both groups. **c)** Comparing CI levels during training on day 1 (early (first 20 trials) vs late (last 20 trials)). **d)** Post assessment performance (day 1; post-R vs post-NR). **e)** Training (day 2; early vs late). **f)** Post assessment performance (day 2; post-R vs post-NR). **g-h)** Scatterplots displaying the relationship between MT and CI levels during post-NR on **g)** day 1 and **h)** day 2 with a linear line fitted across groups. Shaded regions/error bars represent SEM.

Additionally, changes in coarticulation, in contrast to the changes in peak velocities, appear to be training-dependent. Across training, CI levels increased on both days (mixed ANOVA; timepoint x group; main effect for timepoint, day 1: $F = 21.70$, $p < 0.0001$; day 2: $F = 21.45$, $p < 0.0001$; Figure 4b). Although there were no differences between groups in CI levels during baseline (Wilcoxon test; $Z = 1.31$, $p = 0.1908$), we found significantly higher CI levels for the reward group on both days (main effect for group, day 1: $F = 6.81$, $p = 0.0127$; day 2: $F = 9.10$, $p = 0.0044$; Figure 4c,e). Crucially, unlike MT and peak velocity, no significant increases in CI levels were observed for the no reward group on either day during post-R. This suggests that coarticulation cannot be invigorated instantaneously but represents a training-dependent process that is facilitated by reward. Importantly, CI levels were maintained during post-NR on both days for the reward group, highlighting that changes in coarticulation had become reward-independent (mixed-effect ANOVA; condition x group; main effect for group, day 1: $F = 4.91$, $p = 0.0324$; day 2: $F = 7.38$, $p = 0.0097$; Figure 4d,f). To understand whether CI levels are related to the retention of MT performance, we correlated MT values with CI levels during post-NR across participants (Figure 4g,h). We found a significant correlation between CI levels and MT performance for both day 1 (partial correlation controlling for group; post-NR: $\rho = -0.51$; $p < 0.0001$, Figure 4g) and day 2 (partial correlation controlling for group; post-NR: $\rho = -0.66$, $p < 0.0001$, Figure 4h). Although not causal, this indicates that faster MTs during post-NR were associated with higher levels of coarticulation and that the reward group showed better performance for both (Figure 4g,h).

To summarise, these results demonstrate that improvements in MT are driven by two processes which are both reward-sensitive but follow different time courses. The invigoration of peak velocities is instantaneous, whereas coarticulation is training-dependent. Invigoration in the no reward group was mainly driven by increases in peak velocities which were transient in nature, thereby displaying a pronounced ‘on-off’ effect. In contrast, participants in the reward group capitalised on both strategies: the invigoration of peak velocities and additionally increases in coarticulation which led to persistent and ultimately reward-independent improvements in MT performance. Importantly, despite an increase in execution errors on day 1 (reward group), we found that execution errors were not associated with performance levels during late training on day 2 (Supplementary Figure 2 including statistics). In addition, no significant differences between groups were found on day 2 (mixed-effect ANOVA; condition x group; main effect for group, day 2: F = 1.88, p = 0.1785), suggesting that performance gains by the second testing day reflected true improvements in skill.

Reward-based improvements in smoothness are associated with reward-independent maintenance of performance gains

It has been suggested that coarticulation leads to a reduction in jerk and thereby enables smoother and more efficient execution [99,209], which in turn could explain why participants in the reward group maintained their improved performance levels even in the absence of reward. Using spectral arc length as a smoothness metric that is insensitive to movement time[85], we show that the performance in the reward group became progressively smoother (Figure 5a) [209]. Movement smoothness significantly improved across training on both days (mixed ANOVA; timepoint x group; main effect for timepoint, day 1: F = 27.56, p < 0.0001; day 2: F = 18.19, p < 0.0001, Figure 5a). Despite there being no differences between groups during baseline (Wilcoxon test; Z = 1.79, p = 0.0741), the reward group showed a greater improvement in smoothness throughout training (main effect for group, day 1: F = 5.31, p = 0.027; day 2: F = 7.78, p = 0.0080; Figure 5b,d). During post assessments, we found a significant group effect for day 2 (mixed-effect ANOVA; post assessment x group; group, day 1: F = 1.44, p = 0.2369; day 2: F = 6.41, p = 0.0154; Figure 5c,e). This suggests that improvements in smoothness were greater in the reward group and became reward-independent. Additionally, we found that increased smoothness was strongly associated with faster MTs when reward was not available.

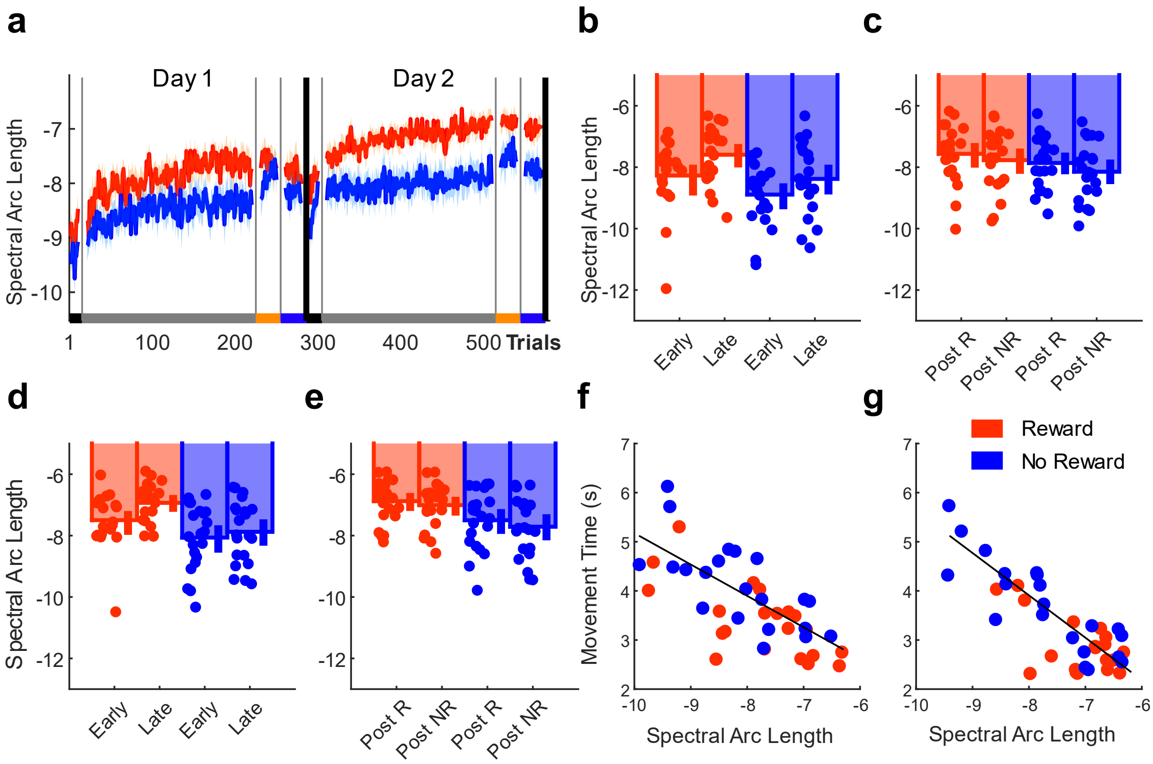


Figure 5 | Reward-based improvements in smoothness. a) Trial-by-trial changes in spectral arc length (smoothness) averaged over participants for both groups. A value closer to zero indicates greater smoothness[85]. **b-e)** Median smoothness values for each participant for both groups. **b)** Comparing smoothness during training on day 1 (early (first 20 trials) vs late (last 20 trials)). **c)** Post assessment performance (day 1; post-R vs post-NR). **d)** Training (day 2; early vs late). **e)** Post assessment performance (day 2; post-R vs post-NR). **f-g)** Scatterplots displaying the relationship between MT and smoothness during post-NR on **f)** day 1 and **g)** day 2 with a linear line fitted across groups. Shaded regions/error bars represent SEM.

We correlated smoothness values with MTs during post-NR on both days (Figure 5f,g), and found a significant correlation between smoothness and MT performance for both day 1 (partial correlation controlling for group; Post-NR: $\rho = -0.69$; $p < 0.0001$, Figure 5f) and day 2 (partial correlation controlling for group; post-NR: $\rho = -0.79$, $p < 0.0001$, Figure 5g). These results align with the aforementioned results showing that CI levels were associated with faster MTs during post-NR, and point to improvements in movement efficiency as a potential mechanism enabling reward-independent maintenance of performance gains during post-NR.

Progressive alignment to the predictions of a minimum jerk model is facilitated by reward

We then assessed whether performance aligned with the predictions of an optimisation model that attempted to minimise jerk given a set of via points, start/end position and movement time [28]. On trial-by-trial basis, mean squared error was calculated between the model and the actual velocity profile (chapter 2, Equation 4; Figure 6a). In comparison to the no reward group, performance in the reward group became significantly more aligned to the predictions of the minimum-jerk model. Mean squared error progressively decreased across days although this was not significant for day 2 (mixed ANOVA; timepoint x group; main effect for timepoint, day 1: $F = 16.19$, $p < 0.0001$; day 2: $F = 2.23$, $p = 0.1429$, Figure 6b). Despite no differences between groups during baseline (Wilcoxon test; $Z = -1.16$, $p = 0.2472$), the reward group's performance showed significantly greater similarity to the model's predictions (main effect for group day 1: $F = 5.61$, $p = 0.0228$; day 2: $F = 7.56$, $p = 0.0089$, Figure 6c,e). Across post assessments, we found a significant interaction for day 2 (mixed-effect ANOVA; condition x group; main effect for interaction, for day 1: $F = 3.92$ $p = 0.0548$; day 2: $F = 7.58$, $p = 0.0088$, Figure 6b,f), however all post-hoc comparisons were not significant. Nevertheless, there was a significant group effect for day 2 (main effect for group, for day 1: $F = 3.26$ $p = 0.0787$; day 2: $F = 4.98$, $p = 0.0313$, Figure 6b,f), suggesting that the degree of similarity to the minimum-jerk model was maintained across post assessments irrespective of reward availability. Within this context, we aimed to assess whether greater similarity to the predictions of the model was related to the maintenance of performance without reward. Using partial linear correlations, we found a correlation between model similarity and MT performance during post-NR for both day 1 (partial correlation controlling for group; Post-NR: $\rho = 0.55$; $p < 0.0001$, Figure 6g) and day 2 (partial correlation controlling for group; post-NR: $\rho = 0.70$, $p < 0.0001$, Figure 6h). These results complement our findings on smoothness and CI.

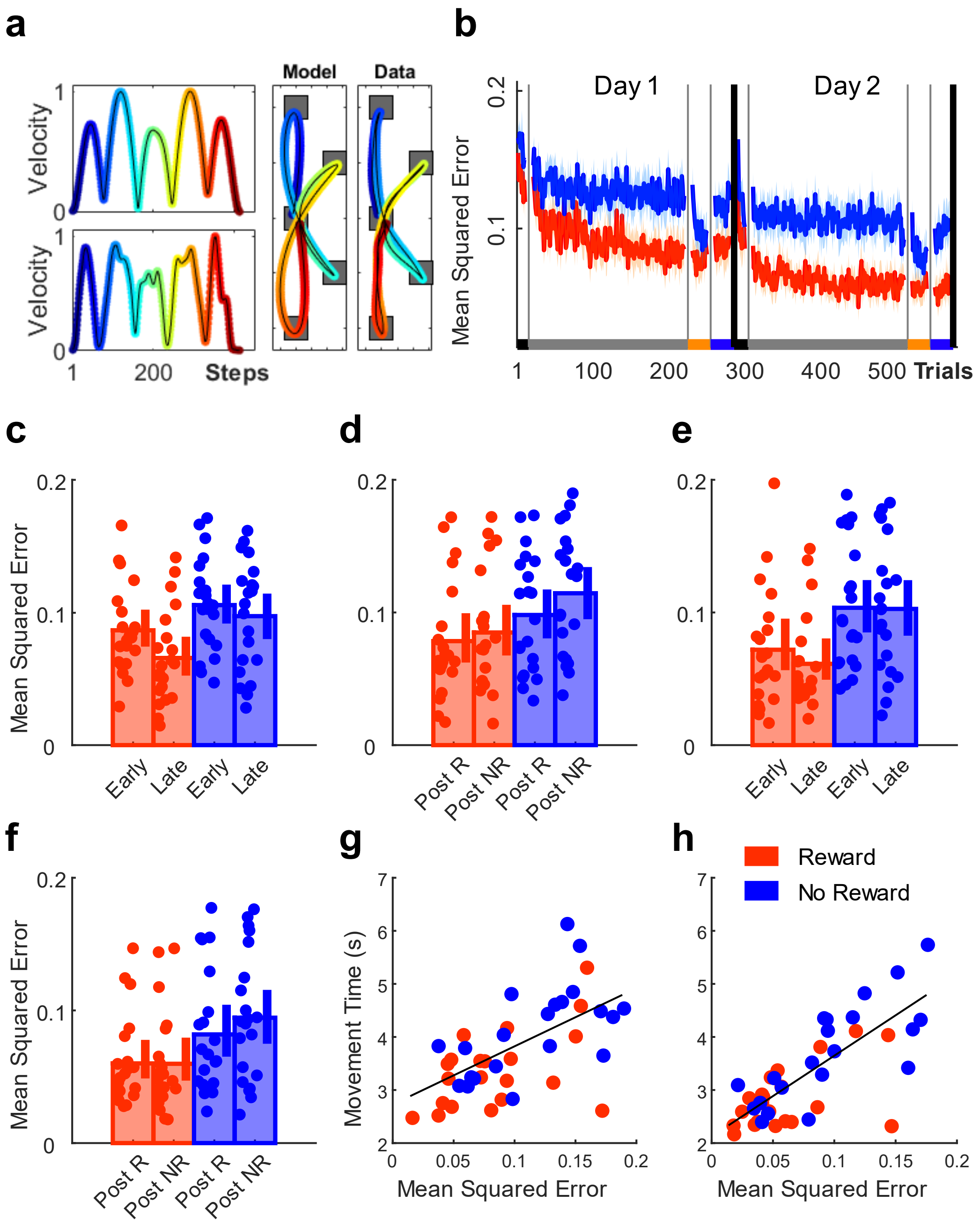


Figure 6 | Progressive alignment to the predictions of a minimum jerk model is facilitated by reward. a) Comparisons between data and the predictions of a minimum jerk model for both trajectory (right panel) and velocity profiles (left panel) for a single participant. **b)** Trial-by-trial changes in mean square error averaged over participants for both groups. **c-f)** Median mean square error for each participant in both groups. **c)** Day 1 early vs late training **d)** Day 1 post assessment (post-R vs post-NR). **e)** Day 2 early vs late training **f)** Day 2 post assessment. **g-h)** Scatterplots displaying the relationship between MT and mean squared error during post NR on **g)** day 1 and **h)** day 2 with a linear line fitted across groups. Shaded regions/error bars represent SEM.

Overall, our results show that over the course of the experiment, performance in the reward group became smoother and showed greater similarity to the predictions of a minimum jerk model. Potentially due to being more energetically efficient [210], participants in the reward group maintained MT performance gains even in the absence of reward.

Spatial reorganisation identifies the final stages of coarticulation and is enhanced by reward

Coarticulation can also be expressed spatially as the radial distance between the peak velocity on the sub-movements and the minimum velocity around the via point (Figure 7a). This distance becomes smaller with high levels of coarticulation [25,26] (Figure 7a; chapter 2, Figure 13c), suggesting that spatial reorganisation reflects the final stages of two movements merging together. We found that the reward group expressed spatial reorganisation with significant decreases in the radial distance between peaks and the via point. No difference between groups in radial distance was observed during baseline (Wilcoxon test; $Z = -1.18$, $p = 0.2371$). However, there was significant main effects for timepoint (mixed ANOVA; timepoint x group; main effect for timepoint, day 1: $F = 7.32$, $p = 0.0100$; day 2: $F = 18.14$, $p < 0.0001$; Figure 7b) as well as group for the second testing day (main effect for group, day 1: $F = 1.95$, $p = 0.1706$; day 2: $F = 7.52$, $p = 0.0091$; Figure 7c,e). This indicates that spatial reorganisation progressively increased across both days, with the reward group showing greater changes on day 2. Across post assessments, the no reward group showed no reward-driven changes in spatial reorganisation, whereas the reward group maintained their performance although the differences between groups was only significant on day 2 (mixed-effect ANOVA; post assessment x group; main effect for group, day 1: $F = 3.08$, $p = 0.0868$; day 2: $F = 5.76$, $p = 0.0211$; Figure 7d,f).

To understand the relationship between CI and spatial reorganisation, we plotted them against each other and detected a pronounced drift in radial distance values (%) with increasing CI levels resulting in a curvilinear shape (Figure 7g).

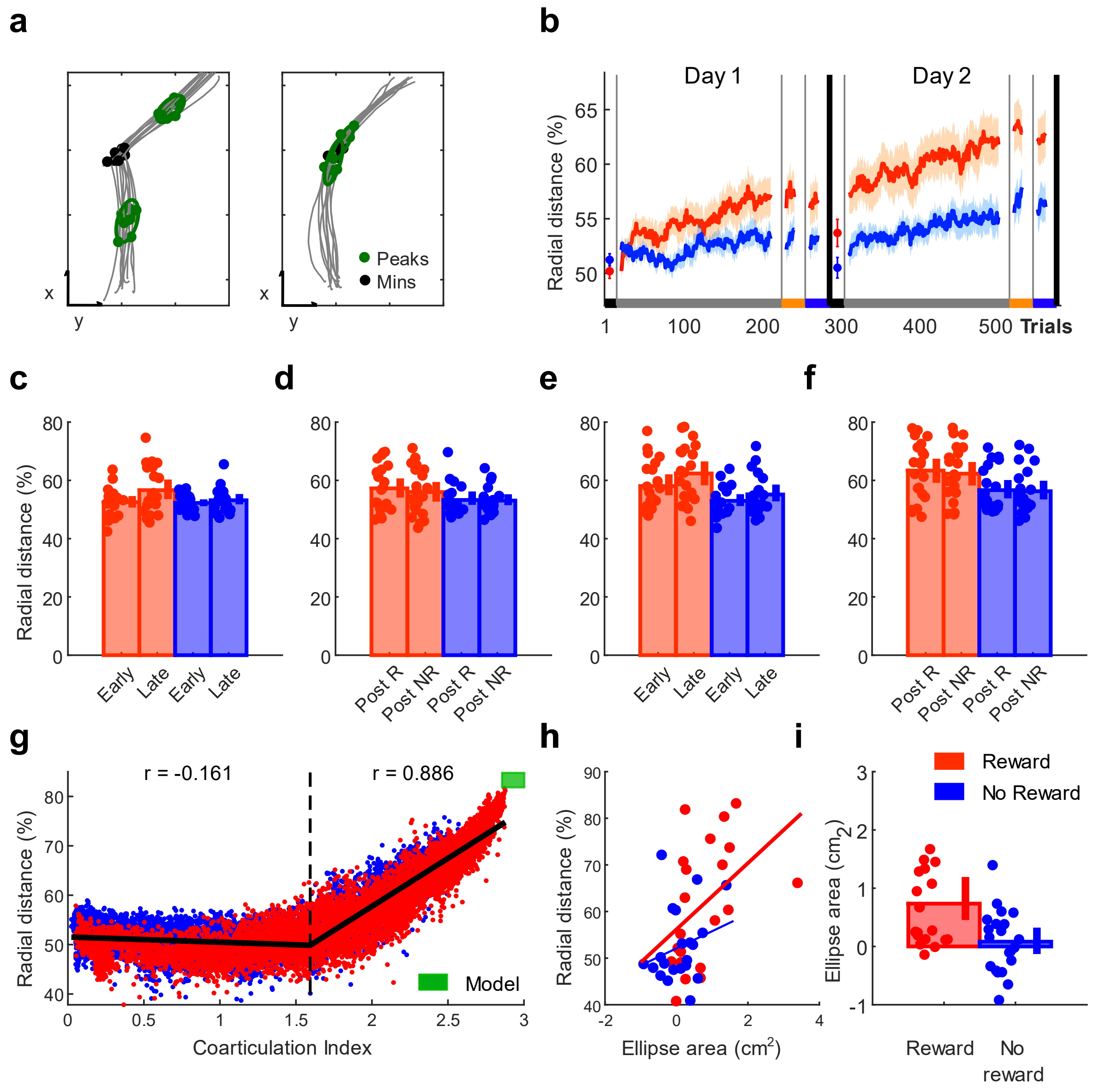


Figure 7| Reward-based improvements in spatial reorganisation. a) Example data for the spatial location of peak velocities when performing two individual (left panel) and one coarticulated (right panel) reaching movement (green and black dots represent peak and min velocities, respectively) **b)** Trial-by-trial changes in radial distance averaged over participants for both groups. **c-f)** Median radial distance for each participant for both groups. **c)** Comparing radial distance during training on day 1 (early vs late). **d)** Post assessment performance (day 1; post-R vs post-NR). **e)** Training (day 2; early vs late). **f)** Post assessment performance (day 2; post-R vs post-NR). **g)** Scatterplot illustrating the relationship between mean CI levels and spatial reorganisation (radial distance %). It includes a two-segment piecewise linear function fitted to the data, while the green area represents the model prediction. **h)** Scatter plot displaying the relationship between variability (ellipse area cm^2) during early training and spatial coarticulation (radial distance %) during late training. **i)** Bar plot comparing the reward and no reward group early training variability including jittered scatter of participant-based median (ellipse area cm^2). Shaded regions/error bars represent SEM.

After fitting a two-segment piecewise linear function to the data, we found an inflection point at ~1.66 (CI) and a strong correlation between CI levels and radial distance values for the second segment (partial correlation controlling for group; segment 1: $\rho = -0.16$; $p < 0.0001$; segment 2: $\rho = 0.89$; $p < 0.0001$).

This suggests that in order to fully coarticulate two consecutive movements, spatial reorganisation is essential [25–27] with this process appearing significantly more pronounced in the reward group. Overall, these findings indicate that improvements in movement efficiency through a quantitative change in how the task is executed may enable the retention of reward-based performance gains.

Within the context of reinforcement learning, it has been suggested that reward leads to improvements in performance via increases in exploration during the early stages of training [82,94]. To assess whether an early increase in spatial variability was associated with improved performance, we measured the area of the confidence ellipses used to determine the radial distance to the via point. We then correlated the ellipse area (cm^2), which was normalised to baseline, over the first 30 trials on day 1 with radial distance (%) over the last 30 trials on day 2 (Figure 7h). We found a significant correlation (partial correlation controlling for group; $\rho = 0.40$; $p = 0.0104$) suggesting that early increases in spatial variability were associated with increases in spatial coarticulation towards the end of the experiment. Comparing ellipse area between groups during early training (first 30 trials), we found that the reward group exhibited higher levels of variability (Wilcoxon test, $Z = 2.52$, $p = 0.0119$, Figure 7i). These results indicate that early reward-driven increases in spatial variability may benefit future coarticulation performance.

Performance gains are maintained across an additional testing day without reward

We next aimed to assess the robustness of these performance gains in an experiment including an additional testing day without reward availability (elongated washout condition) (n=5). In experiment two, participants underwent the same regime as the reward group in experiment 1 on the first two days. On the third day, participants were asked to complete 200 no reward trials. Our results show that even after 24 hours, and over the course of 200 additional unrewarded trials, participants maintained similar MT performance levels. We used a repeated measures ANOVA with timepoint (early vs late across all testing days) as the within factor to assess changes across testing days (repeated measures ANOVA, main effect for timepoint, F = 28.65, p < 0.0001; Figure 8a; Supplementary Figure 3). These results indicate that performance improved over the course of the experiment. However, no changes in MT performance could be observed between late training on day 2 and early training on day 3 (Wilcoxon test, Z = -1.21, p = 0.3016) and between early and late training on day 3 (Wilcoxon test, Z = -1.48, p = 0.4444). Similarly, coarticulation levels appeared stable across the additional testing day without reward (repeated measures ANOVA, main effect for timepoint, F = 19.19 p < 0.0001; Figure 8b; Supplementary Figure 3), with no changes in performance between late day 2 and early day 3 (Wilcoxon test, Z = -0.94, p = 1). Similarly to MT, no changes were observed across day 3 (Wilcoxon test; early vs late; day 3, Z = -0.67, p = 1). In addition, peak velocities were maintained transitioning to and across day 3 (Wilcoxon test, late training day 2 x early training day 3, Z = -1.21, p = 1; early training day 3 x late training day 3, Z = -1.21, p = 1; Figure 8c; Supplementary Figure 3). When assessing changes in smoothness using spectral arc length, we found similar results to experiment 1. Smoothness improved over the course of the experiment (repeated measures ANOVA, main effect for timepoint, F = 57.14 p < 0.0001; Figure 8d; Supplementary Figure 3), while no changes could be observed between late training on day 2 and early training on day 3 (Wilcoxon test, Z = -0.67, p = 1) and between early and late training on day 3 (Wilcoxon test, Z = -1.21 , p = 1). In line with these results, we found that performance aligned progressively with the predictions of the minimum jerk model (repeated measures ANOVA, main effect for timepoint, F = 23.38 p < 0.0001; Figure 8e; Supplementary Figure 3), while no significant changes in similarity could be observed transitioning to and across day 3 (Wilcoxon test, late training day 2 x early training day 3, Z = -2.02, p = 1; early training day 3 x late training day 3, Z = -1.21, p = 1). In addition, we found that participants progressively reorganised their spatial movement output, with a significant decrease in radial distance found across the experiment (repeated measures ANOVA, main effect for timepoint, F = 57.14 p < 0.0001; Figure 8f; Supplementary Figure 3). However, no changes in radial distance could be observed between late training on day 2 and early training on day 3 (Wilcoxon test, Z = -2.02, p = 1) and between early and late training on day 3 (Wilcoxon test, Z = -1.75 , p = 1). These findings support our results from experiment 1 and indicate that improvements in movement efficiency through a quantitative change in how the task is performed may enable long-term retention of reward-based performance gains.

Reward based on performance is most effective in producing behavioural change

Our third experiment was intended to assess whether the observed changes on day 1 were specific to training with reward-based feedback of movement time performance. Participants were allocated to one of the four groups: (1) no reward, (2) reward without performance-based feedback, (3) reward with random feedback and (4) reward with accurate feedback. Only groups 2-4 received monetary reward; reward for groups 1 and 4 were equivalent to that used in experiment 1. Participants within group 3 received feedback about the reward delivered on each trial which was randomly drawn from the feedback experienced during experiment 1. It therefore matched group 4 in terms of reward probability but this did not correspond to actual performance. In contrast, participants in group 2 did not receive any performance-based feedback after completing a given trial. Upcoming reward trials, however, were still cued.

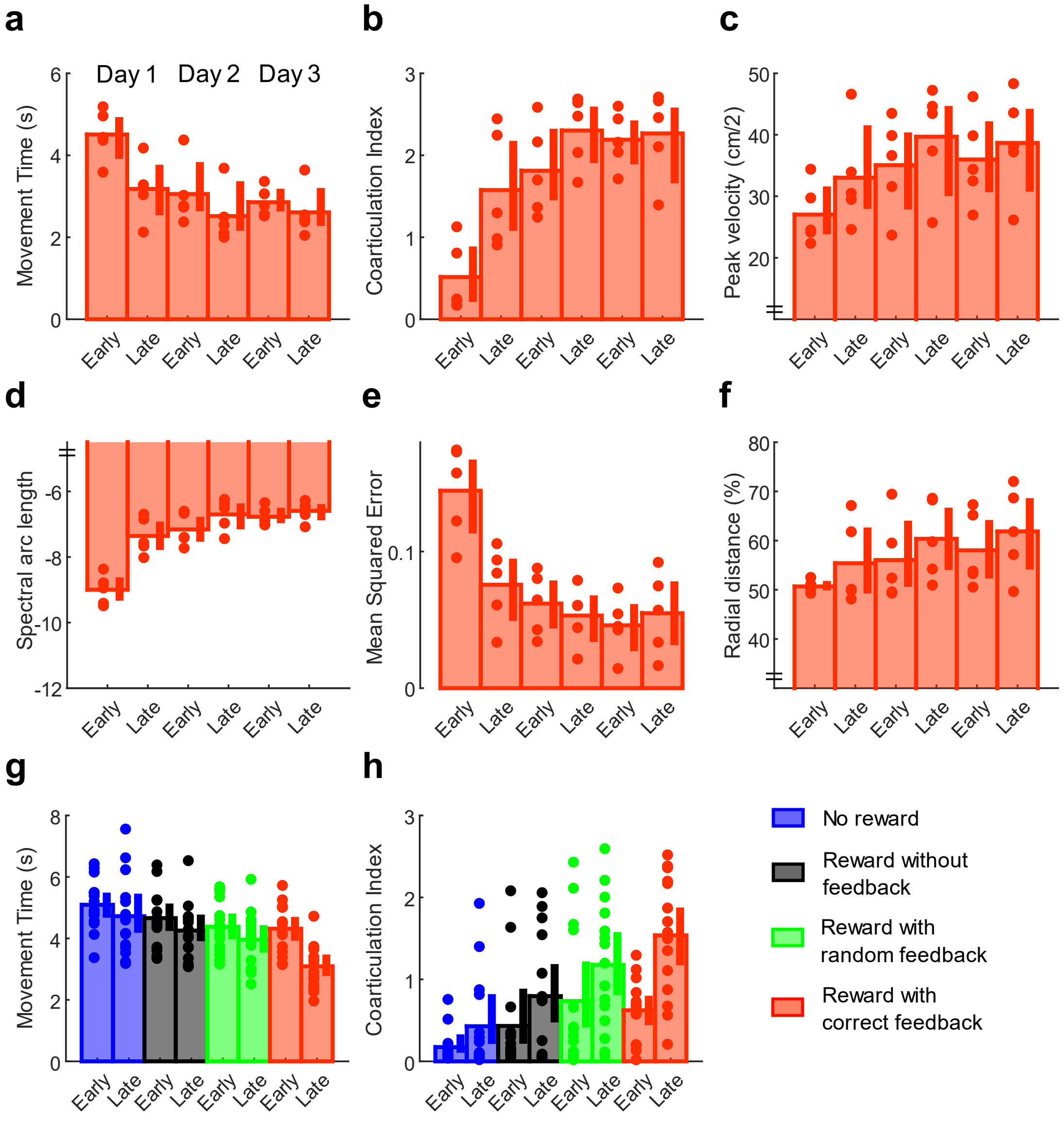


Figure 8| Long-term retention of performance without reward. a-f: Experiment 2. Median data across all 3 days (early (first 20 trials) vs late (last 20 trials) training). **a)** MT **b)** CI level **c)** Peak velocity **d)** Spectral arc length (smoothness) **e)** Mean squared error between data and minimum jerk model prediction **f)** Radial distance (spatial reorganisation) **g-h: Experiment 3. g)** Median MT across groups. **h)** Median CI level. Shaded regions/error bars represent SEM.

Since groups 1 and 4 underwent the same regime as the no reward and reward group in experiment 1 respectively, we were able to test whether our results replicated. Group 4 (reward with accurate feedback) produced the largest improvements in performance, being the only group who showed significant improvements in MT (mixed ANOVA; timepoint x group; interaction, F = 3.76, p = 0.0158; Wilcoxon test; early vs late group 4: Z = 3.40, p = 0.0028; Figure 8g; Supplementary Figure 4) and coarticulation (mixed ANOVA; timepoint x group; interaction, F = 3.18, p = 0.0308; Wilcoxon test; early vs late group 4: Z = -3.24, p = 0.0048; Figure 8h; Supplementary Figure 4). This indicates that the expectation of reward is not sufficient to induce large behavioural improvements, with a link to performance being essential to optimise gains during reward-based training.

These results showed that performance-based reward invigorates sequential reaching. Driven by a reward-based increase in speed, movements also exhibited greater coarticulation, smoothness and a closer alignment to the predictions of a minimum jerk model. Importantly, these performance gains were maintained across multiple days even after the subsequent withdrawal of reward. This highlights the importance of coarticulation to skilful sequential reaching performance, and the potential of this mechanism to produce long-lasting reward-driven improvements in behaviour.

3.4 Discussion

Motor skill learning is integral to everyday life and describes improvements in performance above baseline levels [12]. Improvements in skill have often been assessed at the levels of speed, accuracy and more recently efficiency [211]. In our task, improvements in speed (MT) could be achieved via two strategies: (1) increases in peak velocities of the individual reaching movements and (2) reduction in dwell times around the via points (coarticulation). We were able to show that reward invigorated peak velocities, supporting previous findings on the reward-based invigoration of simple discrete reaching movements [74,75]. Moreover, for the first time, we demonstrate that reward accelerates the naturally slow and difficult to achieve process of coarticulation [25–27,196]. With respect to movement efficiency, these two strategies differ regarding the metabolic costs they incur. Recently, it has been demonstrated that in discrete reaching movements, improvements in peak velocity can be achieved at no cost to accuracy through increased arm stiffness [75]. Although an attractively simple strategy, it comes with a marked escalation in metabolic costs [210]. In fact, it has been shown that in discrete reaching tasks such invigoration is reward-dependent and transient in nature [73–75]. We suggest that this could be due to the energetic demands of the invigorated movement requiring the continued presence of reward to negate this added cost. In contrast, coarticulation enables the performance of sequential movements to become similar to a minimum jerk trajectory through the merging of neighbouring movements into continuous smooth elements [25–28,99]. The subsequent hypothesised improvement in energetic efficiency might explain why the enhanced performance of the reward group became reward-independent. In contrast, performance in the no reward group, who showed little coarticulation, was highly dependent on the availability of reward, i.e., performance gains were lost once reward was removed. Considering that without coarticulation the given reaching sequence will be executed as a series of discrete target reaches, the removal of reward will negate the continued invigoration of performance. This conclusion is in line with previous work showing a transient 'on-off' effect of reward on movement invigoration. Thus, the performance of the no reward group mirrors invigoration patterns observed in simple, discrete reaching tasks.

Despite previous studies demonstrating that reward can enhance retention across a wide variety of sequential and continuous motor tasks, the underlying mechanism for this effect has been described at a very abstract level [158,159,177]. Specifically, it is unclear how reward strengthens a motor memory so that improved performance is maintained even when the incentive is no longer provided. We believe that for reward to induce such long-term improvements in motor skill, it must not only lead to enhanced performance but also improvements in efficiency. A reduction in the cost of this enhanced performance (faster or more accurate movements [73]), enables it to be performed long-term without further incentive/reward. Therefore, reward may not enhance the memory of the action but instead lead to the task being performed in a fundamentally different, and more efficient, manner. In support of this, we found that participants in the reward group produced a quantitative change (spatial reorganisation of the velocity profile between neighbouring movements) in how the task was executed. This highlights that careful analysis of how reward influences the performance of complex actions is essential for dissociating its effects on execution and retention.

Although coarticulation might appear to feature similar movement characteristics as chunking, it is important to emphasise that they represent different processes. Chunking often refers to a series of discrete movements (e.g. button presses in sequence learning paradigms) which are temporally aligned (chunked) over the course of training [20,64,83,189,212]. The chunks are represented at a behavioural level through shorter reaction times between actions, and often allow for a faster execution overall [83,194]. Yet the elements within a chunk are still performed discretely with a clear stop period between them [189,206]. In addition, at a neural level, these elemental movements are planned individually through competitive queuing, highlighting again that some level of independence is maintained [213]. In contrast, coarticulation reflects the merging of neighbouring movements into a single motor primitive that no longer features a pronounced stop period and must run to completion once initiated [25–27,196]. While kinematically appearing to represent a single movement, it remains an open question how these newly formed primitives are represented at a neural level in humans.

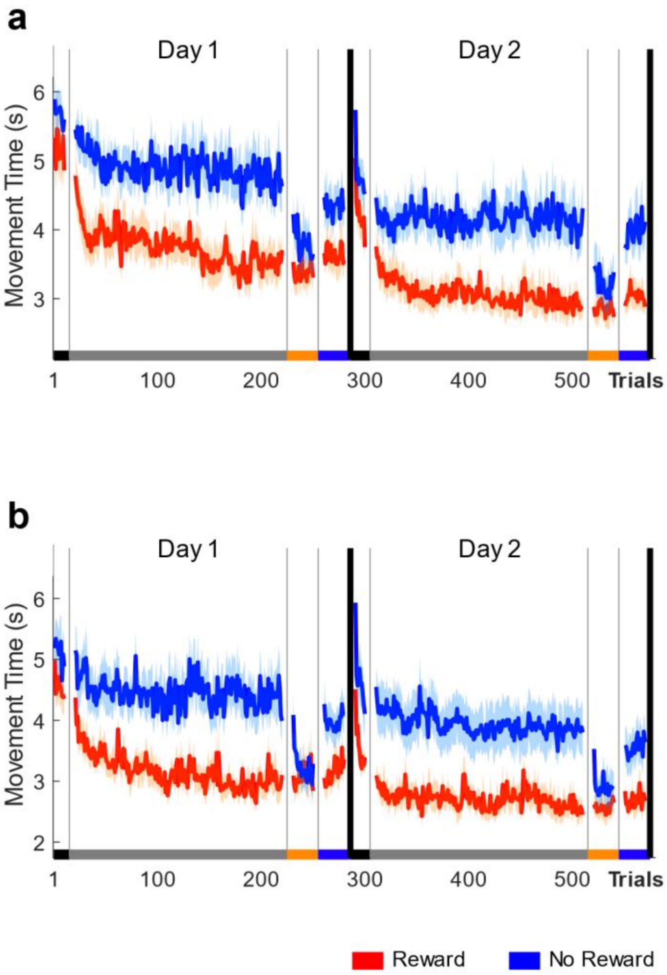
We were able to show that reward facilities the slow and difficult process of coarticulation. Participants in the reward group started to coarticulate after a single day of training which, in comparison to previous work [25–27,194], represents a substantial acceleration in the formation of smooth, continuous reaching movements. Given the link between reward and increased dopamine levels, these results could indicate a critical role for dopamine in coarticulation. In support of this idea, it has been shown in rats that dopamine levels dynamically change with sequence proficiency [62]. In addition, Parkinson's disease (PD) patients OFF medication exhibit specific impairments in movement chunking [1,2]. However, whether dopamine underpins coarticulation similarly to chunking is still unexplored (a notable exception being [22]). Specifically, whether dopamine directly acts on coarticulation or whether any mediating relationship is due to the effects of dopamine on movement invigoration, requires further examination [22].

Here, we show that coarticulation is associated with smoother and, with regards to minimising jerk, more efficient execution. Interestingly, reaching movements performed by stroke patients exhibit reduced smoothness [23,24,214,215], with increases in jerk being due to a decomposition of movement into a series of sub-movements [23,24,214,215]. However, over the course of the recovery process, performance becomes smoother as these sub-movements are progressively blended [23,24,214,215]. Considering this theoretical proximity to the concept of coarticulation, we speculate that stroke recovery and coarticulation may follow similar principles. Consequently, coarticulation facilitated by reward could be a powerful tool in stroke rehabilitation to promote smooth and efficient sequential actions which form an essential component of everyday life activities.

In conclusion, this work highlights that coarticulation could provide a mechanism by which reward can invigorate sequential performance whilst also improving efficiency. This improvement in efficiency through a quantitative change in how the action is performed appears essential for the retention of reward-based improvements in motor behaviour.

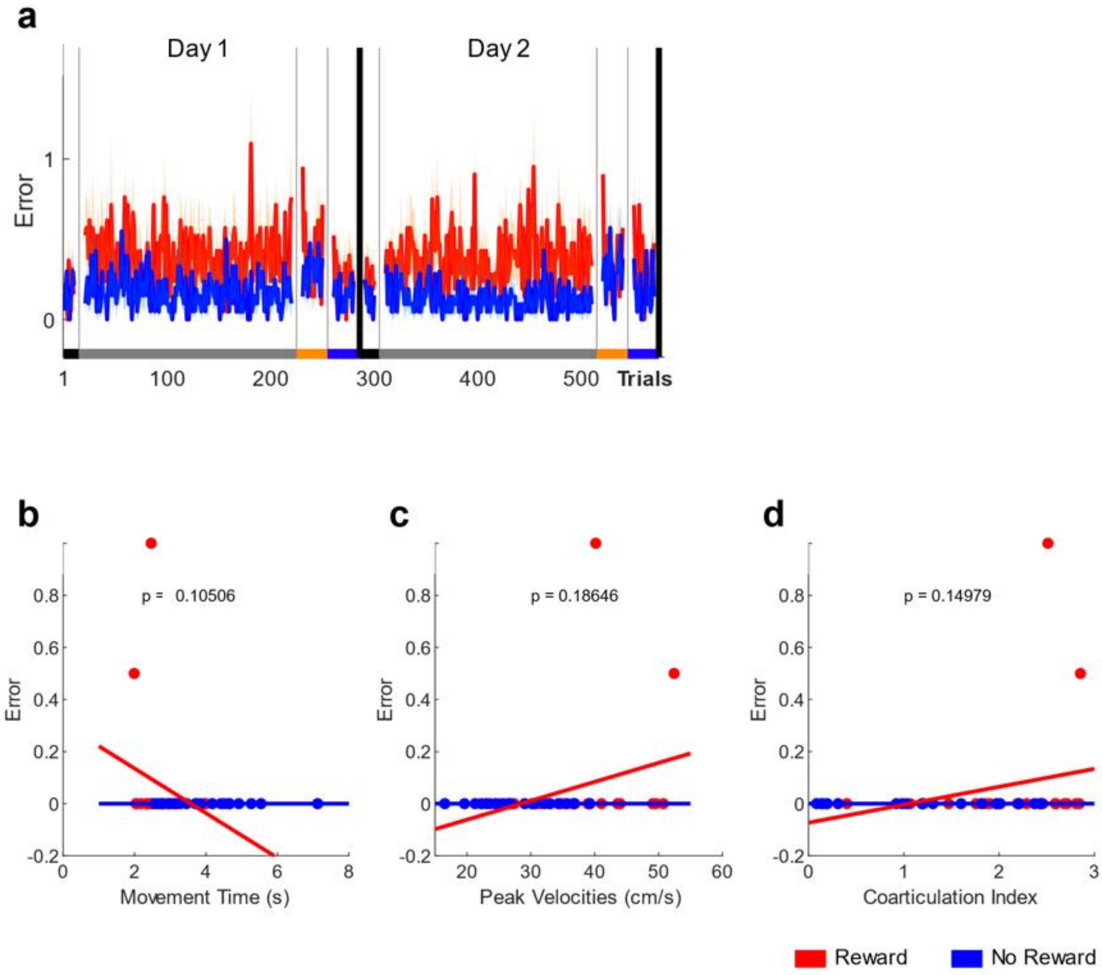
3.5 Supplement

Supplementary Figure 1



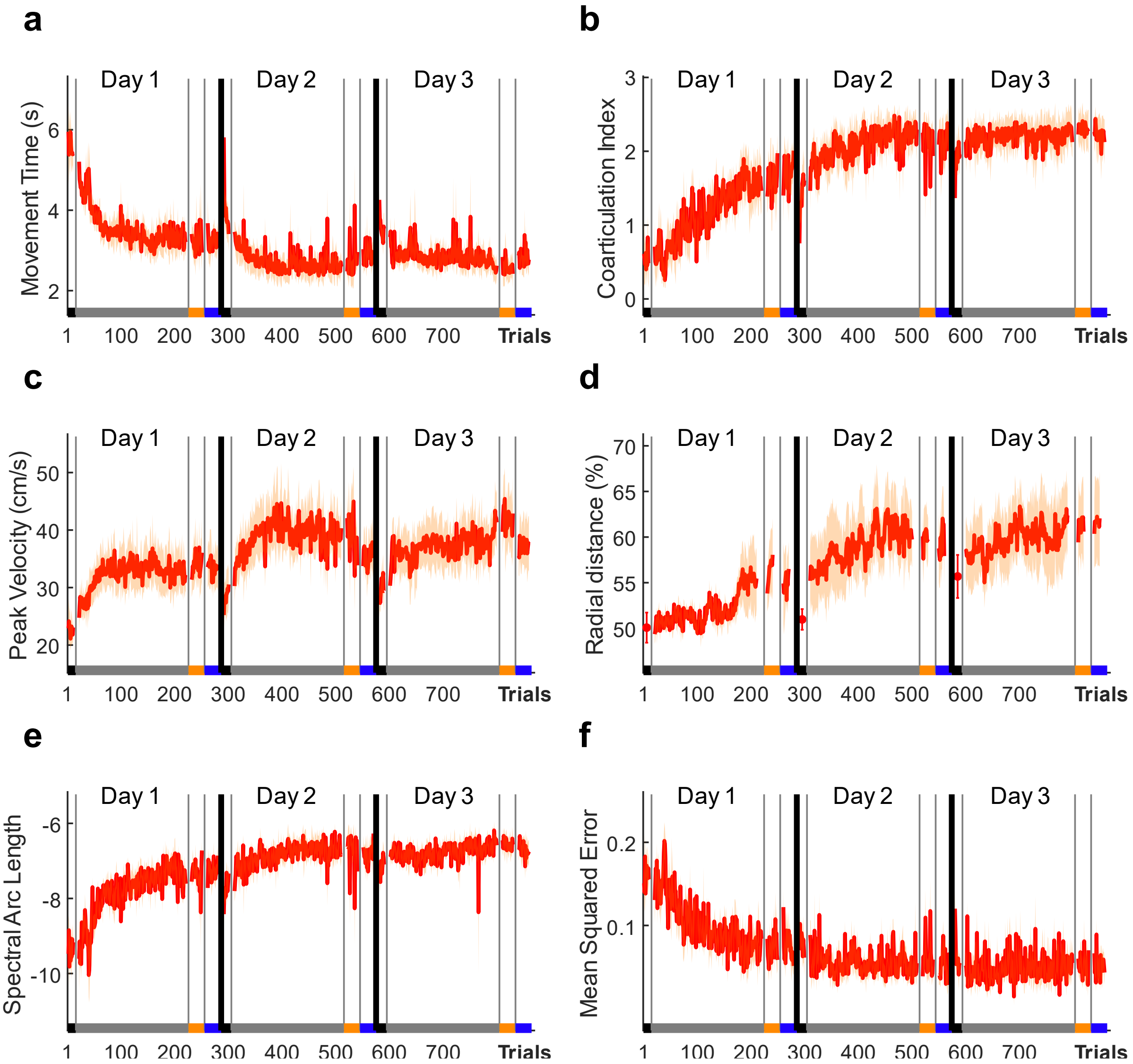
Supplementary Figure 1 | Counterbalancing did not affect MT results. Data was split depending on the counterbalancing. Participants in **a)** completed post-R prior to post-NR, whereas participants encountered the reversed order during their post assessment (**b**).

Supplementary Figure 2



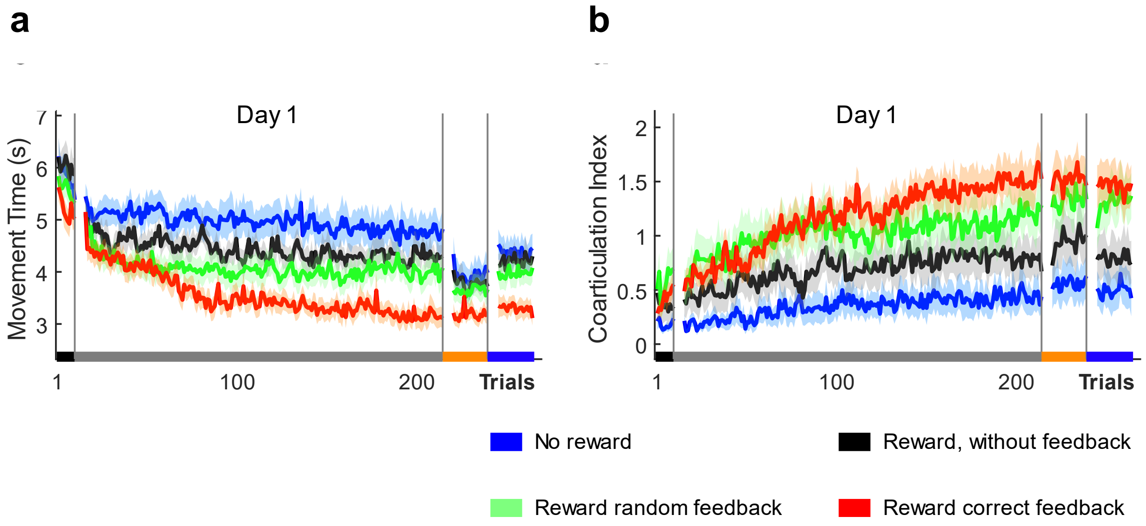
Supplementary Figure 2 | Number of errors did not affect performance during later stages of training. a) Trial-by-trial data showing the number of errors across participants for both groups). **b-d)** Scatterplots comparing median error data during early training with median **(b)** MT, **(c)** peak velocity and **(d)** CI data during late training.

Supplementary Figure 3



Supplementary Figure 3 |Performance gains are stable even over a prolonged washout period. Trial-by-trial data for **a)** MT, **b)** CI levels **c)** peak velocity **d)** radial distance, **e)** spectral arc length, **f)** mean squared error CI for the second experiment which was scheduled on three consecutive days to assess behavioural change over a prolonged washout period. Participants received reward during training on the first two days, however no reward was available on the third day.

Supplementary Figure 4



Supplementary Figure 4 |Reward based on correct performance feedback is most efficient at driving behavioural change. Trial-by-trial data for **a)** MT and **b)** CI levels for the third experiment which included 4 groups that differed with regards to reward availability and feedback type during training to investigate what drove behavioural changes observed in experiment 1.

Chapter 4

The role of dopamine during sequential reaching

4.1 Introduction

Obtaining and maximising reward is a fundamental goal of behaviour and has been shown to both modulate motor vigour and shape learning which has been formally conceptualised as reinforcement learning (RL)[124]. Evidence for its modulatory effect on motor vigour comes from research using saccadic eye movements. It has consistently been demonstrated that reward enhances vigour thereby reducing movement times (MT)[73,155,178]. Expanding on these findings, research found that reward invigorates simple discrete reaching movements leading to improvements in the speed-accuracy trade-off[74,75,156,202]. However, these reward-based improvements in speed and accuracy appear to be transient (i.e., they disappear once reward has been removed), which severely limits the potential use of reward in clinical settings, such as stroke rehabilitation[75]. Recent work has shown that increases in arm stiffness enable faster MTs whilst maintaining similar levels of end-point accuracy[75]. Albeit an attractively simple mechanism to account for the reward-based improvements in the speed-accuracy trade-off, this strategy comes with a marked increase in metabolic cost[209]. Therefore, removing reward might make this strategy unviable, which could explain the pronounced increase in MTs once reward is removed[74,75].

In contrast to simple discrete reaching movements, there is currently a lack of consensus whether reward can invigorate performance on complex sequential movements. In both a serial reaction time task (SRTT) and a force tracking task (FFT), reward did not improve execution via a reduction in MTs[158] or improve accuracy[177], respectively. However, both studies observed an increase in retention, suggesting that reward improved motor retention[158,177]. Yet, a recent study could not replicate these results and instead found that reward did not improve movement execution nor retention on both a SRTT and FFT[159]. Using a continuous sequential reaching task (CSRT), chapter 3 demonstrated that reward improves MTs via an increase in peak velocities and a decrease in dwell times between consecutive reaching movements. Furthermore, reward facilitates the coarticulation of sequential reaching movements leading to these shorter dwell times but also an increase in movement smoothness. Importantly, these reward-based improvements were maintained across an additional testing day without reward. Therefore, these results suggest that reward can improve the speed and efficiency of a sequential reaching action through enhanced coarticulation, and that seems to promote the retention of these reward-based performance gains. However, despite chapter 3 highlighting that both movement time and the learning process involved in coarticulation is reward-sensitive, it is unclear whether the neuromodulator dopamine (DA) underpins these improvements.

A growing body of research has demonstrated that DA plays a central role in processing reward signals[216]. Specifically, while tonic DA appears to modulate motor vigour[167,170,171], phasic DA has been shown to play a role in motor learning[160–166]. A recent study showed that selective depletion of D2 receptors in mice impairs the initiation and vigour of self-paced actions[167]. Additionally, D2 receptors have been suggested to play a role in motor learning. Monkeys that learnt a new discrete motor sequence showed impaired learning when a D2 antagonist (raclopride) was administered prior to training. Similarly, using tiapride, another highly selective D2 antagonist a recent study found that learning on a complex motor sequence task was impaired, hence mirroring previous work on the effect of D2 receptor binding during motor sequence learning[1]. Converging results come from research on motor sequence learning in patients with Parkinson's disease (PD); a neurological condition characterised by severe DA depletion in the striatum. Learning of discrete motor sequence learning in patients OFF medication was impaired, while learning was of similar magnitude ON medication when compared to healthy controls[164,165]. Therefore, DA appears to be involved in both the initiation and vigour of self-paced actions[167] as well as motor sequence learning.

However, despite these advances, it is an outstanding question whether DA and more specifically D2 receptors underpin coarticulation. Upper limb coarticulation represents a hallmark of skilled sequential performance and allows for a smooth and therefore more efficient execution[209]. Evidence for DA playing a role in coarticulation comes from research which has shown that PD patients show impairments in cursive handwriting. Compared to healthy controls, PD patients wrote words as a series of discrete movements with this being independent of movement speed[29]. Additionally, it is unclear whether movement invigoration, which has been shown to be reward-sensitive[167,170,171], is based on a similar DA mechanism as coarticulation or whether these two represent differential neural processes.

Here, we aimed to fill this gap in the literature using haloperidol, a D2 antagonist, to modulate DA levels during our previously established motor sequence learning task (CSRT)[217].

4.2 Methods

Participants: 95 participants (42 males; age range 18-42) were recruited to participate in this experiment, which had been approved by the local research ethics committee of the University of Birmingham. Potential participants were pre-screened and were only invited to the medical exam if they met the following criteria: 1) naïve to the task paradigm ;2) 18–45 years old; 3) no self-reported history of medical disorders; 4) normal or corrected-to-normal vision; 5) no drug allergies; 6) currently taking no medication that interfere with the absorption of haloperidol. Suitable participants were then evaluated by a medical doctor, who reviewed their medical history, evaluated an electrocardiogram taken at rest and took a blood pressure reading. Participants who received medical approval were then scheduled for all experimental sessions. Most participants were self-reportedly right-handed (N = 7 left-handed participants) and gave written informed consent prior to the start of the experiment. Participants were remunerated with money (£18/hour) and were able to earn additional money during the task depending on their performance. Before the start of the experiment, participants were pseudo-randomly allocated to one of the available groups.

Experimental Apparatus: All experiments were performed using a Polhemus 3SPACE Fastrak tracking device (Colchester, Vermont U.S.A; with a sampling rate of 110Hz). Participants were seated in front of the experimental apparatus which included a table, a horizontally placed mirror 25cm above the table and a screen (Figure 1a). The low-latency Apple Cinema placed 25cm above the mirror had a refresh rate of 60Hz and displayed the workspace and participants' hand position (represented by a green cursor – diameter 1cm). On the table, participants were asked to perform 2-D reaching movements. Looking into the mirror, they were able to see the representation of their hand position reflected from the screen above. This setup effectively blocked their hand from sight. The experiment was run using MATLAB (The Mathworks, Natwick, MA), with Psychophysics Toolbox 3.

Task Design: In this experiment we used the same task design as in Chapter 3. In this motor task, participants were asked to hit a series of targets displayed on the screen (Figure 1a). Four circular (1cm diameter) targets were arranged around a centre target ('via target'). Starting in the via target, participants had to perform eight continuous reaching movements to complete a trial. Target 1 and 4 were displaced by 10cm on the y-axis, whereas Target 2 and 3 were 5cm away from the via target with an angle of 126 degrees between them (Figure 1b).

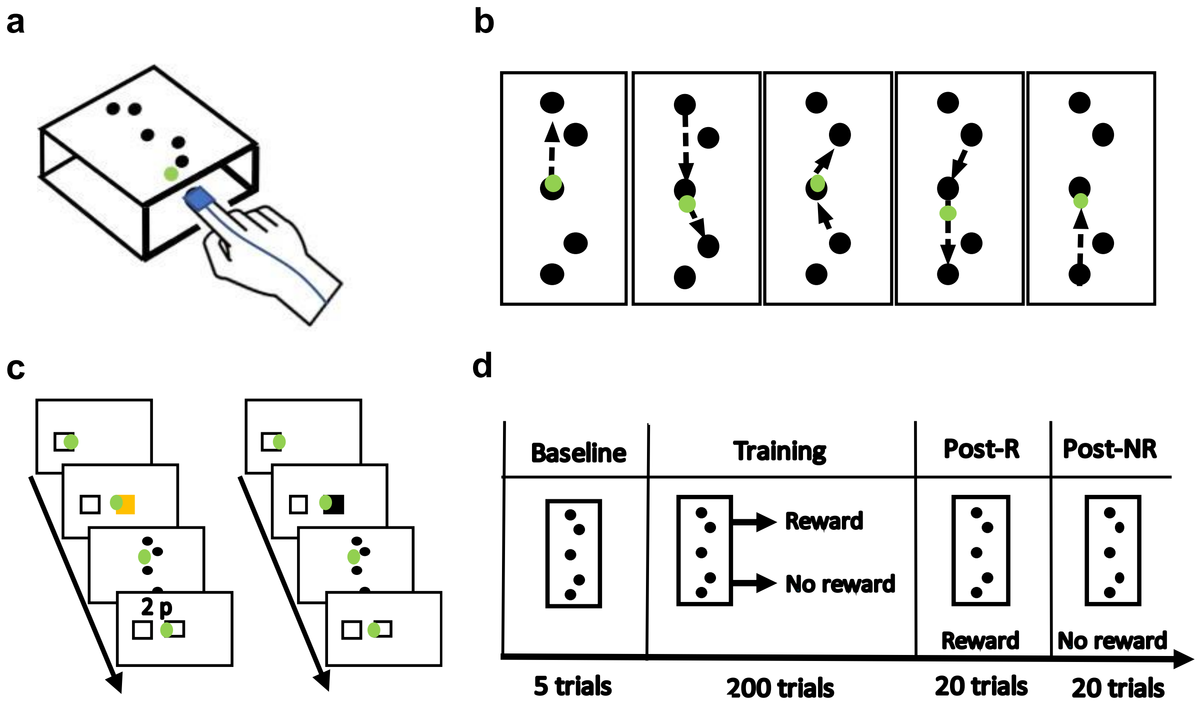


Figure 1 | Experimental Setup and Task Design. a) Participants wore a motion-tracking device on their index finger (Polhemus). **b)** 8 movement sequential reaching task. The participants started from the centre target. **c)** Rewarded trials were cued using a visual stimulus prior to the start of the trial, while feedback was provided at the end (based on a closed-loop design). In no reward trials, participants were instructed to be as fast and accurate as possible. **d)** Study design. Randomly allocated to a reward and no reward group, participants completed 200 trials during training and an additional 20 trials in each post assessment (post-R and post-NR). This design was repeated 24 hours later (day2) as well as 1 week later using a short version of the task (day7).

To start each trial, participants had to pass their cursor though the preparation box (2x2cm) on the left side of the workspace, which triggered the appearance of the start box (2x2cm) in the centre of the screen. After moving the cursor into the start box, participants had to wait for 1.5s for the targets to appear. This ensured that participants were stationary before reaching for the first target. Target appearance served as the go-signal and the start box turned into the via target (circle). Upon reaching the last target (via target), all targets disappeared, and participants had to wait for 1.5s before being allowed to exit the start box to reach for the preparation box to initiate a new trial. Participants had to repeat a trial if they missed a target or performed the reaching order incorrectly. Similarly, exiting the start box too early either at the beginning or at the end of each trial resulted in a missed trial.

Reward Structure and Feedback. We used the same reward structure in this experiment as in our previous experiments (Chapter 3). Participants experienced either reward or no reward trials depending on the current experimental phase: (1) Reward trials were cued using a visual stimulus prior to the start of the trial (Figure 1c). Once participants moved into the preparation box, the start box appeared in yellow (visual stimulus) rather than in black (Figure 1h). Participants were informed that faster MTs will earn them more money, with a maximum amount of 5p available in each trial. While participants moved from the start box to the preparation box to initiate a new trial, the amount earned in the previous trial was displayed on the top of the screen (i.e. 'You have earned 2p out of 5p'). Similarly, to our previous experiment we used a closed-loop design to calculate the amount of reward earned in each trial (see Methods Chapter 1 for more information). When participants started a new experimental block, performance in the first trial was compared to the last 20 trials of the previously completed block. (2) No reward trials were not cued, and no reward was available for participants. However, participants were instructed to 'move as fast and accurately as possible' (Figure 1c).

Task Protocol. Similarly to the previous study, the task protocol included four experimental parts and an initial learning phase was scheduled prior to the start of the experiment (Figure 1d; for more detail see Chapter 2). Additionally, a secondary task was included in this task design, which asked participants to press a force sensor with the index finger of their non-dominant hand. Participants were told to apply pressure in response to an audio signal that changed in amplitude, with higher amplitudes requiring increased force and vice versa. However, the analysis of these secondary-task trials which were scheduled on every 10th trial during training and every 5th trial during the remaining parts are excluded from the analysis presented here.

*Learning***:** We included a learning phase prior to the start of the experiment so that participants could memorise the reaching sequence. This allowed us to attribute any performance gains to improvements in execution rather than memory. After completing 20 learning trials, participants moved on to the main experiment.

*Baseline***:** Participants in both groups completed 10 baseline trials, which were used to assess whether there were any pre-training differences between groups. Both groups were instructed to 'move as fast and accurately as possible', while no performance-based feedback was given at the end of each trial.

Training: Participants in the reward group were informed that during this part they would be able to earn money depending on how fast they complete each trial (200 reward trials). In contrast, participants in the no reward group engaged in 200 no reward trials and were again instructed to move as fast and as accurately as possible.

*Post assessments***:** On both testing days, participants from both groups were asked to complete two post assessments (20 trials each); one with reward trials (post-R) and one with no reward trials (post-NR). The order was counter-balanced across participants.

This design was repeated 24h after the initial session (day2). In addition, we asked participants to return to the lab 1 week after the initial session (day1) to complete a set of post assessments (post-R and post-NR 25 trials each - day7). Consequently, participants completed three experimental sessions (day1, day2 and day7) as well as a medical assessment prior to the start of the study.

Experimental Protocol, Randomisation and Blinding Protocol. In this study, we sought to investigate whether DA modulates both movement vigour (movement time) and coarticulation. To this end, participants were randomly allocated to one of four groups: haloperidol with reward-based feedback (Halo-Rew, N = 25), haloperidol without reward (Halo-NoRew, N = 24), placebo with reward-based feedback (Ctrl-Rew, N = 23) and placebo without reward (Ctrl-NoRew, N = 23) after the medical assessment. Due to the required testing environment this study was single-blind, and both the medical doctor and examiner were aware of the drug group allocation (haloperidol vs placebo). However, to reduce bias all participants were told that they will receive either a placebo tablet or the active drug (haloperidol). Similarly, all participants had to complete a health check on the day of drug/placebo intake (day1) and were checked by the medical doctor in intervals of 1h throughout day1. Additionally, all task instructions were displayed on screen instead of communicated verbally to further reduce bias. The administration of both the active drug and the placebo tablet was performed by the medical doctor.

On the day of drug/placebo intake (day1), participants either received 2.5mg of haloperidol (2 x 0.5mg and 1 x 1.5mg tablet) or three lactose tablets of the same white colouring. In each case, participants were handed an envelope containing either the active drug or placebo and were asked to close their eyes during intake. Haloperidol is a D2-receptor antagonist that shows a limited affinity to D1 receptors and has superior in vivo D2 binding. In addition, it blocks DA D2 binding in the basal ganglia (BG) but not in the prefrontal cortex and as such can be considered to selectively modulate DA levels within the BG pathway[148]. To coincide with the peak plasma concentration, participants were asked to wait in the lab for 120min before engaging in the motor task[148]. The chosen dose of 2.5mg of haloperidol and the waiting time of 2h were similar to previous studies that were able to observe drug-related behavioural and neurophysiological effects of haloperidol[5,148]. After the waiting period, participants were asked to complete the motor task (see Task Design and Task Protocol for details) and upon completion were yet again checked by the medical doctor. Participants were scheduled to return to the lab 24h later (day2) to complete the same motor task again; this time without any drug/placebo manipulation. However, participants received the same feedback as during day1 (i.e., with or without reward-based feedback). Lasty, participants engaged in a short version of the motor task one week after the initial session (day7) and were subsequently debriefed.

Data Analysis. Analysis code will be available on the Open Science Framework website, alongside the experimental datasets at https://osf.io/62wcz/. The analyses were performed in Matlab (Mathworks, Natick, MA) and JASP. We used the same parameters to assess performance as in Chapter 1 (for more information on the individual parameters see chapter 2). Dual-task trials (i.e., trials during which participants engaged in a secondary task) were removed from the analysis.

Movement Time (MT): MT was measured as the time between exiting the start box and reaching the end target. This excludes reaction time, which describes the time between target appearance and when the participants' start position exceeded 2cm.

Peak Velocity: Through the derivative of positional data (*x, y*), we obtained velocity profiles and found the peak velocities of each reaching movement. These were then averaged to provide a peak velocity measure for each trial.

Coarticulation Index (CI): To measure coarticulation, we compared the mean peak velocities of the two sequential reaches with the minimum velocity around the via point. The smaller the difference between these values, the greater coarticulation had occurred between the two movements[193]. We normalised the obtained difference, ranging from 0 to 1, with a CI value of 1 indicating a fully coarticulated movement.

Minimum-jerk model: We used the minimum-jerk model to model trajectories for each angle configuration used in this experiment using the Matlab code provided by Todorov and Jordan (1998)[28]. We then calculated the mean square error (immse function in Matlab) between the predicted and actual velocity profile, which were both normalised and interpolated (N = 500), to estimate the fit on a trial-by-trial basis. Due to the two-dimensional structure of trajectories, we used velocity profiles rather than the trajectories for this comparison.

Spectral Arc Length: To assess movement smoothness, we used spectral arc length as our smoothness metric. Spectral arc length has been shown to be less sensitive to differences in MT and more sensitive to changes in smoothness [85,87]. The spectral arc length is derived from the arc length of the power spectrum of a Fourier transformation of the velocity profile. We used an open-source Matlab toolbox to calculate this value for each trajectory [198].

Statistical Analysis. We excluded participants that on average took longer than 7s to complete a trial during Baseline (day1). This was required to ensure that all groups had a similar level of MT performance during baseline on day 1. Without this, it would have made the interpretation of group differences very difficult, even with normalisation procedure described below. This resulted in 8 participants being removed from further analysis and led to changes in the group sizes (Halo-Rew, N = 0, Halo-NoRew, N = -4, Ctrl-Rew, N = 0, Ctrl-NoRew, N = -4). In addition, data was normalised by performing a baseline correction for each participant and parameter Specifically, for each participant, we subtracted average performance during Baseline (day1) from all subsequent data points across all days.

We used linear mixed-models (LMM) to assess statistical significance of our results. We carried out separate analyses for training on day1 (Training1) and day2 (Training2). Similarly, a separate analysis was conducted for each set of post assessments (including both post-R and post-NR), which were scheduled on day1 (Post1), day2 (Post2) and day7 (Post7). Specifically, the LMM used throughout was:

*Dependent Variable ~ 1 + Reward *Drug *Trial Number + (1|Participant)*

This included fixed main effects for both reward (reward; no reward), drug (haloperidol; placebo) and trial number (number of trial), while accounting for individual differences using individual intercepts (random effect). In our model, we further included all possible interaction terms between the three main effects (i.e., *Reward *Drug, Reward *Trial Number, Drug *Trial Number, Reward *Drug *Trial Number*). This allowed us to assess the effect of both reward and drug condition on performance (dependent variable) across trials. How to best investigate significant interactions observed within LMM has been a heated topic[218]. First, additional LMM were performed to investigate specific interactions such as Reward * Trial Number. To further understand such interactions, median performance over the first 30 trials during training (early) was compared with the median performance over the last 30 trials (late) using Wilcoxon tests with false discovery rate (FDR) being used to correct for multiple comparisons and effect sizes being represented by η^2. The alternative would have been to perform Wilcoxon tests for each individual trial (to determine where Reward and Drug differed) however this was perceived to be an excessive number of statistical tests. Therefore, although a simplified analysis, the comparison of early vs late provided an insight into how reward and drug influenced the changes across trials (as measured by the LMM). 30 instead of 20 trials (as in Chapter 3) were used in this analysis to account for the increased noise in the data (i.e., clinical data). We used one-sample Kolmogorov-Smirnov tests to test our data for normality and found that all measures were non-parametric. Consequently, median values were used as input in all Wilcoxon tests. Linear partial correlations (fitlm function in Matlab) were used to measure the degree of association between the chosen variables, while accounting for the factor reward and drug.

4.2 Results

Across all output measures described in the following section, no significant differences between groups were found during baseline.

Haloperidol impaired reward-based invigoration of sequential reaching movements.

Our results converge with our previous findings that reward invigorates sequential reaching performance (Figure 2a). For both days, we found a significant improvement in MT across training (Training1: $\beta = 0.009$, CI = [0.007, 0.01], $t = -5.4$, $p < 0.0001$; Training2: $\beta = -0.003$, CI = [-0.004, -0.002], $t = -4.2$, $p < 0.0001$) as well as a significant difference in performance between the reward and no reward groups (group; Training1: $\beta = 1.67$, CI = [0.567, 1.8], $t = -9.3$, $p < 0.0001$; Training2: $\beta = 0.91$, CI = [-0.35, 2.17], $t = -6.7$, $p < 0.0001$). Importantly, our analysis revealed a significant interaction between group x trial number x drug on day1 (group x trial number x drug interaction: $\beta = 0.004$, CI = [0.004, 0.005], $t = 2.29$, $p = 0.024$). Fitting two LMM for both the reward and no reward groups, we found a significant interaction between drug and trial number only for the reward groups (trial number x drug interaction: $\beta = -0.003$, CI = [-0.003, -0.002], $t = -2.18$, $p = 0.034$, Figure 2b). This suggests that over the course of Training1, haloperidol selectively affected performance of the reward group. In contrast, haloperidol did not have a significantly negative effect on MT performance in the no reward group. In a subsequent post-hoc test, changes in MT performance between early and late training were analysed. For each participant, the average median MT for each timepoint (early and late training) were subtracted from one another. This comparison confirmed the aforementioned findings (Wilcoxon test; $Z = 2.21$, $p = 0.0272$, $\eta^2 = 0.32$, Figure 2b) and highlighted that haloperidol was associated with a lack of MT improvement only for the reward group. Additionally, we found no significant group x trial number x drug interaction for Training2 ($\beta = -0.003$, CI = [-0.003, -0.002], $t = -1.18$, $p = 0.084$).

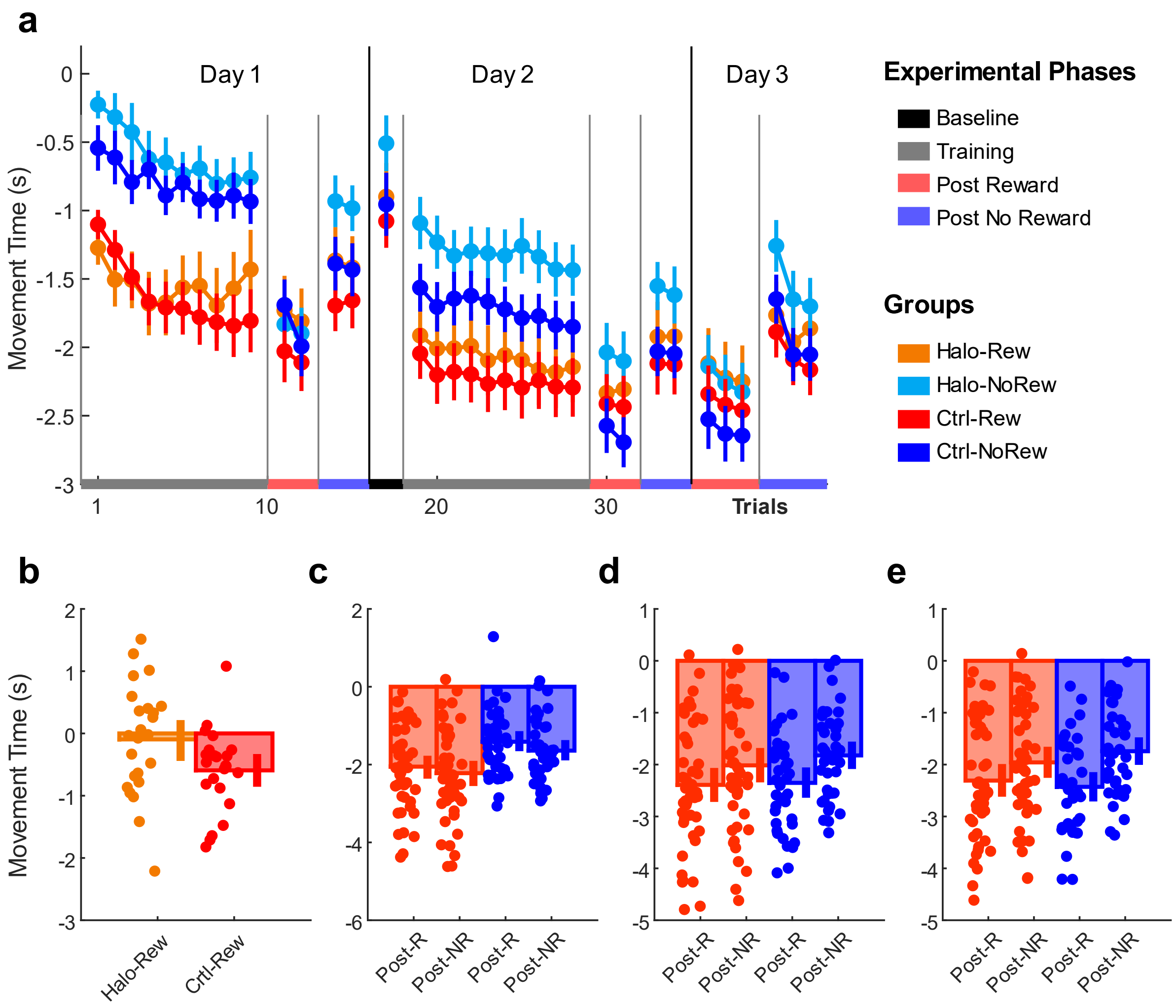


Figure 2 | Haloperidol impaired reward-based invigoration of sequential reaching movements. a) Changes in MT presented in epochs of 20 trials averaged over participants for all groups: **b)** Changes in median MT performance form early to late Training1 for both reward groups. Median MT performance for each group during **c)** Post1, **d)** Post2 and **e)** Post7. Error bars represent SEM.

Additionally, we found no significant group x trial number x drug interaction for Training2. Instead, we found a significant drug effect (β = -0.017, CI = [-1.25, 1.22], t = 3.18, p = 0.002). These results indicate that haloperidol impaired reward-based invigoration of MTs during day 1 but subsequently led to a global slowing in MT performance across groups on day 2. Similarly to our previous results, we observed a significant interaction between trial number (post-R vs post-NR) and group for all post assessments (trial number x group; interaction, Post1: β = -0.031, CI = [-0.059, -0.007], t = 3.38, p = 0.001; Post2: β = -0.0002, CI = [-0.021, 0.021], t = 2.25, p = 0.027; Post7: β = -0.015, CI = [-0.032, -0.0006], t = 3.47, p < 0.0001 Figure c-e). Specifically, we found significant changes in MT performance from post-R to post-NR for the no reward groups (Wilcoxon test; Post1: Z = -3.4, p = 0.0034, η^2 = 0.38; Post2: Z = -2.6, p = 0.0421, η^2 = 0.29; Post7: Z = -3.05, p = 0.0092, η^2 = 0.35) but not the reward groups (Wilcoxon test; Post1: Z = -1.7, p = 0.1796, η^2 = 0.17; Post2: Z = -1.7, p = 0.1853, η^2 = 0.17; Post7: Z = -1.7, p = 0.1939, η^2 = 0.17). This suggests that performance was more stable across reward regimes for the reward groups, whereas MTs changed significantly in the no reward groups depending on reward availability.

Reward-based invigoration of peak velocities was impaired under haloperidol.

Over the course of Training1 and Training2, peak velocities increased (trial number; Training1: β = -0.06, CI = [-0.08, -0.06], t = 3.0, p = 0.004; Training2: β = 0.04, CI = [0.03, 0.05], t = 3.44, p < 0.0001) with the reward groups showing significantly higher peak velocities throughout both training blocks (group; Training1: β = -9.48, CI = [-15.76, -3.22], t = 7.7, p < 0.0001; Training2: β = -1.83, CI = [-10.73, 7.08], t = -5.8, p < 0.0001). Similarly to our MT results, we found a significant interaction between group, trial number and drug on day1 (group x trial number x drug interaction: F = 5.75, p = 0.0189, η^2 = 0.016). Fitting two separate LMM for both the reward and no reward groups, we found a significant interaction between drug and trial number for the reward group (drug x timepoint interaction: β = 0.024, CI = [0.021, 0.028], t = -2.06, p = 0.045, Figure 3a). In contrast, no significant results were found for the no reward groups (drug x timepoint interaction: β = 0.002, CI = [0.0019, 0.022], t = -1.06, p = 0.075). In line with our MT results, this suggests that haloperidol selectively impaired reward-based invigoration of peak velocities, which ultimately led to slower MTs. A subsequent post-hoc test comparing changes in peak velocities between reward groups over Training1 confirmed this conclusion (Wilcoxon test; Z = -2.31, p = 0.0208, η^2 = 0.34, Figure 3b).

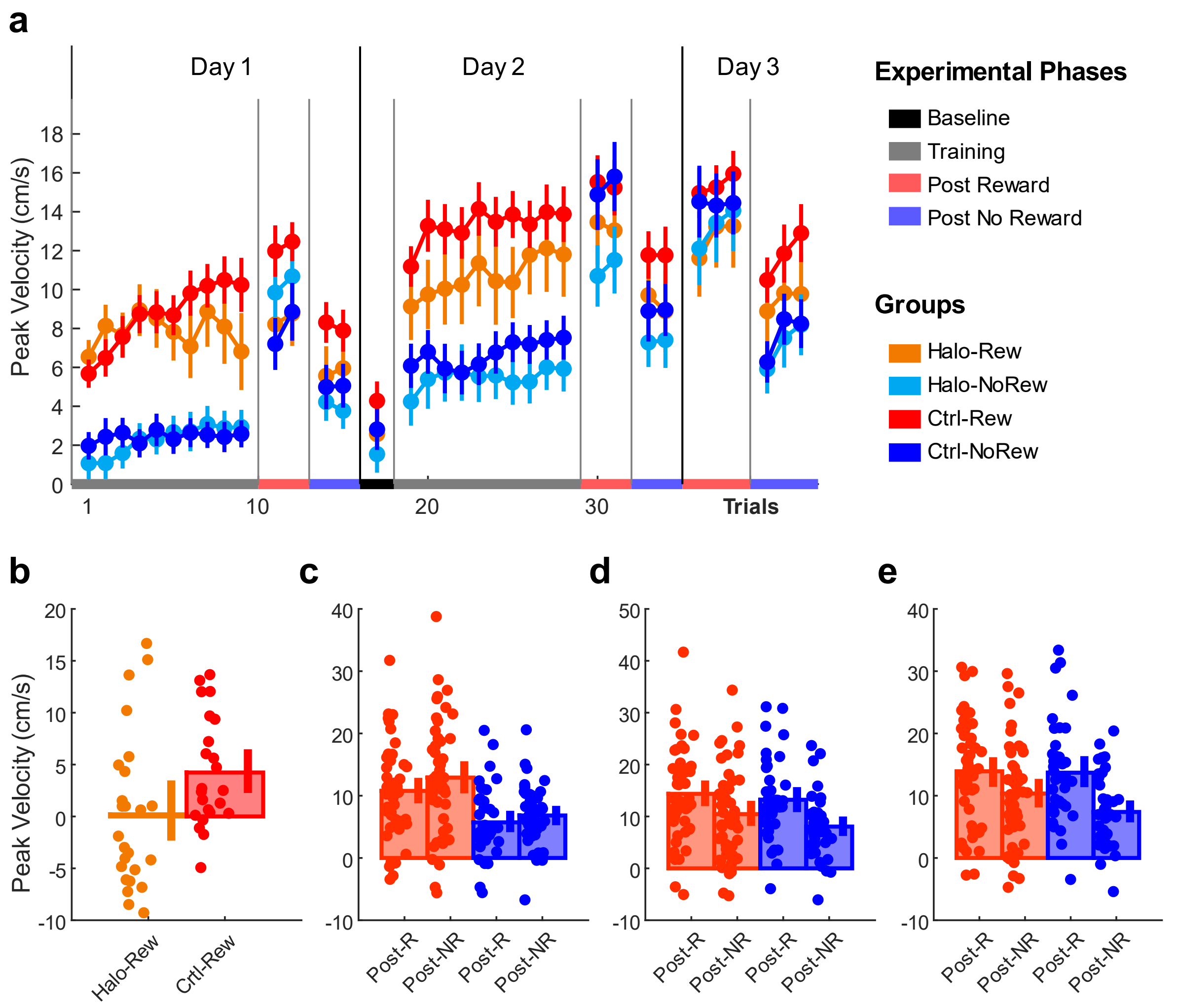


Figure 3 | Reward-based invigoration of peak velocities is impaired with haloperidol. a) Changes in peak velocities presented in epochs of 20 trials averaged over participants for all groups: **b)** Changes in median MT performance form early to late Training1 for both reward groups. Median peak velocities performance for each group during **c)** Post1, **d)** Post2 and **e)** Post7. Error bars represent SEM. Note negative values are possible due to baseline corrections.

Furthermore, our analysis did not reveal a significant group x trial number x drug interaction during Training2. However, in line with the MT results, we found a significant drug effect (β = 5.31, CI = [-3.4, 14.02], t = -2.15, p = 0.033). These findings converge with the MT results and indicate that haloperidol impaired reward-based movement invigoration via a reduction in peak velocities on day 1, and led to global slowing across groups on day 2.

Across post assessments, we found a significant group x trial number interaction for Post1 and Post2 (trial number x group; interaction, Post1: β = 0.18, CI = [-0.0008, -0.36], t = -2.79, p

$= 0.006$; Post2: $\beta = 0.09$, CI = [-0.1, 0.27], $t = -2.39$, $p = 0.019$, Figure 3c-e). Specifically, we found significant changes in MT performance from post-R to post-NR for the no reward groups (Wilcoxon test; Post1: $Z = 3.3$, $p = 0.0039$, $\eta^2 = 0.37$; Post2: $Z = 3.49$, $p = 0.0019$, $\eta^2 = 0.39$) as well as the reward groups (Post1: $Z = 2.39$, $p = 0.0341$, $\eta^2 = 0.24$; Post2: $Z = 2.31$, $p = 0.0424$, $\eta^2 = 0.24$). Albeit, less significant and smaller in effect size, these results suggest that similarly to the no reward groups, performance in the reward groups changed according to reward availability.

Haloperidol impaired coarticulation without reward-based feedback.

Our results converge with our previous findings that reward facilitates coarticulation of sequential reaching movements (Figure 4a,b). For both days, we found a significant improvement in CI levels across trials (trial number ; Training1: $\beta = 0.002$, CI = [0.001, 0.003], $t = 6.03$, $p < 0.0001$; Training2: $\beta = 0.003$, CI = [0.001, 0.0004], $t = 3.92$, $p < 0.0001$) as well as a significant difference in performance between the reward and no reward groups (group; Training1: $\beta = -1.0$, CI = [-1.66, -0.35], $t = 4.97$, $p < 0.0001$; Training2: $\beta = -1.39$, CI = [-2.28, -0.5], $t = 4.35$, $p < 0.0001$). Crucially, we found a significant interaction between group and drug for both days (group x drug interaction; Training1: $\beta = 0.42$, CI = [0.007, 0.84], $t = 2.85$, $p = 0.005$; Training2: $\beta = 0.59$, CI = [0.02, 1.16], $t = 2.51$, $p = 0.013$, Figure 4b,c). Specifically, we observed a significant difference in CI levels between the no reward groups on day2 (Wilcoxon test early vs late; Training1: $Z = -1.8$, $p = 0.119$, $\eta^2 = 0.21$; Training2: $Z = -2.8$, $p = 0.0118$, $\eta^2 = 0.31$). In contrast, we did not find differences in coarticulation for the reward groups on either day (Training1: $Z = 1.03$, $p = 0.3028$, $\eta^2 = 0.11$; Training2: $Z = 0.77$, $p = 0.4391$, $\eta^2 = 0.08$). Despite the no reward group differences failing to reach significance for Training1, this trend

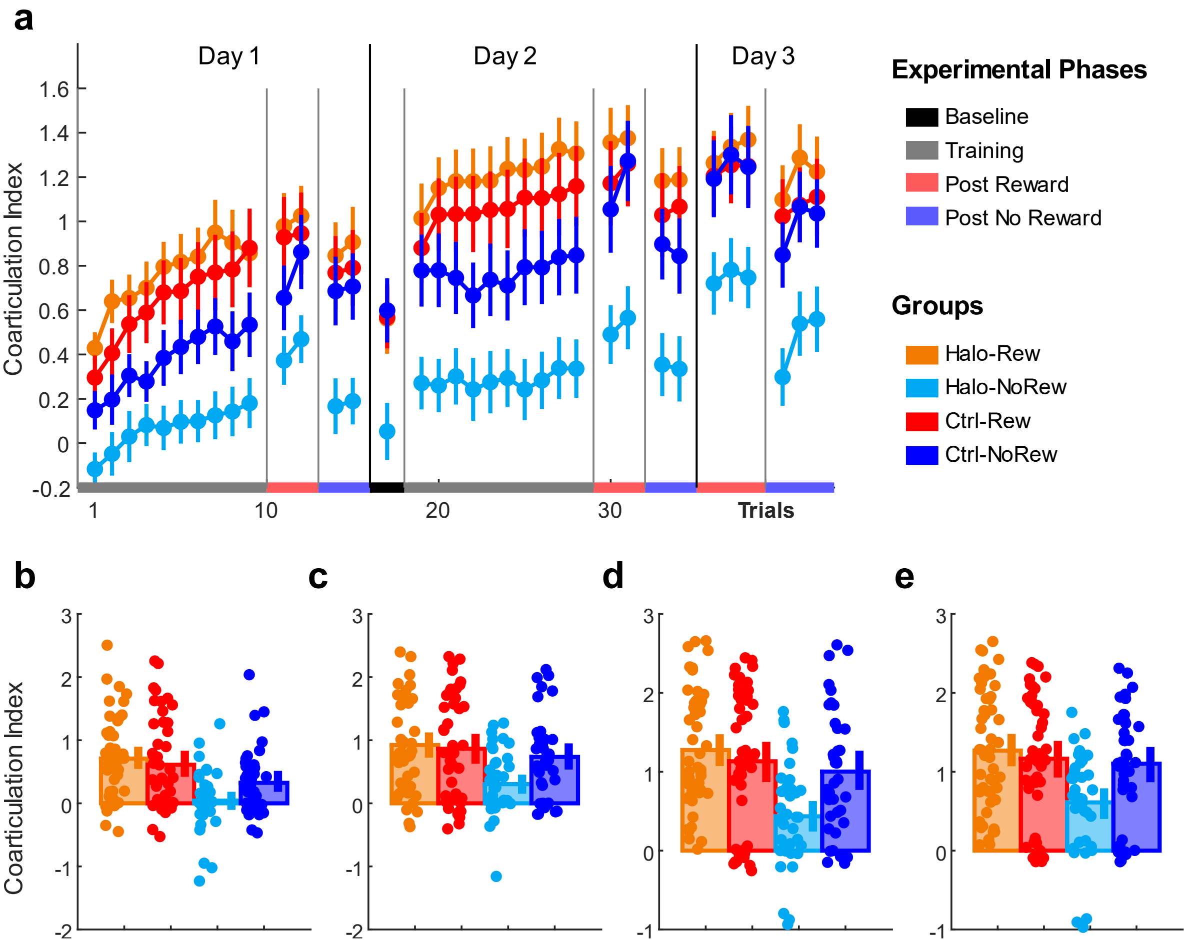


Figure 4 | Haloperidol impaired coarticulation without reward-based feedback. a) Changes in CI levels presented in epochs of 20 trials averaged over participants for all groups: Median peak velocities performance for each group during **b)** Training1 **c)** Training2, **d)** Post2 and **e)** Post7. Error bars represent SEM. Note negative values are possible due to baseline corrections.

suggests that haloperidol impaired coarticulation in the no reward group during initial learning (day1) and led to a progressively larger deficit across days (Figure 4a). While performance differences between groups were persistent across all post assessments (group; Post1: β = -1.18, CI = [-2.05, -0.31], t = -4.06, $p < 0.0001$; Post2: β = -1.49, CI = [-2.44, -0.54], t = 3.03, p = 0.003; Post7: β = -1.17, CI = [-2.08, -0.26], t = 2.84, p = 0.006), we also found a significant group x drug interaction for Post2 (β = 0.68, CI = [0.07, 1.28], t = 2.15, p = 0.035, Figure 4d). Yet again, we observed a significant difference in CI levels between the no reward groups (Wilcoxon test: Z = -2.89, p = 0.0076, $\eta^2 = 0.33$) but not the reward groups (Z = 0.48, p

= 0.631, η^2 = 0.04). Additionally, across Post7 the group x drug interaction was close to significance (β = 0.51, CI = [-0.04, 1.09], t = 1.81, p = 0.064, Figure 4e). These findings indicate that haloperidol impaired coarticulation when no reward was available; an impairment which became more apparent over time. In contrast, haloperidol had no effect when reward-based feedback was provided.

Increases in movement smoothness were driven by reward but were not modulated by haloperidol.

On both days, movement smoothness improved significantly across training (trial number; Training1: β = -0.001, CI = [-1.83, 0.3], t = 4.5, p < 0.0001; Training2: β = 0.004, CI = [0.002, 0.005], t = 5.1, p < 0.0001, Figure 5a). Similarly, to previous results, we found that the reward groups produced significantly smoother movements throughout both days (group; Training1: β = -0.76, CI = [-0.003, 0.008], t = 4.2, p < 0.0001; Training2: β = -0.57, CI = [-1.78, 0.65], t = 4.6, p < 0.0001). However, in contrast to the CI results, we failed to observe a significant interaction between group x drug on either day (Figure 5b,c). Specifically, we found no evidence that haloperidol modulated movement smoothness during training. In line with these results, we found no group x drug interaction across all post assessments. Despite a lack of significance, smoothness was lowest in the no reward group that received haloperidol, which is especially prominent on day2. Considering that our previous results suggest that coarticulation is associated with movement smoothness, we aimed to assess whether this holds true in the current experiment. Correlating median CI and smoothness values for both early and late Training1 and Training2, we found a strong association between them (partial correlation controlling for group and drug; ρ = 0.64; p < 0.0001, Figure 5d,e). This finding highlights coarticulation is related to smoothness. However, here we did not observe that drug-induced impairments in CI levels also led to a decrease in smoothness.

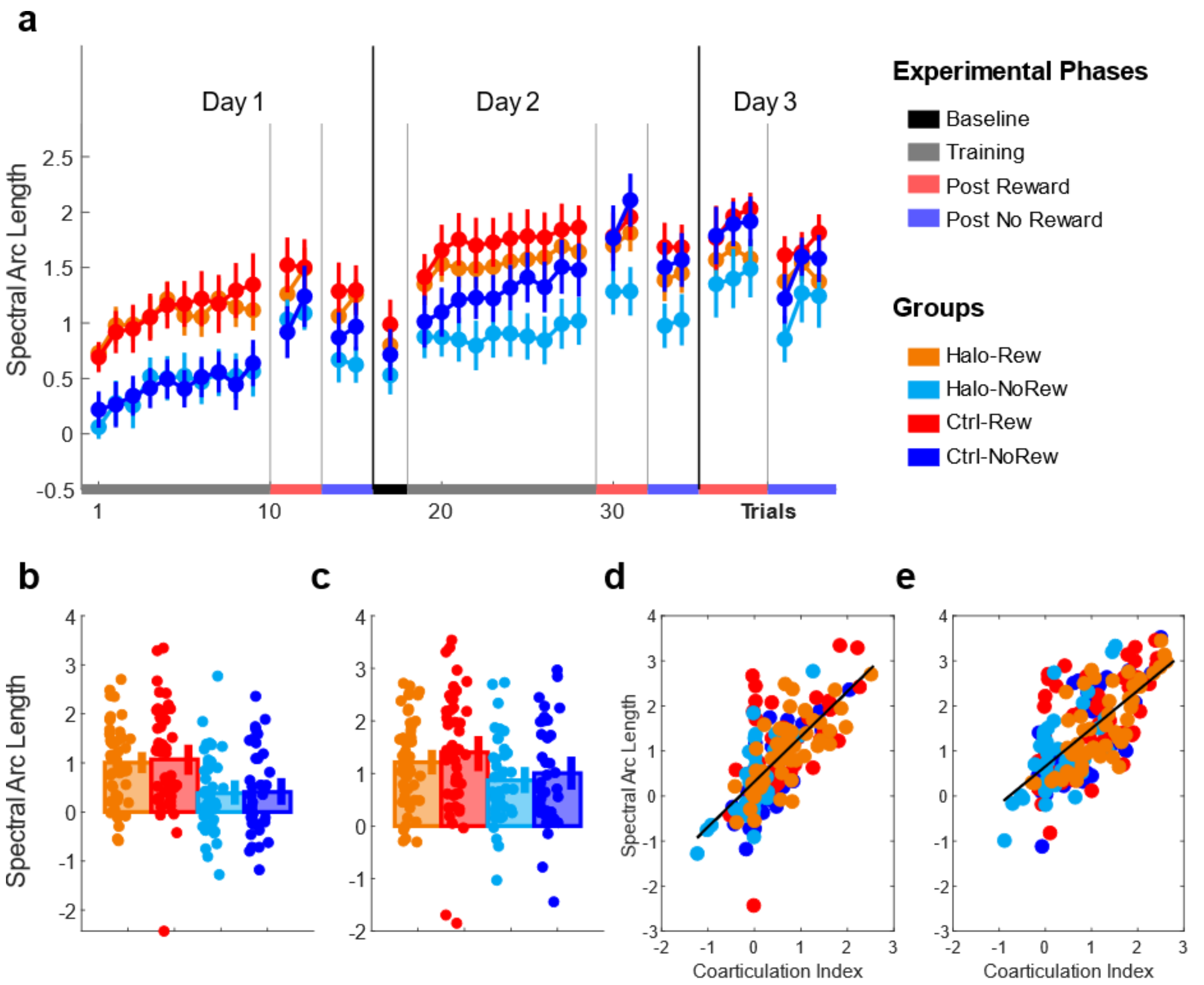


Figure 5 | Increases in movement smoothness were driven by reward but were not modulated by haloperidol. **a)** Changes in smoothness presented in epochs of 20 trials averaged over participants for all groups: Median smoothness levels for each group during **b)** Training1 and **c)** Training2. Scatterplots displaying the relationship between median spectral arc length values and median CI levels during **d)** Training1 and **e)** Training2. Error bars represent SEM. Note negative values are possible due to baseline corrections.

Haloperidol impaired the progressive alignment to the predictions of a minimum jerk model when no reward-based feedback was available.

Assessing whether performance aligned with the predictions of an optimisation model (minimum jerk model), we calculated the mean squared error between the model's prediction and the actual data on a trial-by-trial basis. Across Training1, we observed a significant decrease in mean squared error (trial number; Training1: β = -0.0003, CI = [-0.0002

0.0005], t = 4.96, p < 0.0001, Figure 6a), suggesting that performance aligned progressively more with the predictions of the minimum jerk model.

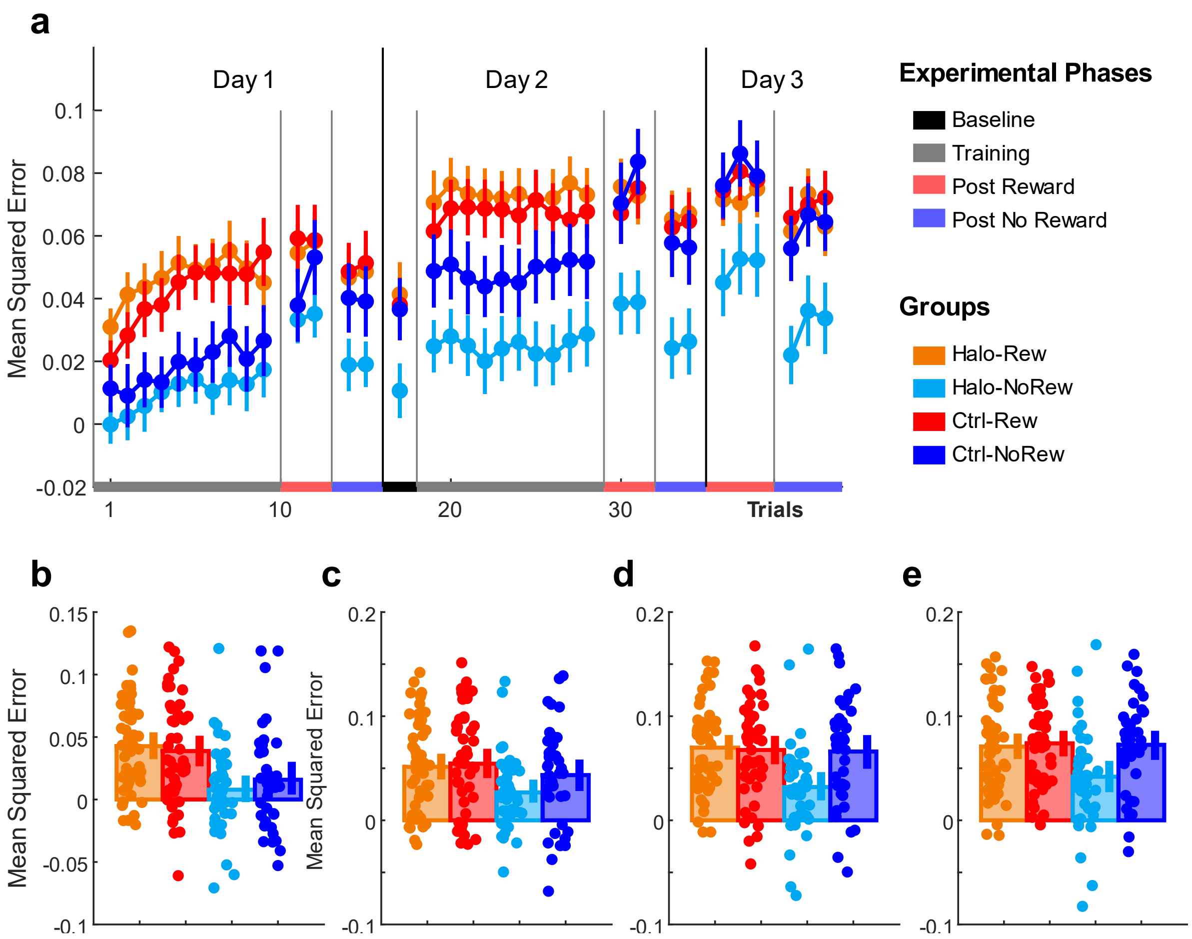


Figure 6 | Haloperidol impaired the progressive alignment to the predictions of a minimum jerk model when no reward-based feedback was available. a) Changes in mean squared error presented in epochs of 20 trials averaged over participants for all groups: Median mean squared error for each group during **b)** Training1 **c)** Training2, **d)** Post1 and **e)** Post2. Error bars represent SEM. Note negative values are possible due to baseline corrections.

Additionally, we found that mean squared error was significantly lower for the reward groups across both trainings (group; Training1: β = -0.05, CI = [-0.098, -0.007], t = 6.67, p < 0.0001; Training2: β = -0.07, CI = [-0.13, 0.02], t = 6.02, p < 0.0001). Importantly, a group x drug interaction was significant for both days (group x drug interaction; Training1: β = 0.02, CI = [-0.01, 0.05], t = 2.67, p = 0.008; Training2: β = 0.03, CI = [-0.007, 0.06], t = 2.57, p = 0.011,

Figure 6b,c). Post-hoc analysis revealed a significant difference in mean squared error for the no reward groups on day2 (Wilcoxon test; Training1: Z = -0.6, p = 0.684, $\eta^2 = 0.04$; Training2: Z = -2.4, p = 0.0324, $\eta^2 = 0.27$). In contrast, we did not find performance differences for the reward groups on either day (Training1: Z = 0.41, p = 0.684, $\eta^2 = 0.04$; Training2: Z = 0.83, p = 0.4093, $\eta^2 = 0.08$).

Despite differences not reaching significance during Training1 for the no reward groups, this trend suggests that haloperidol impaired this optimisation process in the no reward group during initial learning (day1) which then led to a progressively larger deficit across days. Across all post assessments, we found significant group differences (group; Post1: β = -0.04, CI = [-0.1, 0.014], t = 2.41, p = 0.018; Post2: β = -0.15, CI = [-0.003, 0.07], t = 2.34, p = 0.022; Post7: β = -0.06, CI = [-0.117, 0.006], t = 2.22, p = 0.029). These results converge with previous findings suggesting that reward facilitates this optimisation process (i.e., reduces jerk). Additionally, we found significant a group x drug interaction for Post1 and near significant interaction for Post2 (group x drug interaction; Post1: β = 0.04, CI = [-0.003, 0.072], t = -2.09 p = 0.04; Post2: β = 0.03, CI = [-0.007, 0.063], t = 1.9, p = 0.059). This suggests that haloperidol impaired performance in the no reward group and aligns with the results during training.

4.4 Discussion

In summary, reward led to a clear improvement in movement time, peak velocity, coarticulation, movement smoothness and similarity to an optimal minimum-jerk performance. This replicates our work in chapter 3 and reaffirms that reward is associated with improved speed and efficiency during sequential reaching. When under the influence of a dopamine D2-receptor antagonist (haloperidol), it was found that haloperidol affected both movement vigour and coarticulation. While both haloperidol groups showed a global slowing (slower movement time) during the day of drug intake, this effect was only significant for the Halo-Rew group. Interestingly, a significant drug effect was found on day2 suggesting that haloperidol led to a global slowing on the second testing day. However, it is important to note that the MT performance differences are far more pronounced in the no reward groups than in the reward groups. This stratification in MT performance cannot be fully explained by differences in peak velocities, which were performed at a similar level across both no reward groups on day2 (here performance differences are more pronounced in the reward groups). Instead, the marked impairment in coarticulation which was apparent on both testing days in the Halo-NoRew group could explain this pronounced difference in MTs between no reward groups. While the Ctrl-NoRew group gradually increased coarticulation levels across Training2, leading to further reductions in dwell time around the via points, the Halo-NoRew group exhibited sustained impairment in coarticulation. Interestingly, this impairment in coarticulation could not be observed in the Halo-Rew group with similar CI levels across reward groups, suggesting that haloperidol impaired coarticulation only when reward-based feedback was not provided.

Importantly, these results indicate that reward has dissociable effects on movement invigoration and coarticulation. While haloperidol led to a global slowing of movement time, coarticulation was only impaired in the Halo-NoRew group. This suggests that while reward driven by DA modulates both movement invigoration and coarticulation, the neural underpinnings appear to be somewhat dissociable. Support for this dissociation comes from recent work showing that tonic DA appears to modulate motor vigour[167,170,171], while phasic DA has been shown to play a role in motor learning[160–166]. Therefore, haloperidol appears to affect tonic DA irrespective of reward-based feedback leading to increases in MT. These findings agree with a recent study which showed that a selective depletion of D2 receptors in mice impairs the initiation and vigour of self-paced actions[167]. Further evidence comes from studies which assessed response vigour in PD patients OFF medication and showed a reduction in response vigour in both a saccadic task[168] and an effort-based decision-making task[169]. Consequently, DA appears to modulate motor vigour, while a D2 antagonist leads to an increase in MTs irrespective or reward-based feedback.

In contrast to tonic DA, phasic DA has been found to play a role in motor learning[160–166]. Specifically, phasic firing of DA neurons increases if the obtained reward is higher than expected. In contrast, a dip in phasic DA burst can be observed for the opposite[142]. Consequently, phasic DA firing in response to reward appear to explain well how action values are updated thereby enabling learning[128]. In line with these findings, our results suggest that DA plays a role in coarticulation. However, interestingly this effect was only present in the Crtl-NoRew group. The impairment in the Crtl-NoRew group mirrors recent neuropharmacological work that has shown that tiapride, a highly selective D2 antagonist, impaired learning on a complex motor sequence task[166]. Further support comes from a study which assessed chunking in PD patients ON and OFF medication. This study found that patients OFF medication showed severe impairment in chunking even for well-trained motor sequences. In comparison, patients ON medication did not differ significantly from healthy controls[164]. Consequently, our results support recent findings and suggest that DA plays a role in motor sequence learning and particularly in learning to skilfully execute movement sequences through coarticulation. Importantly, the detrimental effects of a D2 antagonist on coarticulation appear to be compensated for by reward-based feedback. There is ample evidence that both the direct and indirect pathway contribute to motor sequence learning[160–162], with some work suggesting that the direct (D1) pathway is more involved when learning from rewards[148]. Consequently, blocking D2 receptor binding could lead to increased D1 receptor binding, which may increase learning from positive reinforcement[148]. However, in the current study, blocking D2 receptor binding through haloperidol did not enhance coarticulation in the presence of reward, indicating that this proposed dichotomy may be overly simplistic. Further research is needed to explain why reward-based feedback preserved coarticulation.

Although sometimes lacking significance, there was an overall trend for the Halo-NoRew group to express impaired smoothness and a reduced similarity to the predictions of a minimum-jerk model throughout both testing days. This suggests that impairing coarticulation leads to impairments in movement quality that propagate and cannot be compensated for during a subsequent training. Further research is needed to understand why participants in the Halo-Rew do not show improvements in coarticulation on day2.

These results provide evidence that both movement invigoration of sequential reaching movements and coarticulation are DA-dependent processes. However, the underlying neural mechanisms appear to be different. While haloperidol led to a global slowing, coarticulation was only impaired in the Halo-NoRew group. Based on prior work, this suggests that tonic DA modulates movement invigoration, while phasic dopamine underpins coarticulation. Importantly, the deleterious effects of haloperidol on coarticulation can be compensated for with reward-based feedback.

Chapter 5

Coarticulation is guided by the principle of smoothness maximisation

5.1 Introduction

Humans have the remarkable ability to produce complex sequential movements such as drinking a cup of coffee or driving a car with reliability, smoothness and speed. Recent behavioural evidence suggests that the production of such efficient upper limb action sequences requires coarticulation[25–27,217]. Ubiquitous in our daily life but often impaired in clinical populations[21–24], coarticulation describes the slow process of blending discrete sub-movements into continuous actions[25,27,28,99]. Imagine, you want to drink a sip of your coffee. Rather than reaching for your cup with your hand still closed, you start opening your hand while reaching for it. Similarly, you will tilt your cup while moving it to your already opened mouth. This simple example of drinking a sip of coffee highlights the importance of coarticulation to daily activities and illustrates its main feature: it allows for discrete sub-movements to be merged into a single continuous action[194,195,205]. Thus, upper limb coarticulation represents a hallmark of skilled sequential performance[25–28,83,194,217].

However, despite showing in chapters 3-4 that reward led to a clear improvement in coarticulation, this could only be observed in three pairs of in-centre-out reaching movements. In contrast, no increases in coarticulation were evident in the other reaching movement transitions. Similarly, no coarticulation could be observed in the pilot studies described in chapter 2. In these experiments, reward only led to reductions in MTs which were transient (i.e., they disappeared once reward was removed). These results mirror previous findings by Sosnik et al. (2004, 2007, 2015)[25–27]. In their work, participants were asked to execute four sequential reaching movements. However, even after nine training sessions, only two pairs of reaching movements were coarticulated, while pronounced stop periods could be observed between the other movement transitions[25–27]. Consequently, coarticulation appears to be transition-specific with some transitions allowing for coarticulation, while others do not. Hence, it remains unclear which sequential actions benefit from coarticulation and what guides this selection process.

Seminal work by Flash and Hogan (1985) introduced a kinematic model which predicts human behaviour based on the criterion of maximum smoothness[99]. It posits that the most optimal way of completing a given task is to choose the smoothest motor command[99]. Therefore, the aim of the motor system is to optimise the cost of jerkiness. Indeed, the so-called minimum jerk model[99,100] has been shown to account for a variety of reaching movements[99,100,111,112]. In line with these results, in both the work presented in this thesis and earlier work by Sosnik er al. (2004, 2007, 2015)[25–27] task performance progressively aligned with the predictions of a minimum-jerk model[28,99]. This suggests that a minimum-jerk model can accurately predict which movement transitions will be coarticulated. Therefore, coarticulation appears to be guided by the criterion of maximum smoothness. Considering that reductions in jerk have been linked to decreases in metabolic effort[209], maximising smoothness appears to be a biologically plausible strategy to guide human behaviour.

Here we tested the hypothesis that coarticulation is guided by the criterion of maximum smoothness using a novel reaching task design. In this experiment, participants were asked to perform a sequence of two reaching movements with varying reaching angles between them ranging from acute to obtuse. Based on the predictions of a minimum-jerk model, in obtuse angles performance is maximally smooth if the sequential reaching movements are fully coarticulated. In contrast, executing two discrete movements (i.e., no coarticulation) maximises smoothness in acute angles. Hence, this experiment, by systematically varying the angle of the movement transition, explores whether the principle of smoothness maximisation guides the coarticulation of reaching movements.

Furthermore, two further research questions were addressed in this study: 1) Does a delay in receiving rewarded training influence coarticulation?; and 2) Do individual differences affect coarticulation?

It has been shown that increases in task space exploration prior to a subsequent reward-based training, led to steeper learning curves[84,91,219,220]. In contrast, such an exposure could also slow down subsequent learning due to having to unlearn an original behaviour[61]. Consequently, task exposure prior to reward-based training could be either beneficial or detrimental to performance. To address this question, two training phases were included in this experiment and participants either received reward during the first or second phase.

In both the work presented in this thesis and earlier work by Sosnik er al. (2004, 2007, 2015)[25–27] , a small but significant number of participants fail to coarticulate sequential movements even with reward and/or after days of training[25–27,217]. To date, these individual differences in coarticulation have not been accounted for. Based on recent work that demonstrated working memory capacity influences both motor skill learning and adaptation[221–226], participants were asked to complete a working memory (WM) task prior to the start of the experiment. Additionally, to investigate whether these individual differences in coarticulation could be mapped onto naturally occurring differences in movement smoothness[86] participants performed a doodling task in-between the working memory task and the main experiment.

5.2 Methods

Participants: 59 participants (18 males; age range 18-24) were recruited to participate in this experiment, which had been approved by the local research ethics committee of the University of Birmingham. All participants were novices to the task paradigm and were free of motor, visual and cognitive impairment. Most participants were self-reportedly right-handed (N = 8 left-handed participants) and gave written informed consent prior to the start of the experiment. For their participation, participants were remunerated with either course credits or money (£7.5/hour) and were able to earn additional money during the task depending on their performance. Before the start of the experiment, participants were pseudo-randomly allocated to one of the available groups.

Experimental Apparatus: All experiments were performed using a Polhemus 3SPACE Fastrak tracking device (Colchester, Vermont U.S.A; with a sampling rate of 110Hz). Participants were seated in front of the experimental apparatus which included a table, a horizontally placed mirror 25cm above the table and a screen (Figure 1a). The low-latency monitor placed 25cm above the mirror had a refresh rate of 60Hz and displayed the workspace and participants' hand position (represented by a green cursor – diameter 1cm). On the table, participants were asked to perform 2-D reaching movements. Looking into the mirror, they were able to see the representation of their hand position reflected from the screen above. This setup effectively blocked their hand from sight. The experiment was run using MATLAB (The Mathworks, Natwick, MA), with Psychophysics Toolbox 3.

Task design: In this experiment, we aimed to assess whether coarticulation is guided by smoothness maximation. To this end, a task was designed in which participants were asked to hit a series of targets displayed on the screen consisting of a via point and an end target (both 1cm diameter). The angle between these two reaching movements (both 10cm) ranged from acute to obtuse (10°,27.5°,45°,62.5°,80°; Figure 1b). Consequently, participants were asked to execute two equally long continuous reaching movements with 5 different reaching angles between them. Based on the predictions of a minimum-jerk model (Figure 1c,d), in obtuse angles performance is maximally smooth if the sequential reaching movements are fully coarticulated. In contrast, executing two discrete movements (i.e., no coarticulation) maximises smoothness in acute angles (Figure 1e,f). Hence, this experiment, by systematically varying the angle of the movement transition, explores whether the principle of smoothness maximisation guides coarticulation of reaching movements. To start each trial, participants had to pass their cursor though the preparation box (2x2cm) on the left side of the workspace, which triggered the appearance of the start box (2x2cm) in the centre of the screen. After moving the cursor into the start box, participants had to wait for 1.5s for the targets to appear. This ensured that participants were stationary before reaching for the first target. Target appearance served as the go-signal. Upon reaching the last target, all targets disappeared, and participants had to wait for 1.5s before being allowed to exit the start box to reach for the preparation box to initiate a new trial. Participants had to repeat a trial if they missed a target. Similarly, exiting the start box too early either at the beginning or at the end of each trial resulted in a missed trial.

Reward Structure and Feedback: The same reward structure as in chapters 3-4 was used in this experiment. Participants experienced either reward or no reward trials depending on the current experimental phase: (1) Reward trials were cued using a visual stimulus prior to the start of the trial (Figure 1c). Once participants moved into the preparation box, the start box appeared in yellow (visual stimulus) rather than in black (Figure 1f). Participants were informed that faster MTs will earn them more money, with a maximum amount of 5p available in each trial. While participants moved from the start box to the preparation box to initiate a new trial, the amount earned in the previous trial was displayed on the top of the screen (i.e. 'You have earned 2p out of 5p'). Similarly to our previous experiment, we used a closed-loop design to calculate the amount of reward earned in each trial (see Chapter 2 for more information). When participants started a new experimental block, performance in the first trial was compared to the last 20 trials of the previously completed block. (2) No reward trials were not cued, and no reward was available for participants. However, participants were instructed to 'move as fast and accurately as possible' (Figure 1f).

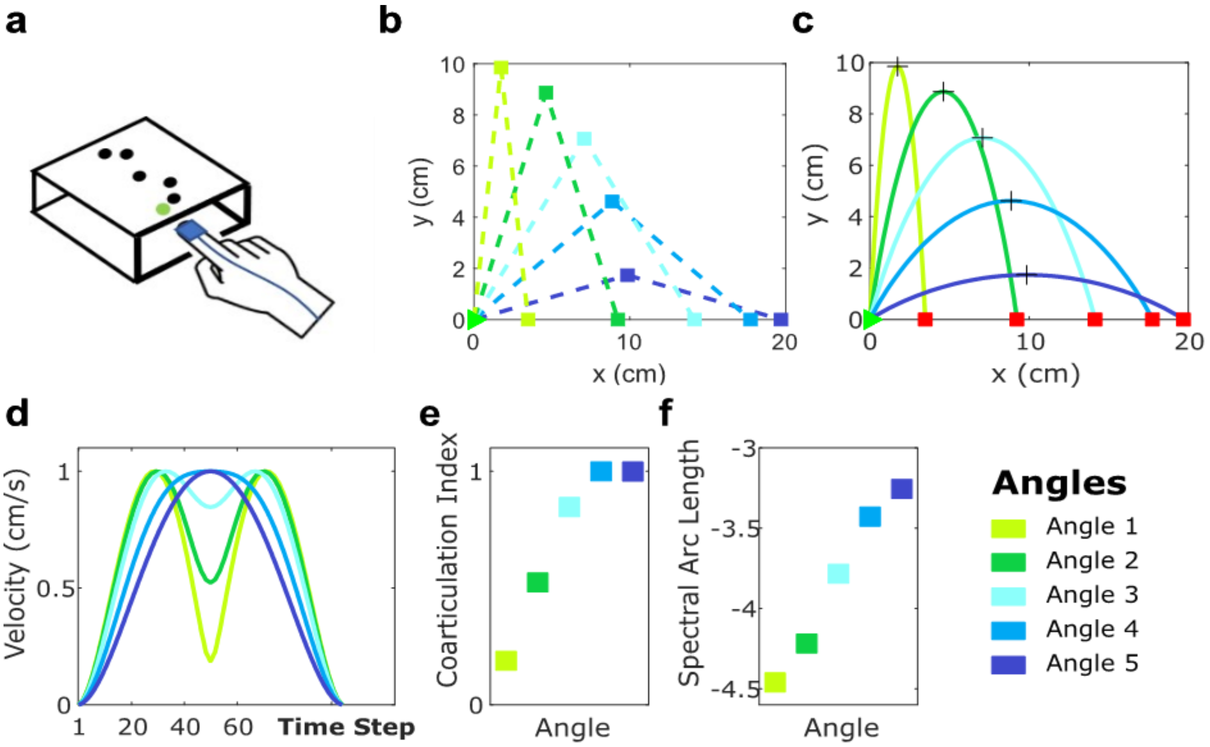


Figure 1 | Experimental setup and task design. a) Participants wore a motion-tracking device on their index finger (Polhemus). **b)** Location of the via and end target for each angle configuration. **c)** Predicted minimum jerk trajectories for each angle configuration. Dotted lines represent discrete reaches to each target. **d)** Velocity profile for each angle configuration derived from the positional data (x,y) of the minimum jerk model. **e)** The level of coarticulation and **f)** Smoothness (spectral arc length) present in each minimum jerk model prediction, with values closer to 0 indicating smoother movements.

Experimental Procedure

Participants were pseudo-randomly allocated to either the Reward1 (N = 29) or Reward2 (N = 30) group. Both groups completed two training blocks (Training1 and Training2) and were then asked to complete two Post assessments afterwards (post-R where reward was available and post-NR without reward). Participants in the Reward1 group received reward during Training1 and no reward during Training2. In contrast, participants in the Reward2 group received no reward first and were then able to earn reward during Training2. Differences in performance during the Post assessment could provide insight whether the timing of reward affects future task performance. To address whether individual differences influence coarticulation participants were asked to complete both a working memory (cognitive) and a free doodling task (motor) at the beginning of the experiment (Figure 2a).

Working memory task (WM): Participants were asked to complete a WM task prior to the start of the main experiment[225]. Across 60 trials, participants were asked to memorise an array of consonants during the encoding period (1s) and remember them during a subsequent maintenance period (3s). During recall (4s), participants were presented with a single letter and were asked to decide whether this letter was part of the array or not. Participants were able to respond by pressing one of three keys: “1” key indicated that the letter presented in the recall period was part of the array (‘match’). “2” key indicated the opposite (‘nonmatch’); and “3” indicated that the participant was unsure as to the correct answer (Figure 2b). This WM task included 5 levels of difficulty with 12 trials presented for each. Half of these (6 trials) were trials in which ‘match’ was the correct answer, whereas in the remaining 6 trials ‘nonmatch’ was the correct answer. The order of presentation was pseudorandomized across participants, who had 10 practice trials prior to the start of the WM task to familiarise themselves with the task and instructions. Difficulty in this task was determined by the length of the array to be remembered, ranging from 5 to 9[225].

Doodling Task: Across 5 trials participants were asked to doodle freely for 6s. Participants were seated in front of the experimental apparatus used for the main experiment wearing the Polhemus tracking device on their index finger. To start a trial, participants moved the cursor into the start box positioned in the centre of the workspace. After a wait period (1.5s), a white rectangle (40x20cm) appeared around the start box, which served as the go signal. Participants were instructed to freely doodle within the confines of the square until it disappeared, which also marked the end of the current trial. To initiate a new trial, participants had to move back into the start box, which appeared once the trial was over (Figure 2c).

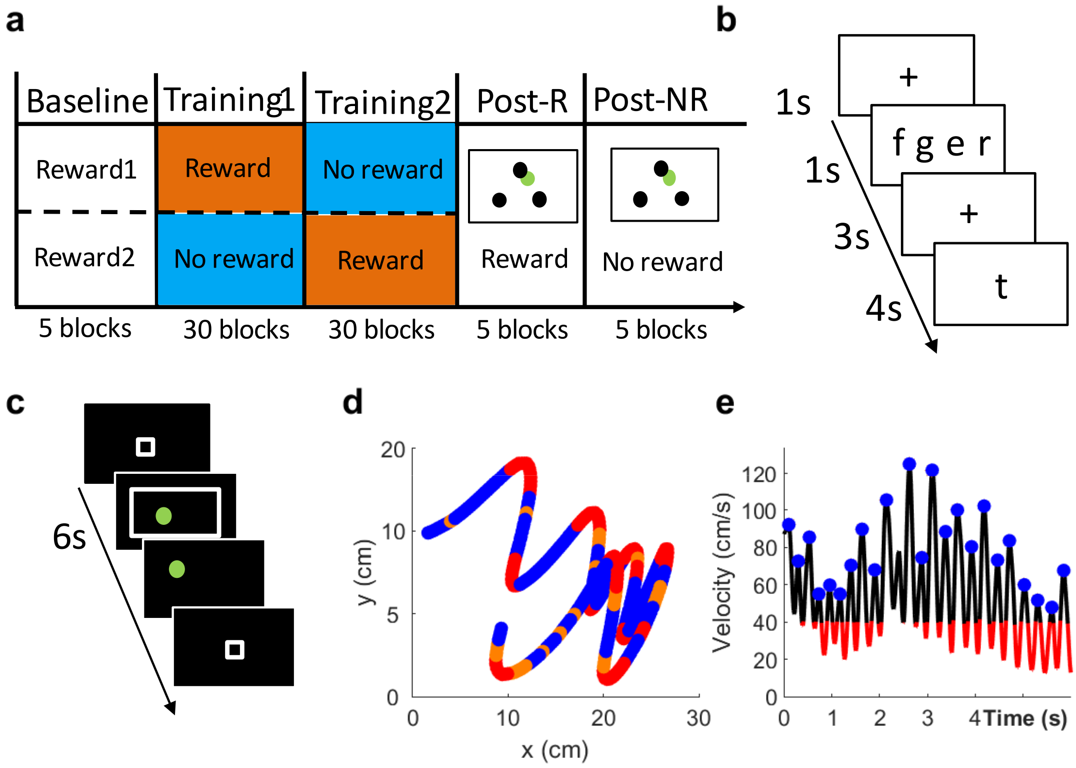


Figure 2 | Experimental setup and kinematic markers. a) Study design. Participants were randomly allocated to the Reward1 and Reward2 group and completed two trainings (Training1 and Training2) and two Post assessments (post-R and post-NR). **b)** Illustration of the working memory task. **c)** Illustration of the doodling task. **d)** Trial data from the doodling task, which was coloured according to the calculations of the curvature analysis. Blue segments represent segments, in which angle changes were below the threshold (straight lines), whereas angle changes above the threshold were coloured in red and orange (curves) **e)** Velocity profile derived from doodling trial data. Coloured in red are velocity values that fall below the peak velocity threshold (20%) and were used to calculate the mean arrest period ratio. Blue dots represent peaks in the velocity profile to determine the number of peaks for the kinematic analysis of the doodling data.

Baseline: Participants in both groups completed 5 blocks of baseline trials prior to the start of the main experiment. Each block contained 5 trials, one for each angle configuration. The order within each block was randomised. No reward was available during Baseline, instead both groups were instructed to 'move as fast and accurately as possible'

Training1: Participants in the Reward1 group were informed that during this part they would be able to earn money depending on how fast they complete each trial (25 blocks of 5 trials). In contrast, participants in the Reward2 group engaged in 25 blocks of no reward trials and were again instructed to move as fast and as accurately as possible.

Training2: Participants in the Reward1 group were not able to earn any reward during this part. Instead, they were again instructed to move as fast and as accurately as possible. In contrast, participants in the Reward2 group were informed that during this part they would be able to earn money depending on how fast they complete each trial (25 blocks of 5 trials).

Post assessments: Participants from both groups were asked to complete two post assessments (5 block of 5 trials); one with reward trials (post-R) and one with no reward trials (post-NR). The order was counter-balanced across participants.

Data Analysis

Analysis code will be available on the Open Science Framework website, alongside the experimental datasets at https://osf.io/62wcz/. The analyses were performed in Matlab (Mathworks, Natick, MA). A similar array of parameters to assess performance as in Chapters 3-4 were used here (for more information on the individual parameters see Chapter 2):

Movement Time (MT): MT was measured as the time between exiting the start box and reaching the end target. This excludes reaction time, which describes the time between target appearance and when the participants has left the start box.

Peak and Minimum Velocity: Through the derivative of positional data (*x, y*), the velocity profiles were obtained and both the peak and minimum velocities of each reaching movement were subsequently determined. These were then averaged for each trial.

Coarticulation Index (CI): To measure coarticulation, we compared the mean peak velocities of the two sequential reaches with the minimum velocity around the via point. The smaller the difference between these values, the greater coarticulation had occurred between the two movements[193]. We normalised the obtained difference, ranging from 0 to 1, with a CI value of 1 indicating a fully coarticulated movement.

Spectral Arc Length: To assess movement smoothness, we used spectral arc length as our smoothness metric. Spectral arc length has been shown to be less sensitive to differences in MT and more sensitive to changes in smoothness [85,87]. The spectral arc length is derived from the arc length of the power spectrum of a Fourier transformation of the velocity profile. We used an open-source Matlab toolbox to calculate this value for each trajectory [198].

Coarticulation Magnitude: To understand how much participants coarticulated sequential reaching movements across all five angle configurations, we determined their coarticulation magnitude. To this end, we aggregated participants' CI scores across four time windows, which included the last 5 blocks of Training1 and Training2 as well as all 5 blocks of post-R and post-NR. We then z-scored the CI values for each angle separately and found the mean for each participant. This analysis stratifies participant's performance into participants that comparatively coarticulated more across all angle configurations and participants that did not. Positive values for this coarticulation magnitude metric indicate greater expression of coarticulation, whereas negative values suggest that a given participant coarticulated less in magnitude across all angles and timepoints of interest.

Working Memory Performance: WM performance was defined as the average percentage of correct responses across the three highest levels of difficulty for each task[225].

Kinematic Analysis of Doodling Task: To analyse the doodling task, we compiled a battery of kinematic markers including mean speed, spectral arc length and three additional smoothness metrics. This approach was inspired by recent work assessing kinematic changes during stroke recovery[23,24].

Curvature ratio: Movement smoothness has been traditionally quantified using the velocity profile (temporal smoothness). However, here we sought to use the trajectory of the movement to assess smoothness. Spatial smoothness, represented by the curvature of the trajectory, has been shown to capture both stroke patient' recovery and movement quality in healthy subjects[227]. To measure curvature, we first used spline interpolation (MATLAB function interp1) to evenly space data points (unit distance) and then measured the angle between consecutive data points. Subsequently, we subtracted consecutive angles values and defined curvature as differences between angle values exceeding 5°. Finally, we measured the ratio of curved movements to straight movements, with higher curvature ratio values indicating higher spatial smoothness (Figure 2d).

Mean Speed: We calculated the mean speed of each doodling trial from the velocity profile, which was derived from the recorded positional data (x, y).

Spectral Arc Length: We used the same approach to calculate the trial-based spectral arc length as described above.

Mean arrest period ratio: We determined the mean arrest period ratio in each doodling trial by measuring the proportion of time speed was below a given percentage of the peak velocity. The threshold used in this analysis was 20%, which has been shown to be informative when studying movement smoothness in healthy participants (Figure 2e)[23,228].

Number of velocity peaks: Using the same threshold as in the mean arrest period ratio analysis we obtained the number of peaks in the velocity profile that exceeded the given threshold.

This smoothness metric has been used in both healthy subjects and stroke patients to quantify movement smoothness (Figure 2e)[229,230].

Statistical Analysis

Throughout we used mixed model ANOVAs to assess statistical significance. A mixed model ANOVA with group (Reward1 and Reward2) and angle (angle1 – angle5) as factors was used to measure differences in performance during baseline. We then carried out three separate analyses to analyse the performance across Training1 and Training2. All mixed model ANOVAs included group (Reward1 and Reward2) and angle (angle1 – angle5) as factors. In addition, timepoint was added as a factor. To assess changes across Training1, we compared the first 5 blocks during early Training1 with the last 5 blocks of Training1. Similarly, to measure changes in performance from Training1 to Training2 (from here on called Transition) we compared the last 5 blocks of Training1 with the first 5 blocks of Training2. Finally, to assess performance across Training2 we compared the first 5 blocks of Training2 with the last 5 blocks of Training2. Hence, we used three separate mixed model ANOVAs with different timepoints (Training1, Transition and Training2) to answer our main research question. To address our second research question (reward timing on task performance), we carried out a mixed model ANOVA with group (Reward1 and Reward2), angle (angle1 – angle5) and timepoint (Post (i.e., post-R and post-NR)) as factors. We used one-sample Kolmogorov-Smirnov tests to test our data for normality and found that all measures were non-parametric. Median values were therefore used as input in all mixed model ANOVAs. Wilcoxon tests were employed when a significant interaction and/or main effects were reported and corrections for multiple comparisons were performed using Bonferroni correction. Linear partial correlations (fitlm function in Matlab) were used to measure the degree of association between the chosen variables, while accounting for the factor group and angle.

5.3 Results

Reward invigorated sequential reaching movements irrespective of reaching angle.

The results converge with results from Chapters 3-4 showing that reward invigorated sequential reaching performance (Figure 3a,b). No differences between groups or angles were observed during baseline (mixed-effect ANOVA; group: F = 0.92, p = 0.3416; angle: F = 2.09, p = 0.0834). However, significant interactions between timepoint and group for both Training1 (timepoint x group; F = 7.65, p = 0.0076) and Transition (timepoint x group; F = 36.34, p < 0.0001) were found. Specifically, significant difference between groups during late Training1 (Wilcoxon test; Z = -6.2, p < 0.0001), but not during early Training1 (Z = -1.88, p = 0.302, Figure 3c) were observed. Similarly, a significant decrease in MTs for Reward2 once they received reward during Training2 was observed (Z = 6.22, p < 0.00001), whereas Reward1 increased MTs (Z = -2.99, p = 0.0136, Figure 3d). These results suggest that, reward invigorates sequential reaching. However, we only found significant group differences during Training1 (group; F = 4.82, p = 0.0322), while no significant differences were observed across the other timepoints of interest (group; Transition: F = 1.7, p = 0.1972; Training2: F = 1.3, p = 0.2596,). Similarly, the interaction between timepoint and group was not significant during Training2 (timepoint x group; F = 2.02, p = 0.1612, Figure 3e), which indicates that despite Reward1 increasing their MTs during early Training2, overall group performances did not differ across the rest of the experiment.

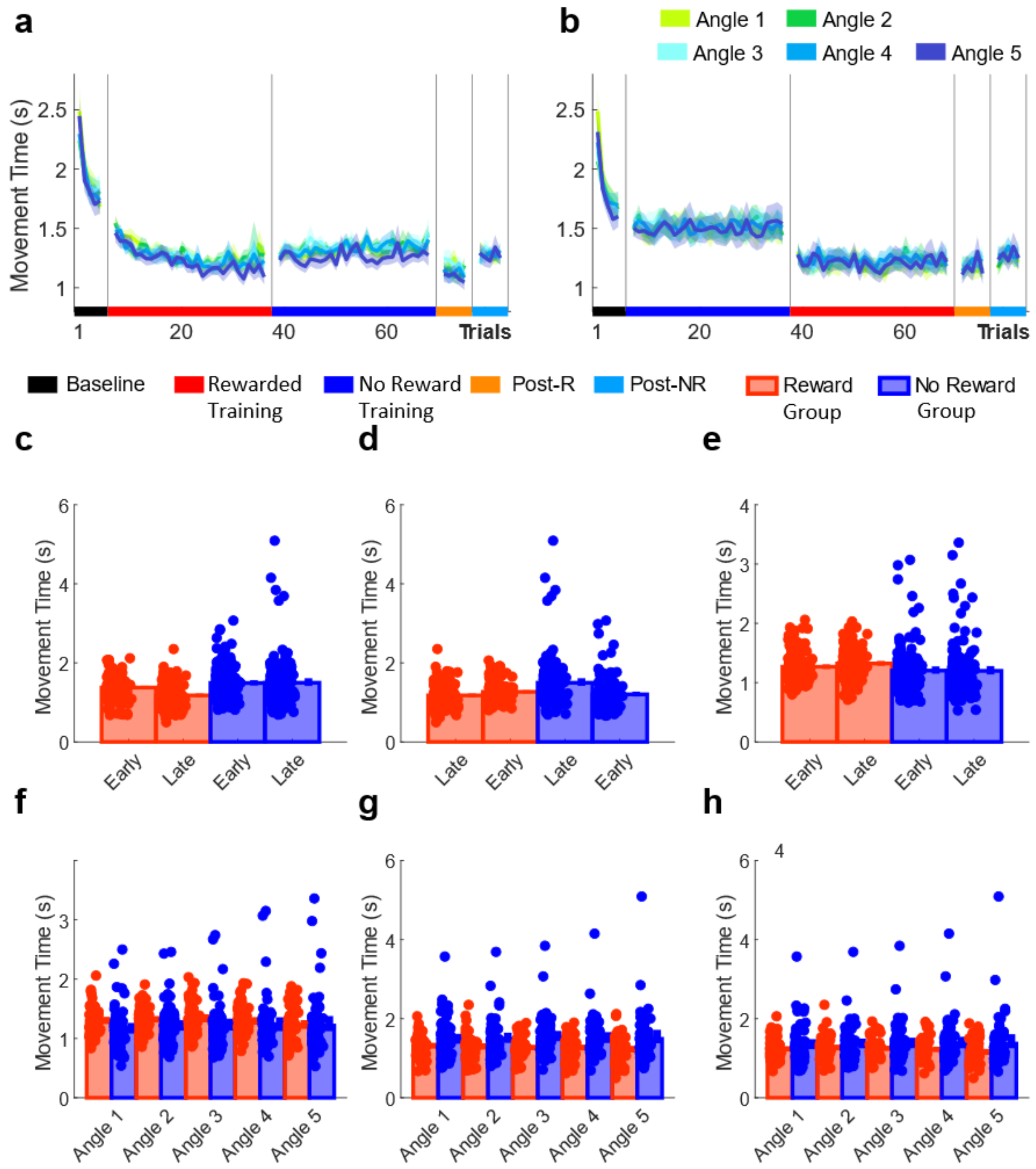


Figure 3 | Reward invigorates sequential reaching movements irrespective of reaching angle (◦). Trial-by-trial changes in MT averaged over participants for each angle and group: **a)** Reward1 and **b)** Reward2. Median change in MT between groups for timepoint: **c)** Training1, **d)** Transition and **e)** Training2. Median change in MT for each angle between groups for timepoint: **f)** Training1, **g)** Transition and **h)** Training2. Shaded regions/error bars represent SEM.

This suggests that Reward1 maintained performance gains across Training2, in which they were not able to receive any reward. A significant interaction between group and angle during

Training2 (group x angle; F = 3.71, p = 0.0061) as well as a significant main effect for angle (angle; F = 2.9, p = 0.023) indicates that this retention of reward-based improvements is angle-specific. Specifically, we found that group performance on all angles apart from angle5 were significantly different during Training2 (Wilcoxon test; angle1: Z = 3.37, p = 0.0037; angle2: Z = 2.93, p = 0.0169; angle3: Z = 3.14, p = 0.0086; angle4: Z = 3.02, p = 0.0128: angle5: Z = 1.56, p = 0.6021, Figure 3f). These findings indicate that the retention of reward-based performance gains was greatest in angle5.

Additionally, we found significant interactions between group and angle during Training1 and Transition (group x angle; F = 3.26, p = 0.0126, Figure 3g; F = 3.8, p = 0.0052, Figure 3h, respectively). During Training1, only the performance on angle5 was significantly different between groups suggesting that here reward invigoration was the greatest (Wilcoxon test; angle1: Z = -1.53, p = 0.6283; angle2: Z = -2.24, p = 0.1248; angle3: Z = -2.14, p = 0.2101; angle4: Z = -2.41, p = 0.0788: angle5: Z = -2.79, p = 0.0267, Figure 3g). In contrast, changes in performance during Transition were not significant (Wilcoxon test; angle1: Z = -0.79, p = 1; angle2: Z = -0.67, p = 1; angle3: Z = -0.9, p = 1; angle4: Z = -1.19, p = 1: angle5: Z = -1.97, p = 0.2556, Figure 3h). Overall, these results show that reward invigorates sequential reaching movements, and that the magnitude and retention of these reward-based performance gains was greatest in the most obtuse angle (angle5).

Reward invigorated peak velocities but changes in peak velocities did not fully explain reward-based retention of performance gains.

The results show that reward invigorates peak velocities (Figure 4a,b), however fails to fully explain the retention of MT performance gains during Training2. Despite no differences in peak velocity during baseline (mixed-effect ANOVA; group: F = 0.17, p = 0.6755; angle: F = 0.28, p = 0.4356) a significant interaction between all factors during Training1 (timepoint x group x angle; F = 5.29, p < 0.0001) and Transition (timepoint x group x angle; F = 4.66, p =

0.0012) was found. Specifically, we found a significant difference between groups during late Training1 (Wilcoxon test; Z = 5.23, p < 0.00001), but not during early Training1 (Z = 2.24, p = 0.1251, Figure 4c). Similarly, a significant increase in peak velocities for Reward2 once they received reward was observed during Training2 (Z = -5.75, p < 0.0001), whereas Reward1 exhibited a decrease in peak velocities (Z = 2.76, p = 0.029, Figure 4d). This supports existing findings that reward-based invigoration of MT can be driven by increases in peak velocities. However, in contrast to the MT results, no significant differences between groups could be found during Training1 and Transition (main effect for group; F = 2.95, p = 0.0915, F = 0.12, p = 0.7287, respectively).

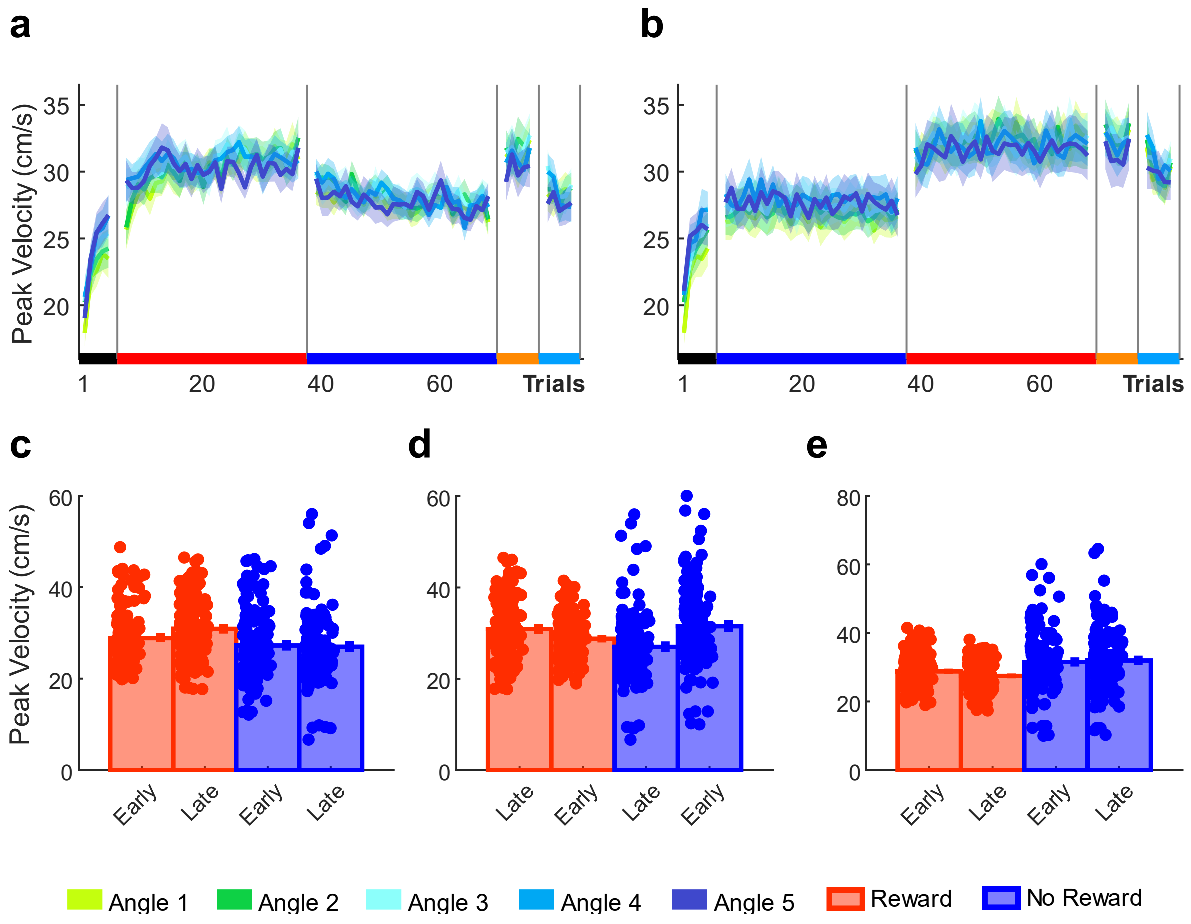


Figure 4 | Reward invigorates peak velocities but changes in peak velocities do not fully explain reward-based retention of performance gains. Trial-by-trial changes in peak velocities averaged over participants for each angle and group: **a)** Reward1 and **b)** Reward2. Median change in peak velocities between groups for timepoint: **c)** Training1, **d)** Transition and **e)** Training2. Shaded regions/error bars represent SEM.

In contrast, a significant group difference was found during Training2 (for group; F = 4.81, p = 0.0325), suggesting that here, unlike in the MT analysis, groups exhibited a significant difference in peak velocities. Additionally, no interaction between group and angle was observed (group x angle; F = 0.37, p = 0.8298, Figure 4e) suggesting that this change in peak velocities across Training2 was not angle-specific. Overall, these results suggest that the invigoration of peak-velocities is reward-sensitive but follows the previously observed 'on-off effect' of reward on movement invigoration. Consequently, the observed retention of Reward1's performance gains during Training2 cannot be fully explained by peak velocities.

Coarticulation is guided by the criterion of smoothness maximisation while increased coarticulation led to a retention of reward-based decreases in MT.

The results show that more obtuse angles are being coarticulated faster and to a greater extent than more acute angles (Figure 1a,b). While no differences between groups or angles were observed during baseline (mixed-effect ANOVA, group: F = 0.54, p = 0.4661; angle: F = 0.36, p = 0.8374), a significant interaction for both Training1 (timepoint x group; F = 5.93, p = 0.0181) and Transition (timepoint x group F = 11.77, p = 0.0011) was found. Post hoc test revealed a significant difference between groups during late Training1 (Wilcoxon test; Z = 4.17, p < 0.0001, Figure 1c), but not during early Training1 (Z = 1.23, p = 0.98, Figure 1d). Similarly, a significant increase in CI levels for Reward2 during Training2 was observed (Z = -3.78, p < 0.00001), suggesting that reward facilitates coarticulation. Importantly, a lack in interaction during Training2 (timepoint x group; F = 0.23, p = 0.6313, Figure 1e), may suggest that this is a stable change in performance. Additionally, no significant group differences were found during Training2 (group; F = 0.01, p = 0.9521) indicating that a delay in reward-based training did not influence coarticulation.

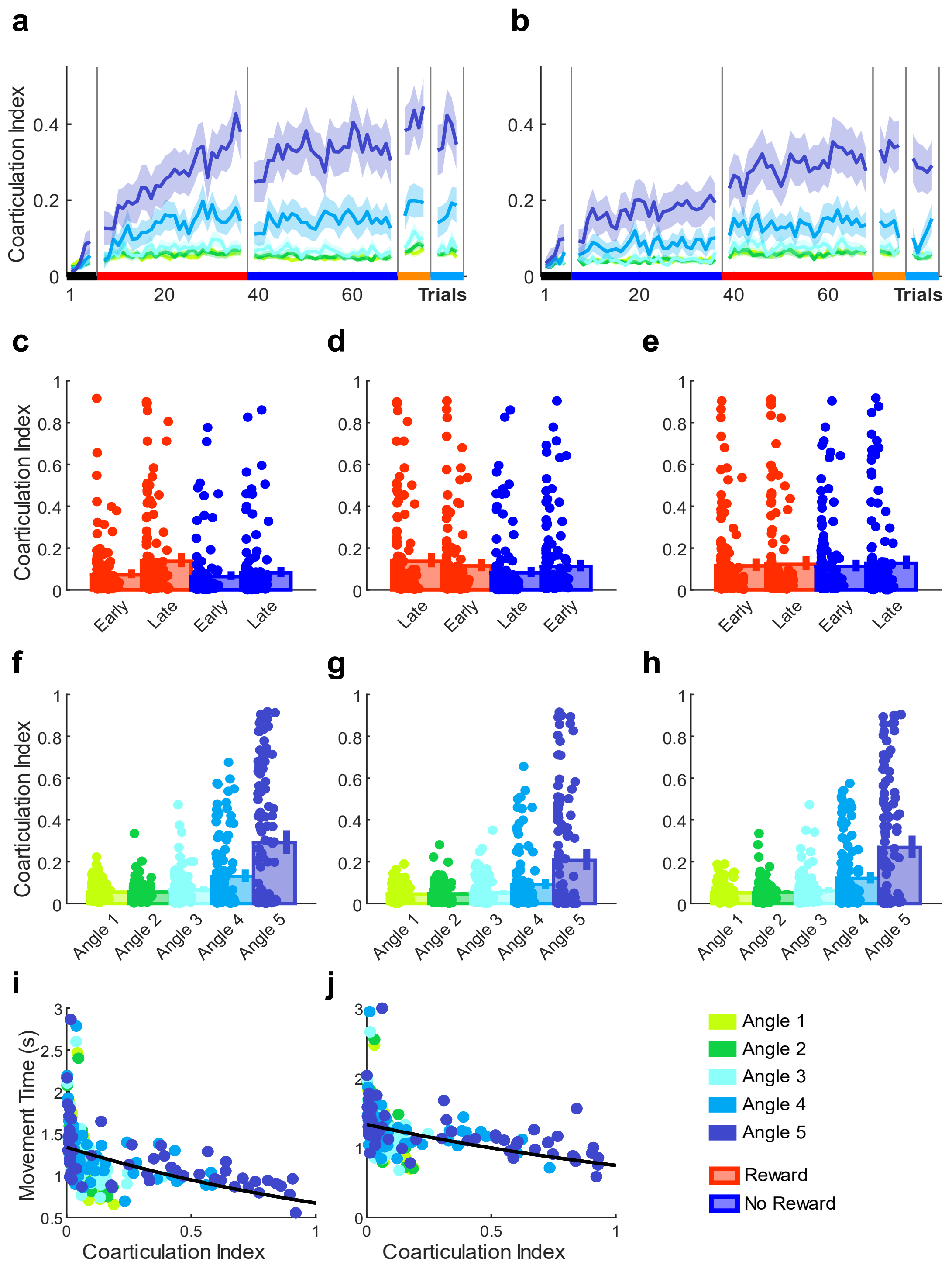


Figure 5 | Coarticulation is angle-specific with obtuse reaching angles allowing for more coarticulation; a process which is facilitated by reward. Trial-by-trial changes in CI levels averaged over participants for each angle and group: **a)** Reward1 and **b)** Reward2. Median change in CI levels between groups for timepoint: **c)** Training1, **d)** Transition and **e)** Training2. Median values of CI levels for each angle across groups for timepoint: **f)** Training1, **g)** Transition and **h)** Training2. Scatterplots displaying the relationship between MT and CI levels during **i)** late Training2 and **j)** post NR with a linear line fitted across groups. Shaded regions/error bars represent SEM.

Notably, significant differences between angles across all timepoints of interest were observed (angle; Training1: F = 28.99, $p < 0.0001$; Transition: F = 36.87, $p < 0.0001$; Training2: F = 37.39, $p < 0.0001$; Figure 5f-h). Specifically, CI levels for angle5 were significantly higher than all other angles for each timepoint (see Supplement Table 1). Similarly, CI levels for angle4 were significantly higher than angle1-angle3, whereas no differences could be observed between the most acute angles (angle1-angle3; see Supplement Table 1). These results highlight that more obtuse angles can be coarticulated faster and to a greater extent, which represents a change in performance that becomes reward-independent over time. It also indicates that coarticulation on different angles can be predicted by the minimum-jerk model. Based on predictions of a minimum-jerk model, CI level calculations revealed that angle5 can be fully coarticulated, whereas no (or very little) coarticulation is present in angle1. The presented results fit the model predictions and suggest that smoothness maximisation guides coarticulation. Considering that the retention of reward-based performance gains was greatest in magnitude in angle5, we sought to understand whether increases in CI levels are related to this observed retention. To this end, MT values with CI levels were correlated during Training2 and post-NR, and a significant correlation was found between them (partial correlation controlling for both group and angle; Training2: $\rho = -0.44$, $p < 0.0001$, Figure 5i, post-NR: $\rho = -0.41$, $p < 0.0001$, Figure 5j). Although not causal, this indicates that faster MTs during both Training2 and post-NR were associated with higher levels of coarticulation, which was most prominent in angle5.

Reward invigorated minimum velocities around the via point which represented a stable change in motor output irrespective of reward availability.

Coarticulation leads to discrete actions being progressively merged into a single movement, which comes with increases in minimum velocities around the via point. Such an increase in minimum velocity reduces the dwell time in the target and ultimately decreases overall MTs. Hence, here minimum velocities were assessed to investigate whether the retention of performance gains can be linked to stable increases of minimum velocities. Despite no differences in peak velocity during baseline (mixed-effect ANOVA; group: $F = 0.66$, $p = 0.4206$; angle: $F = 0.23$, $p = 0.9178$), we found significant increases in minimum velocities when participants received reward (Figure 6a,b). A significant interaction between timepoint and group for both Training1 (timepoint x group; $F = 9.63$, $p = 0.003$, Figure 6c) and Transition (timepoint x group $F = 17.20$, $p < 0.0001$, Figure 6d) was found. Specifically, a significant difference between groups during late Training1 (Wilcoxon test; $Z = 4.70$, $p < 0.00001$), but not during early Training1 ($Z = 1.69$, $p = 0.46$, Figure 6c) was found. Similarly, a significant increase in minimum velocities for Reward2 were observed once they received reward during Training2 ($Z = 4.70$, $p < 0.00001$), however no changes were found for Reward1 ($Z = 1.62$, $p = 0.52$, Figure 6d). These results suggest that reward facilitates coarticulation and in turn increases minimum velocities around the via point. Importantly, similarly to the coarticulation analysis no significant interaction between timepoint and group for Training2 (timepoint x group; $F = 1.71$, $p = 0.1951$, Figure 6e) as well as no significant group differences (group; $F = 0.38$, $p = 0.5399$) were found. These results align with the coarticulation results suggesting that the observed changes in performance are robust and become reward-independent with training. Additionally, we observed significant differences between angles across all timepoints of interest (angle; Training1: $F = 24.48$, $p < 0.0001$; Transition: $F = 32.75$, $p < 0.0001$; Training2: $F = 31.14$, $p < 0.0001$; Figure 6f-h). Specifically, the CI levels for angle5 were significantly higher than all other angles for each timepoint (see Supplement Table 2). Similarly, CI levels for angle4 were significantly higher than angle1-angle3 for all timepoints apart from Training1, whereas no differences could be observed between the most acute angles (angle1-angle3; see Supplement Table 2). To ascertain whether increases in minimum velocities were related to faster MTs during periods of no reward, both values for Training2 and post-NR were correlated. A significant correlation between minimum velocities and MTs for both timepoints (partial correlation controlling for both group and angle; Training2: $\rho = -$

0.47, p < 0.0001, Figure 6i, post-NR: ρ = -0.45, p < 0.0001, Figure 6j) was found. Hence, decreases in dwell time via increases in minimum velocities around the via point were associated with the retention of performance gains during periods of no reward.

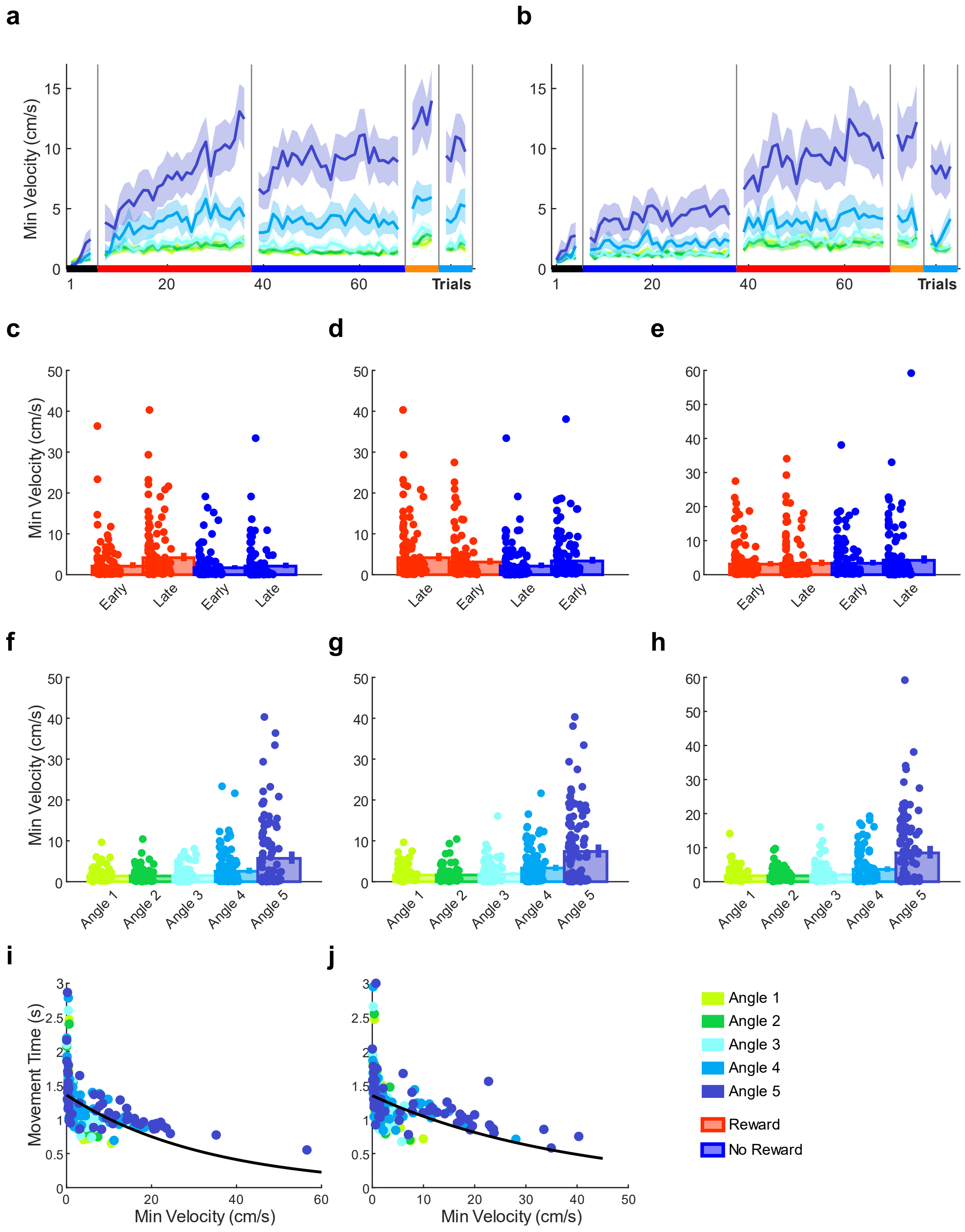


Figure 6 | Reward invigorates minimum velocities around the via point which represents a stable change in motor output irrespective of reward availability. Trial-by-trial changes in minimum velocities averaged over participants for each angle and group: **a)** Reward1 and **b)** Reward2. Median change in minimum velocities between groups for timepoint: **c)** Training1, **d)** Transition and **e)** Training2. Median values of minimum velocities for each angle across groups for timepoint: **f)** Training1, **g)** Transition and **h)** Training2. Scatterplots displaying the relationship between minimum velocities and CI levels during **i)** late Training2 and **j)** post NR with a linear line fitted across groups. Shaded regions/error bars represent SEM.

To summarise, these results demonstrate that improvements in MT were driven by two processes which are both reward-sensitive: 1) an increase in peak velocities and 2) an increase in coarticulation which represented a reduction in dwell time via an increase in minimum velocities around the via point. Importantly, increases in coarticulation represented a stable change in behaviour that was associated with the retention of reward-based performance gains during periods of no reward availability. In line with the predictions of the minimum-jerk model, this was particularly apparent in the most obtuse angle (angle5), which was coarticulated faster and to a greater extent, and hence indicates that coarticulation is guided by the criterion of smoothness maximisation. This increase in coarticulation allowed for faster execution during both periods with and without reward (as seen with angle 5). In contrast, the invigoration of peak velocities followed the previously observed ‘on-off’ principle and here performance gains remained transient in nature (as seen with angles 1-3).

Coarticulation is guided by smoothness maximation.

To further explore whether coarticulation is guided by smoothness maximation, actual and model smoothness data was compared. It was hypothesised that coarticulation should only occur if it leads to increases in movement smoothness. Considering that coarticulation was most pronounced in angle5, it should follow that during early Training1 movement smoothness was least optimal with regards to the predictions of a minimum-jerk model. Over training and with coarticulation, improvements in movement smoothness should be observed that approach optimal values. Conversely, movement smoothness should be closer to optimal in angle1 over the whole experiment considering that here no increases in coarticulation were observed. Furthermore, over the course of training smoothness across angles should increase depending on how much the reaching movements are coarticulated. To test this hypothesis, using spectral arc length as the smoothness metric the predicted movement smoothness of each angle was calculated (based on the predictions of a minimum-jerk model, Figure 1f). This value was then subtracted by the spectral arc length for each trial performed by each participant . Hence, optimal smoothness represented the difference between the actual and predicted smoothness data, with optimal smoothness values close to zero. Indeed, the results indicate that performance was significantly different across angles during Training1 (mixed-effect ANOVA; angle: $F = 122.68$, $p < 0.0001$, Figure 7a,b). Post hoc analysis revealed that optimal smoothness values were significantly smaller in angle1 when compared to any other angle, while values in angle5 were significantly greater (Figure 7c, see Supplement Table 3). This suggests that movement smoothness was near optimal in angle1, whereas in angle5 smoothness was sub-optimal. Furthermore, a significant timepoint x angle interaction (timepoint x angle interaction; $F = 4.44$, $p = 0.0018$, Figure 7d) suggests that over the course of training improvements in optimal smoothness were angle-specific. Indeed, only performance on angle5 showed significant improvements, while improvements in angle4 approached significance (Wilcoxon test; angle1: $Z = -1.58$, $p = 0.5617$; angle2: $Z = -1.46$, $p = 1$; angle3: $Z = -1.85$ $p = 0.32$; angle4: $Z = -2.34$, $p = 0.0654$: angle5: $Z = -2.79$, $p = 0.0367$, Figure 7d). Results for Training2 mirror these results. Both a significant main effect for angle (angle: $F = 89.15$, $p < 0.0001$, Figure 7e) as well as a significant interaction was found (timepoint x angle interaction; $F = 4.22$, $p = 0.0031$, Figure 7f). However, no Post hoc comparison reached significance. These results suggest that performance on angle1 was closer to optimal throughout the experiment, while optimal smoothness was lowest in angle5.

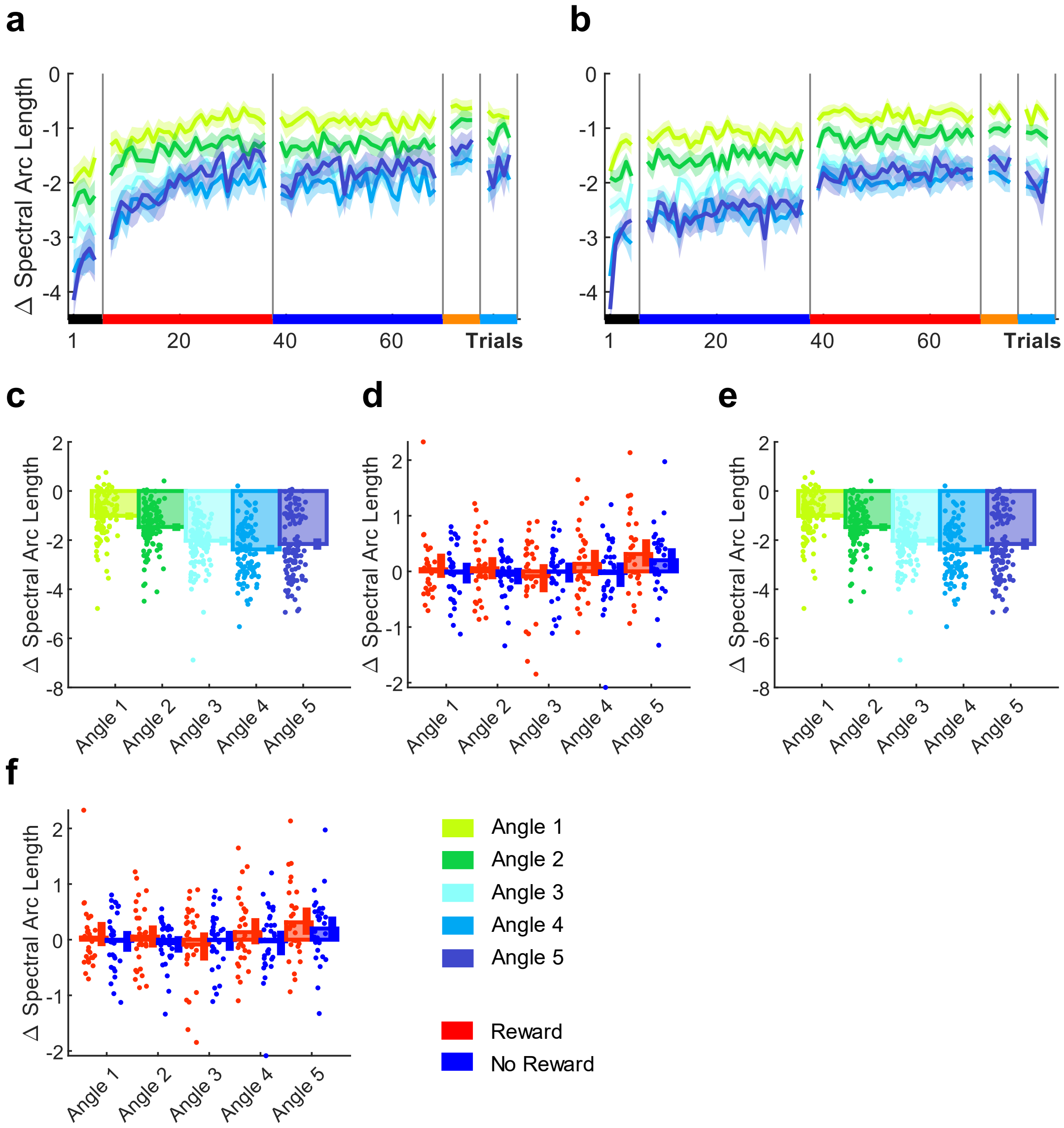


Figure 7| Coarticulation is guided by smoothness maximation. Trial-by-trial changes in optimal smoothness determined by taking the difference between actual and predicted movement smoothness (based on minimum-jerk model) calculated using spectral arc length for each angle and group: **a)** Reward1 and **b)** Reward2. Median change in optimal smoothness between angles for timepoint: **c)** Training1 and **e)** Training2. Difference in median optimal smoothness values (early vs late) for each angle across groups for timepoint: **d)** Training1 and **f)** Training2. Shaded regions/error bars represent SEM.

Furthermore, improvements in optimal smoothness were greatest in angle5, which matches the observed increases in coarticulation. Consequently, coarticulation appears to be only expressed if movement smoothness is sub-optimal.

A delay in rewarded training does not affect future task proficiency.

To assess whether a delay in receiving a rewarded training affects task performance, participants engaged in two counter-balanced Post assessments after Training1 and Training2. During post-R, participants received reward-based feedback, while during post-NR no reward was available, and participants were instructed to complete each trial as fast and accurately as possible.

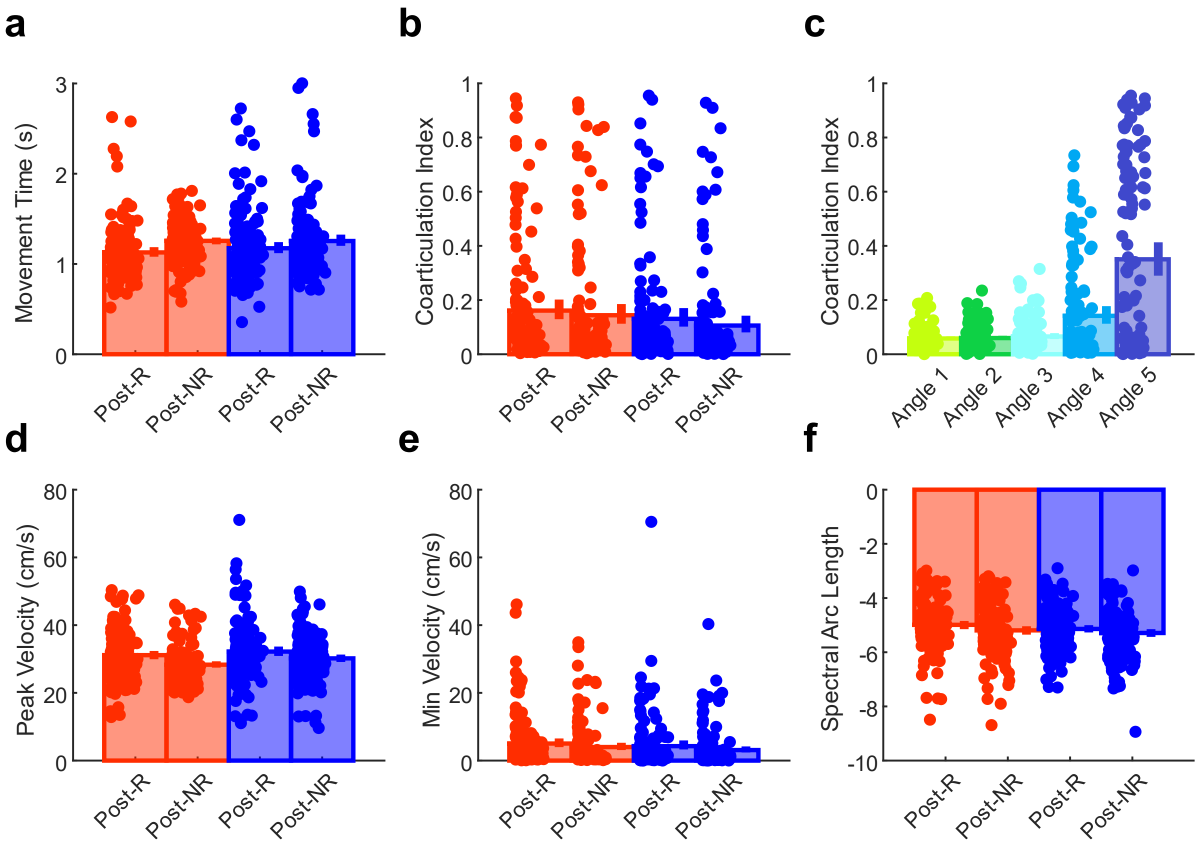


Figure 8 | A delay in rewarded training does not affect future task proficiency. a) Median change in MT values between groups during Post. **b)** Median change in CI levels between groups during Post. **c)** Median CI values for each angle across groups during Post. Median change in **d)** peak velocities, **e)** minimum velocities. Error bars represent SEM.

Assessing differences in performance between groups while under the same reward regime (post-R and post-NR) can provide insight into whether the timing of a reward training affects future task performance. It was hypothesised that the Reward2 group, which received a rewarded training after the non-rewarded training, may show lower performance levels compared to Reward1 especially during post-NR. No significant group differences for MT across Post assessments (mixed-model ANOVA; group; F = 0.78, p = 0.7812) nor a significant interaction between timepoint (post-R vs post-NR) and group (Post; F = 0.76, p = 0.386, Figure 8a) were found, suggesting that both groups scored similar MTs across Post assessments. Similarly, no significant interaction between group and angle (Post; F = 2.03, p = 0.0911) was found. However, we observed a significant main effect for timepoint (Post; F = 14.19, p < 0.0001), which indicates that both groups decreased their MTs once reward was removed during post-NR. These results suggest that the timing of the rewarded training did not affect MT performance.

Assessing changes in CI levels across Post assessments, we yet again could not observe any significant group differences (mixed-model ANOVA; group; F = 1.31, p = 0.2569) nor a significant interaction between timepoint and group (Post; F = 0.21, p = 0.6532, Figure 8b). Similarly, no significant interaction between group and angle (Post; F = 0.92, p = 0.45) was found. However, a significant main effect for angle (Post; F = 49.49, p < 0.0001, Figure 8c) was observed, which indicates that CI levels were different depending on the angle in both groups. In line with the results from our Training2 analysis, we found higher CI levels for angle5 than for any other angle. Similarly, CI levels for angle4 were significantly higher than angle1-angle3, whereas no differences could be observed between the most acute angles (angle1-angle3; see Supplement Table 4).

Converging with these results, no significant group differences across Post assessments for were found for either peak velocities (mixed-model ANOVA; group; F = 0.69, p = 0.4066, Figure 8d) or minimum velocities (mixed-model ANOVA; group; F = 0.62, p = 0.4347, Figure 8e). In summary, these results highlight that Reward1 and Reward2 reached similar levels of task proficiency despite different reward regimes. Thus, the timing of reward does not affect future task performance in this task.

Individual differences within the motor and cognitive domain do not predict ability to coarticulate.

To assess whether individual differences predict participants' ability to coarticulate participants were asked to complete both a working memory (WM) and a doodling task. Specifically, we aimed to investigate whether individual differences within the cognitive (WM) and/or motor (doodling) domain predict coarticulation ability in our main experiment. We determined coarticulation magnitude (CI_Mag) for each participant by comparing CI levels across all timepoints of interest for each angle individually using z-scores. Subsequently, mean performance for each participant was calculated (see Methods for further information on CI_Mag).

Performance on a working memory (WM) task does not predict coarticulation ability.

Correlating participants WM performance with their CI_Mag values a positive relationship was found, suggesting that participants with higher WM values tend to be able to coarticulate more. However, this relationship was not significant (partial correlation controlling for group: $\rho = 0.24$, $p = 0.114$, Figure 9a).

Performance on a doodling task did not predict coarticulation ability.

Using a doodling task, we sought to investigate whether differences in coarticulation magnitude are related to a natural disposition to produce smooth movements. To this end, participants were asked to complete 5 trials in which they had to doodle for 5s. The obtained data was then analysed using a battery of kinematic markers to determine whether individual differences within the motor domain influence coarticulation magnitude.

Using spectral arc length as the smoothness metric, we determined the participants' mean movement smoothness during the doodling task and correlated them to the CI_Mag values.

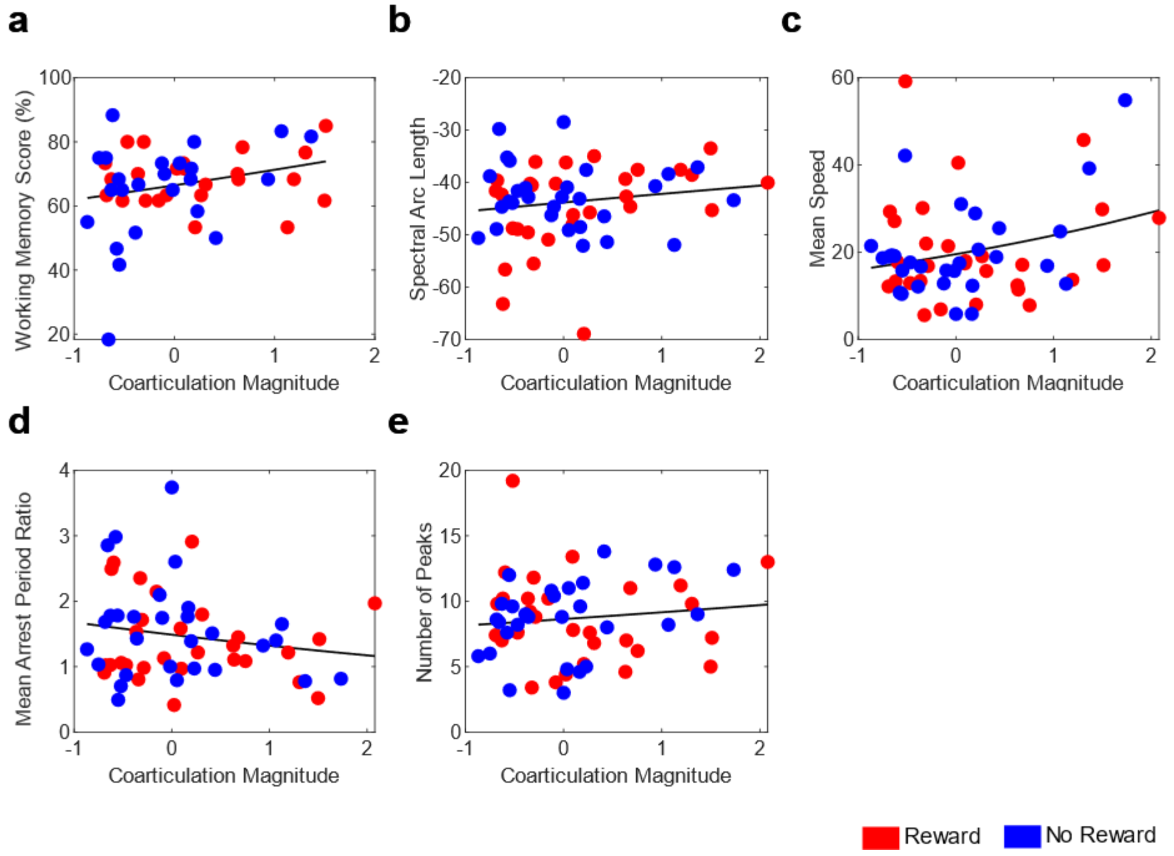


Figure 9 | Individual differences within the motor and cognitive domain do not predict ability to coarticulate. Scatterplots displaying the relationship between coarticulation magnitude (CI_Mag) and **a)** working memory scores (%), **b)** mean smoothness values, **c)** mean speed, **d)** mean arrest period ratio, **e)** number of peaks.

A positive relationship between movement smoothness during free scribbling and coarticulation magnitude during the main task was observed. However, this relationship was not significant (partial correlation controlling for group: ρ = 0.18, p = 0.177, Figure 9b). Similarly, correlating participants' mean speed during doodling with CI_Mag, a positive relationship between mean speed while doodling and coarticulation ability was found. Yet again, this relationship was not significant (partial correlation controlling for group: ρ = 0.23, p = 0.089, Figure 9c). Following participants' mean arrest period ratio was determined, which is defined as the proportion of time that a participant's movement speed is below a given percentage of the peak velocity. We correlated participants' mean arrest period ration during the doodling task with their CI_Mag values and found a negative relationship, which was not significant (partial correlation controlling for group: ρ = -0.17, p = 0.202, Figure 9d). Subsequently, the total number of peaks in a speed profile was determined, with fewer peaks suggesting that participants execute smoother movements due to fewer periods of acceleration and deceleration. The mean number of peaks was used as the peak metric and was correlated with participants' CI_Mag values. A non-significant positive relationship between the peak metric and CI_Mag values was found (partial correlation controlling for group: ρ = 0.11, p = 0.403, Figure 9e). Lastly, we assessed whether coarticulation magnitude during the main experiment is related to the ratio of curved to straight movements during the doodling task. To this end, the ratio of curved to non-curved (straight) movements was determined and used as the curvature metric. We found that the mean ratio of movement curvature during the doodling task is not related to coarticulation magnitude (partial correlation controlling for group: ρ = -0.08, p = 0.576, Figure 9f). Overall, these results indicate that movement smoothness during a free doodling task was not associated with coarticulation magnitude.

5.4 Discussion

The results highlight that reward invigorated performance leading to equal improvements in MTs across angles. Reductions in MTs were driven by two processes which were both reward-sensitive: 1) an increase in peak velocities and 2) an increase in coarticulation which was represented by a reduction in dwell time via an increase in minimum velocities around the via point. Importantly, increases in coarticulation were stable and were associated with the retention of reward-based performance gains during periods without reward. However, reward only facilitated coarticulation on the obtuse angles while no increases in coarticulation were observed in the acute angles. Therefore, coarticulation is angle-specific and improved retention of reward-based reductions in MTs was only observed in the most obtuse angle (angle5). This suggests that decreases in MTs on the acute angles were mainly driven by increases in peak velocities, while both increases in peak and minimum velocities led to the observed improvement in MTs on the obtuse angles. Recent work has shown that increases in arm stiffness enable faster MTs through increases in peak velocities whilst maintaining similar levels of end-point accuracy[75]. Albeit an attractively simple mechanism to account for the reward-based improvements in the speed-accuracy trade-off, this strategy comes with a marked increase in metabolic cost[209]. Therefore, removing reward might make this strategy unviable. This could explain the lack of retention in the acute angles and highlights that increases in minimum velocities through coarticulation, which was shown to be a stable change in behaviour, enables retention in the obtuse angles.

The results show that coarticulation is angle-specific suggesting that coarticulation is constrained by the transition angle between reaching movements. Based on seminal work by Flash and Hogan[99] demonstrating that the minimum jerk model can account for a variety of reaching movements[99,100,111,112], it was hypothesised that coarticulation is guided by the criterion of maximum smoothness. Considering that reductions in jerk have been linked to decreases in metabolic effort[209], maximising smoothness appears to be a biologically plausible strategy to guide human behaviour. Indeed, the results demonstrate that, compared to obtuse angles, movement smoothness within acute angles was significantly closer to being optimally smooth (as determined by the predictions of a minimum-jerk model). Therefore, changes in behaviour through coarticulation will only have led to small improvements in smoothness. In contrast, performance on the most obtuse angle was sub-optimal. Here movement smoothness was improved through coarticulation becoming more optimal over the course of training. Hence, these results suggest that the principle of maximising smoothness guides coarticulation and more specifically highlights that coarticulation is constrained by the transition angle between reaching movements. These results explain the fact that no improvements in coarticulation were seen in the pilot studies. The transition angles between the reaching movements were all acute and therefore, executing discrete reaching movements is more optimally smooth than coarticulating the movement sequences. However, this also results in reward-based invigoration effects remaining transient.

However, it is important to stress that motor goals other than movement smoothness can guide behaviour and motor skill learning. In this task, both the primary goal (decreases in MT) and the secondary goal (movement smoothness) can be achieved through increases in coarticulation in obtuse and the lack thereof in acute angles, with coarticulation enabling retention. Yet, other primary goals (such as path length or increases in MT) may warrant distinct behavioural solutions that are optimal given the underlying reward structure. Despite this, the results highlight that kinematic models such as the minimum-jerk model are very accurate at predicting skilful execution of reaching movement sequences irrespective of the task goals. Additionally, the finding that coarticulation is guided by the principle of smoothness maximation suggests that the motor system incorporates smoothness as an inherent goal to optimise behaviour[99,100]. The minimum-jerk model was challenged because it was deemed biologically unfeasible for the motor system to calculate and keep track of the 3rd derivative of positional data which has to be integrated over the whole movement[101]. However, reductions in jerk have been linked to decreases in metabolic effort[209], which might suggest that rather than keeping track of jerk per se, the motor system simply stores a value for muscular effort. It has consistently been shown that effort influences both decision-making and execution during reward-based motor tasks[231,232]. Therefore, effort based on smoothness maximisation could guide coarticulation. However, this will have to be proved experimentally in future.

The results additionally showed that a delay in receiving rewarded training does not affect future task performance. Specifically, no group differences during Post were found, which suggests that both groups performed similarly across all measured variables. Moreover, we observed a steep improvement in MTs, peak and minimum velocities for the Reward2 group once they received reward during Training2. In contrast, improvements were more gradual for the Reward1 group during Training1 as seen in a lack of group differences during early Training1. Therefore, the results align with recent work showing that participants exhibited steeper learning curves in reward-based training, if it followed non-rewarded training[84,91,219,220]. Similarly, within latent learning it has been shown that rats find the rewarded location within a maze faster (rewarded phase) if they were pretrained first without reward[130,233]. However, despite Reward2 showing a steep improvement, no additional performance gains were observed for Reward2 during Training2.

Furthermore, the results from the WM task and doodling task suggest that individual differences do not predict coarticulation in this task. Despite, recent research showing that WM ability modulates motor skill learning and adaptation[221–226], the here presented analysis did not reveal a significant relationship between WM scores and coarticulation. In this task, only two movements had to be planned in each trial, which ultimately might not require extensive cognitive resources. Similarly, no behavioural measures to assess doodling performance were associated with coarticulation. This could either suggest that the measures were not sensitive enough to detect individual differences in smoothness during doodling, or that coarticulation is not linked to a natural tendency to produce smooth movements.

In summary, these results suggest that coarticulation is guided by the principle of smoothness maximation which leads to increases in coarticulation in obtuse angles. These increases in coarticulation result in increases in minimum velocities which appear to drive reward-based retention of reductions in MTs.

5.5 Supplement

Supplementary Table 1

x	Angle1	Angle2	Angle3	Angle4	Angle5
Angle1					
Angle2	Z = 1.18 , p = 1				
Angle3	Z = 1.5, p = 1	Z = 0.34 , p = 1			
Angle4	Z = 1.1 , p = 0.03	Z = 0.26, p = 0.02	Z = -0.08, p = 0.07		
Angle5	Z = -0.07, p < 0.001	Z = -0.84, p < 0.001	Z = -1.21, p < 0.001	Z = -1.41 , p < 0.001	

x	Angle1	Angle2	Angle3	Angle4	Angle5
Angle1					
Angle2	Z = 0.54 , p = 1				
Angle3	Z = 0.64, p = 0.95	Z = 0.01 , p = 1			
Angle4	Z = -0.67, p < 0.001	Z = -1.02, p < 0.001	Z = -0.88, p = 0.004		
Angle5	Z = -2.22, p < 0.001	Z = -2.25, p < 0.001	Z = -2.62, p < 0.001	Z = -2.32 , p < 0.001	

x	Angle1	Angle2	Angle3	Angle4	Angle5
Angle1					
Angle2	Z = 0.33 , p = 1				
Angle3	Z = 0.74, p = 0.97	Z = 0.23 , p = 0.97			
Angle4	Z = -0.51, p < 0.001	Z = -1.01, p < 0.001	Z = -0.95, p = 0.001		
Angle5	Z = -1.99, p < 0.001	Z = -2.41, p < 0.001	Z = -244, p < 0.001	Z = -1.97 , p < 0.001	

Supplementary Table 1: Coarticulation is most pronounced in obtuse angles. Using the MATLAB functions anova2 and multcompare to assess differences in median CI levels between angles across for timepoint: **top)** Training1, **middle)** Transition and **bottom)** Training2. Highlighted in yellow are significant comparisons.

Supplementary Table 2

x	Angle1	Angle2	Angle3	Angle4	Angle5
Angle1					
Angle2	Z = 1.11 , p = 1				
Angle3	Z = 1.12, p = 1	Z = 0.06 , p = 0.92			
Angle4	Z = 0.59 , p = 0.06	Z = -0.22 = 0.062	Z = -0.38, p = 0.11		
Angle5	Z = -0.41, p < 0.001	Z = -1.11, p < 0.001	Z = -1.32, p < 0.001	Z = -1.31, p < 0.001	

x	Angle1	Angle2	Angle3	Angle4	Angle5
Angle1					
Angle2	Z = 0.54 , p = 1				
Angle3	Z = 0.45, p = 0.98	Z = -0.12 , p = 0.99			
Angle4	Z = -1.07, p = 0.03	Z = -1.4, p = 0.036	Z = -1.25, p = 0.019		
Angle5	Z = -2.24, p < 0.001	Z = -2.77, p < 0.001	Z = -2.69, p < 0.001	Z = -2.23 , p < 0.001	

x	Angle1	Angle2	Angle3	Angle4	Angle5
Angle1					
Angle2	Z = 0.17 , p = 1				
Angle3	Z = 0.41, p = 0.99	Z = 0.17 , p = 0.99			
Angle4	Z = -0.82, p = 0.004	Z = -1.11, p = 0.004	Z = -1.21, p = 0.024		
Angle5	Z = -2.27, p < 0.001	Z = -2.58, p < 0.001	Z = -2.66, p < 0.001	Z = -2.02 , p < 0.001	

Supplementary Table 2: Minimum velocities are fastest in obtuse angles. Using the MATLAB functions anova2 and multcompare to assess differences in median minimum velocity between angles across for timepoint: **top)** Training1, **middle)** Transition and **bottom)** Training2. Highlighted in yellow are significant comparisons.

Supplementary Table 3

x	Angle1	Angle2	Angle3	Angle4	Angle5
Angle1					
Angle2	Z = 4.4 , p < 0.0001				
Angle3	Z = 8.16 , p < 0.0001	Z = 4.75 , p < 0.0001			
Angle4	Z =9.32, p < 0.0001	Z = 6.79 , p < 0.0001	Z = 2.88 , p < 0.0001		
Angle5	Z =6.79, p < 0.0001	Z = 4.13 , p < 0.0001	Z = 1.16 , p = 0.2672	Z = -1.32, p = 0.0081	

x	Angle1	Angle2	Angle3	Angle4	Angle5
Angle1					
Angle2	Z = 4.86 , p < 0.0001				
Angle3	Z = 8.87 , p < 0.0001	Z = 6.37 , p < 0.0001			
Angle4	Z = 9.53, p < 0.0001	Z = 3.39 , p < 0.0001	Z = 2.88 , p < 0.0001		
Angle5	Z = 4.84, p < 0.0001	Z = 2.68 , p < 0.0001	Z = 0.34 , p = 0.2816	Z = -1.26, p = 0.1525	

Supplementary Table 3: Optimal smoothness using spectral arc length is greatest in the most acute angle. Using the MATLAB functions anova2 and multcompare to assess differences in median smoothness vlaues between angles across for timepoint: **top)** Training1, and **bottom)** Training2. Highlighted in yellow are significant comparisons.

Supplementary Table 4

x	Angle1	Angle2	Angle3	Angle4	Angle5
Angle1					
Angle2	Z = 0.049 , p = 1				
Angle3	Z = 0.644, p = 0.99	Z = 0.41 , p = 1			
Angle4	Z = -0.621, p < 0.001	Z = -0.27, p < 0.001	Z = -0.79, p = 0.004		
Angle5	Z = -3.76, p < 0.001	Z = -3.73, p < 0.001	Z = -4.01, p < 0.001	Z = -3.39 , p < 0.001	

Supplementary Table 4: Coarticulation is most pronounced in obtuse angles. Using the MATLAB functions anova2 and multcompare to assess differences in median CI levels between angles across for timepoint Post. Highlighted in yellow are significant comparisons.

Chapter 6

General discussion

6.1 Summary of thesis

In chapter 2, a novel sequential reaching task (CSRT) was introduced, which was based on previous pilot studies and on important work by Sosnik et al. (2004, 2007, 2015)[25–27]. In comparison to the task design used in the pilot studies, the CSRT allowed for coarticulation of movements sequences, which was confirmed a priori by the predictions of a minimum-jerk model[28,99]. Additionally, results from the pilot studies provided evidence that reward can invigorate sequential reaching performance as seen by greater decreases in movement time (MT).

In chapter 3, results were presented that demonstrate that reward invigorates performance on the CSRT through reductions in MT via both increases in peak velocities and decreases in dwell time around the via points. Specifically, these reductions in dwell time represented an increase in coarticulation. Importantly, the results suggest that reward facilitates coarticulation, which also leads to improvements in movement smoothness and MTs that persist even over an additional testing day without reward.

In chapter 4, a neuropharmacological manipulation was used to investigate whether coarticulation, which was shown to be reward-sensitive, is a dopamine-dependent process. The findings highlight that coarticulation is severely impaired by haloperidol (D2 antagonist) if no reward-based feedback is available. In contrast, haloperidol had no deleterious effect on coarticulation if reward was provided. This finding complements existing work showing that a D2 antagonist impairs chunking in a discrete sequence learning task[164] and expands the literature by showing that reward-based feedback appears to preserve coarticulation. Additionally, haloperidol led to a global slowing, as seen by increases in MTs and peak velocities across groups, which aligns with previous results showing that tonic dopamine modulates motor vigour.

In chapter 5, two hypotheses were tested: 1)coarticulation is both transition and is guided by the principle of maximising smoothness, and 2) coarticulation is related to individual differences in working memory (WM) capacity and baseline movement. The results confirmed that coarticulation is transition specific and is guided by smoothness maximisation which led to increases in coarticulation only in the more obtuse angles. These findings explain the lack of coarticulation in the pilot studies in which angles between reaching movements were acute. Furthermore, both individual differences in WM and baseline movements smoothness did not explain differences in coarticulation across participants. Consequently, it remains an open question why some participants, even with reward-based feedback, do not coarticulate.

6.2 Impact of this work on the emerging field of upper limb coarticulation

Coarticulation has been studied in speech production for over 100 years[180–182] and more recently been described in a range of upper limb movement sequences such as typing[186–188], sign language[184] and piano playing[185]. Furthermore, coarticulation has also been observed during various 2D and 3D upper limb actions[111,189–193]. Consequently, upper limb coarticulation appears to be a central mechanism to human movement that is ubiquitous in our daily life and represents a hallmark of skilled sequential performance through which behaviour becomes more temporally and spatially efficient[25–27,99,194,217]. Crucially, coarticulation breaks down in stroke[23,24] and PD patients[29], which severely affects their daily life. Considering its importance to both daily life tasks, and in describing movement impairment in clinical populations, the recent increase in scientific interest in upper limb coarticulation is no surprise. The work presented in this thesis contributes to this emerging field by showing that coarticulation can be facilitated through reward. Therefore, this work provides evidence that coarticulation can be phrased in a reinforcement learning (RL) context. Compared to other computational approaches such as optimal control, RL requires substantially less informational and computational resources (which the brain simply might not have)[234,235]. Consequently, phrasing coarticulation as a RL problem provides the exciting opportunity to build a biologically plausible model to study the component processes underlying coarticulation. Further support for this approach comes from the result that dopamine plays a role in coarticulation and that reward-based feedback appears to compensate for the deleterious effects of a D2 antagonist on coarticulation. This suggest that DA plays a role in coarticulation and provides a basis to explore whether coarticulation can be increased through DA agonists such as levodopa.

6.3 Coarticulation vs chunking

In this thesis, coarticulation has been defined as an optimisation process that allows for sequential reaching movements to temporally and spatially converge. This definition is based on the seminal work by Sosnik et al. (2004, 2007, 2015)[25–27]. Specifically, on a temporal level this manifests itself in a progressive reduction in dwell time (stop period between reaching movements) resulting in temporally overlapping movements. In fully coarticulated movements, no stopping can be observed and the two reaching movements are assumed to be planned and executed as a single continuous movement[25–27]. For this to occur, the reaching trajectory has to be spatially modified. Similarly, to anticipatory coarticulation described in speech production[180–182], the trajectory of the preceding reaching movement is adapted to allow for a fluid transition between movements[25–27]. Consequently, in coarticulation a kinematic change in execution can be observed that aligns with the predictions of a minimum-jerk model. According to previous work, this kinematic change results in the development of a new motor primitive that is globally planned and has to run to completion once initiated[25–27]. This potential to form a new motor primitive to substitute executing two discrete primitives is the feature that distinguishes coarticulation from chunking[25–27,176]. Therefore, during chunking individual motor primitives (i.e., button presses) can be aligned temporally resulting in a reduction in response times. Similarly to coarticulation this will lead to in part temporally overlapping movements, which over time will be executed as a holistic unit[176] (i.e., chunk). However, the motor primitives underlying the formed chunk will remain stable and are not kinematically modified[176]. This working definition was based on previous work[25–27,176,180–182] and suggests that coarticulation may represent a distinct yet extreme version of chunking. However, consensus within the field is currently lacking and a recent publication on coarticulation of reaching movements in monkeys termed this optimisation process chunking[194]. This lack of consistent usage of terminology may be in part due to the fractured literature within the field of upper limb coarticulation. Specifically, different schools of thought use different terminology and definitions of chunking and coarticulation might have been informed by the motor unit that is being assessed. For example, coarticulation is predominantly used in the speech production literature[180–182], whereas chunking is more common in the discrete sequence learning literature[20,64,83,190,212] to describe this optimisation process. The motor units that are being assessed (i.e., tongue and lip movements compared to button-pressing, which has been the focus in the discrete sequence learning literature[20,64,83,190,212]) might have informed what chunking and coarticulation represent. In discrete button-pressing tasks, kinematic changes are limited and are often further prevented by enforcing button-presses to be completed before allowing for a subsequent press[236–238]. Furthermore, in most task designs, each finger only ever pressed the same button[189,190,236–239], which severely restricts how the hand is configurated in anticipation for a subsequent button-press in a different location. Therefore, the focus on temporal aspects of chunking might be due to the task design used to investigate this optimisation process. Conversely, in speech production the kinematic modification of both tongue and lip movements in anticipation of a subsequent phoneme is naturally present and therefore will inform the definition of coarticulation[180–182]. Therefore, accounting for these kinematic changes will inform the definition. In an attempt to align both concepts, Godøy et al. (2010) assessed music related actions and defined chunking and coarticulation as optimisation processes that occur simultaneously on a macro and micro level, respectively. Here chunking relates to the parsing and concatenation of sequence elements, whereas coarticulation represents the low-level integration of these sequence elements to improve fluidity[176]. However, to date, a clear differentiation between these two concepts does not exist and it remains an open question whether they represent altogether different optimisation processes or whether coarticulation is an extreme form of chunking.

6.4 Coarticulation as a marker of movement quality

The results presented in this thesis highlight that increases in coarticulation and movement smoothness are closely related. Additionally, with increases in coarticulation, performance showed greater alignment to the predictions of a minimum-jerk model[25,28,99]. However, note here that these results were based on fitting the velocity profile of the minimum-jerk prediction to the actual data and are not a trajectory-based outcome measure. However, existing jerk metrics such as log dimensionless jerk are also based on the velocity profile, which is often preferred due to the reduction of dimensionality that comes with the first derivative of positional data (x, y)[23,85,87]. Movement smoothness is severely impaired in stroke patients, while recovery is characterised by improvements in smoothness[23,24]. Therefore, movement smoothness represents a clinically relevant marker to assess progress during rehabilitation[23,24]. Particularly, movement smoothness has been shown to reduce effort via a decreases in the metabolic cost of the movement[209]. This suggests that coarticulation leads to a smoother and thus more efficient execution of sequential movements. Importantly, the results from Chapter 2 and Chapter 3 highlight that these improvements are robust and reward-independent as seen by a retention of smoothness levels across an additional testing day without reward. Furthermore, coarticulation allows for a reduction in MTs, via a reduction in dwell times around via points, whilst maintaining high levels of accuracy. Hence, coarticulation leads to improvements in the movement smoothness while also enhancing the speed-accuracy trade-off. The results from Chapter 2 indicate that this improvement in the speed-accuracy trade-off is retained during long periods without reward. Taken together, coarticulation could be used as a global marker of movement quality due to its relation to various characteristics of skill. Previously, motor acuity has been proposed as a global marker to assess skill learning[12,240]. Motor acuity was used to describe improvements in the speed-accuracy trade-off and other skill related concepts such as decreases in variability. It represents an effort to summarise several outcome measures into a single parameter; an approach which could be particularly useful in clinical settings and in identifying the neural correlates of motor skill learning[240]. However, compared to motor acuity, coarticulation occurs naturally and represents a strategy to decrease motor effort via a reduction jerk while improving the speed-accuracy trade-off. Therefore, coarticulation does not only represent a summation of several outcome measures but is a naturally occurring process where increases directly lead to improvements in inherent aspects of skill such as movement smoothness and speed. Consequently, measuring upper limb coarticulation could be an important and clinically relevant marker to assess movement quality in both healthy and clinical populations such as stroke patients to track and evaluate recovery.

6.5 Coarticulation within the context of RL

Despite its importance to skilfully executing upper limb reaching sequences, how humans learn to coarticulate is an outstanding research question. This is a crucial gap in the literature because it precludes the development of targeted interventions to enhance coarticulation. The results from this thesis provide novel evidence that coarticulation can be facilitated with reward and that DA plays a role in it. Consequently, coarticulation could be phrased as a RL problem in which the agent aims to maximise future reward. RL has been successful in modelling and predicting animal and human behaviour across a broad range of tasks (for reviews see [128,134,135]). Recently, it was demonstrated that using a RL framework an anthropomorphic robot arm could learn to hit a baseball[241]. Consequently, using RL could provide a framework to understand how humans learn to coarticulate. Further support comes from work showing that phasic DA bursting in the striatum appears to be the neurobiological substrate of the RPE and more specifically its value (i.e., positive, negative or neutral)[138,142]. Work in PD patients ON and OFF medication[164], as well as neuropharmacological studies using D2 antagonists[166], demonstrated that learning on a discrete movement sequence task is impaired. Specifically, they found that chunking was impaired[164,166]. Hence, there is a mounting evidence that DA indeed subserves the computations in RL. Recent computational work found that a robot with more degrees of freedom than necessary to complete a sequential reaching task, coarticulated movements[235]. A hierarchical RL was employed that aimed to achieve two hierarchically organised goals. The primary goal was to complete the given action sequence irrespective of the quality of the chosen actions which ultimately ensured reward (i.e., task completion). The quality of the chosen actions concerned the secondary goal, which centred around minimising the costs associated with movements such as jerk (i.e., coarticulation)[235].

The authors showed that using undirected search with a hierarchical optimisation approach resulted in behaviour that resembles coarticulation[235]. This work could also shed light on the observation that while coarticulation leads to overall improvements in movement quality, it produces sub-optimal solution on the level of individual movements. This is because coarticulation, via increases in curvature, increases the path length of the individual movements. Therefore, potentially more effort has to be invested to complete individual reaching movements which is however paid for by the overall reduction in jerk and improvement in the speed-accuracy trade-off. These considerations will have to be confirmed experimentally and similarly whether such a hierarchical RL framework could predict coarticulation in humans will have to be tested.

6.6 Implications of this work for stroke rehabilitation

Executing upper limb movements, stroke patients exhibit marked impairment in performance. Reaching movements are decomposed into jerky sub-movements, which severely affects their daily life[4,23,215]. Successful recovery is associated with improvements in movement smoothness[23] which has been demonstrated using an array of smoothness metrics[23]. Importantly, it has also been found that over the course of rehabilitation sub-movements appear to decrease in number while the remaining sub-movements grow larger showing pronounced overlapping[24,214]. This analysis of the mechanism underlying stoke recovery is in theory very close to the concept of coarticulation[24,25,214]. Therefore, stroke recovery and coarticulation may follow similar principles. However, even state-of-the-art high intensity upper limb stroke rehabilitation programs prescribe training in which activities of daily living are decomposed into individual movements[242]. Stroke patients are trained on these decomposed movements while the training on movement transitions is minimal[242]. The results from this thesis highlight that coarticulation, and therefore the transition between movements, enhances smoothness and allows for the retention of reward-based improvements in the speed-accuracy trade-off. Consequently, based on these results, stoke recovery may benefit from re-directing training of decomposed, individual movements to entire movements sequences. Additionally, as described in the previous section, coarticulation leads to improvements in efficiency across the entire sequence, while efficiency on individual movement elements might be sub-optimal. Therefore, an entire movement sequence will have to be trained to allow for such overall improvements in movement quality.

Importantly, reward facilitates coarticulation while a D2 antagonist impairs coarticulation when reward-based feedback is not provided. This suggest that DA plays a role in coarticulation and provides a basis to explore whether coarticulation can be increased through a DA agonist such as levodopa.

6.7 Neural representation of coarticulated movements

The important work by Sosnik et al. (2004, 2007, 2015) demonstrated that coarticulation of sequential reaching movements led to the formation of stable continuous actions[25–27]. In one study, they leveraged that human actions cannot be stopped once they have been initiated to investigate whether coarticulated movements are indeed treated as a single motor primitive[27]. Their results show that a stop signal after movement initiation did not lead to the break-down of the coarticulated action. Instead, participants completed the coarticulated movement and stopped on the next target. Therefore, this work suggests that coarticulation leads to the formation of new motor primitives that do not simply represent the summation of two discrete movements[27]. Further evidence comes from work highlighting that new coarticulated primitives are highly generalisable and do not break down when executed slowly. Furthermore, the work presented here agrees with these previous findings showing that with coarticulation individual reaching kinematics are spatially modified allowing for a continuous execution. To address whether such coarticulated movements are encoded as a single movement or as a series of discrete movements, Zimnik and Churchland (2021) took neural recordings of two monkeys while they performed single, a series of two discrete and compound actions[243]. They concluded that neural events during execution of both discrete and compound actions elicits the same neural events in the dorsal premotor and primary motor cortex. Specifically, they found that in compared to single reaching movements, the second movement in the compound action was prepared on the fly during the execution of the first movement[243]. Therefore, it appears that compound movements (i.e., coarticulated movements) are encoded discretely and initiated successively. However, it is important to highlight that the behavioural results do not suggest that reaching movements were coarticulated in the compound action. Specifically, the mean velocity profiles of the compound action are of a distinct two-peaked shape, with a pronounced drop in velocities around the via point. Furthermore, the design predominantly includes reaching transitions with acute angles. Results from chapter 5 highlight in acute angles movement smoothness is optimal when executing two discrete reaching movements. Taken together, this indicates that these sequential reaching movements were not coarticulated[243], but were executed discretely in rapid succession. This would align with their results showing that the second reaching movement was prepared while executing the first reaching movement. Consequently, it remains an open question whether coarticulated reaching movements that exhibit temporal and spatial alignment are encoded as a single or set of discrete movements at a neural level.

6.8 Limitations of the presented work

A central finding presented in this thesis resolves around reaching movements becoming smoother with increases in coarticulation. This has been shown using the jerk metric spectral arc length and by comparing reaching performance to the predictions of a minimum jerk model (Chapter 3); findings that are in line with previous work by Sosnik et al. (2004, 2007, 2015)[25–27]. Based on recent work demonstrating that movement smoothness is related to energetic efficiency[209], it has been concluded that increases in coarticulation lead to more efficient movement patterns via increases in smoothness. However, within the scope of the thesis this has not been assessed experimentally. It is, therefore, paramount to conduct experiments to provide causal evidence that demonstrates that coarticulation indeed leads to increases in energetic efficiency.

Similarly, results in Chapter 3 show that improvements in MT are maintained during long periods without reward; a finding that was correlated with coarticulation. Therefore, improvements in the retention of reward-based performance gains were related to coarticulation. Based on previous work[209] it was hypothesised that the observed improvements in smoothness may account for these findings considering that there is theoretical evidence that increases in smoothness lead to improvements in movement efficiency, which may explain why participants continued to score fast MTs even when not being rewarded (i.e., only because it is energetically feasible to do so). However, despite this intuitive logic, factors other than efficiency via coarticulation could account for these results. Within this context it is of relevance to highlight that coarticulation describes a change in the underlying reaching kinematics that is independent of movement speed. Thus, two consecutive movements can be executed in a coarticulated manner (i.e., as a single movement) irrespective of how fast the movement is executed. Therefore, despite the results presented in this thesis, other factors other than coarticulation could account for the observed retention of reward-based improvements in MTs (however, note that coarticulation will always reduce the dwell time around the via point which will reduce MTs regardless of movement speed). Work on use-dependent plasticity (UDP) has shown that repeated movements towards a specific target introduce a bias towards the same target during execution in a subsequent trial[79]. Therefore, executing the CSR task repeatedly fast may shape UDP and introduce a bias to continue to perform fast movements even without the availability of reward. Future experiments will have to be designed to causally understand whether coarticulation, UDP or other factors drive the observed improvements in the retention of reward-based performance gains.

6.9 Conclusions

While reward and its effect on shaping behaviour has been a research focus for many decades, research on upper limb coarticulation has emerged more recently. Due to its impact on movement efficiency via increases in smoothness, upper limb coarticulation underlies fluid generation of sequential movements. Crucially, coarticulation breaks down in clinical populations such as Parkinson's disease and stroke patients, which severely affects their daily life[23,43,164]. Capitalising on the effect of reward in enhancing complex motor skill learning[158,177,244] the work presented in this thesis aimed to investigate whether coarticulation can be enhance with reward. Results across experiments have found that reward can enhance coarticulation, which led to increases in movement smoothness and speed which were retained across a long period without reward. Extending these findings, a neuropharmacological modulation (chapter 4) was used to investigate whether dopamine plays a role in coarticulation. The results suggest that the dopamine antagonist haloperidol impairs coarticulation which can be compensated for with reward-based feedback. Finally, it was shown that coarticulation is guided by the principal of maximum smoothness which explains why coarticulation is constrained by the transition angle between reaching movements (chapter 5). The work presented here provides evidence that coarticulation could be a useful tool in rehabilitation especially if it is paired with reward and furthermore advocates for a more widespread use of measuring coarticulation to evaluate movement efficiency in both healthy and clinical populations such as Parkinson’s disease and stroke patients.

References

1. Hamann, S. B., Ely, T. D., Grafton, S. T. & Kilts, C. D. Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience* **2**, 289–293 (1999).
2. Dayan, P. & Daw, N. D. Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience* **8**, 429–453 (2008).
3. Chen, X., Holland, P. & Galea, J. M. The effects of reward and punishment on motor skill learning. *Current Opinion in Behavioral Sciences* **20**, 83–88 (2018).
4. Goodman, R. N. *et al.* Increased reward in ankle robotics training enhances motor control and cortical efficiency in stroke. *J Rehabil Res Dev* **51**, 213–227 (2014).
5. Quattrocchi, G., Greenwood, R., Rothwell, J. C., Galea, J. M. & Bestmann, S. Reward and punishment enhance motor adaptation in stroke. *J. Neurol. Neurosurg. Psychiatry* **88**, 730–736 (2017).
6. Halsband, U. & Lange, R. K. Motor learning in man: A review of functional and clinical studies. *Journal of Physiology-Paris* **99**, 414–424 (2006).
7. Luft, A. R. & Buitrago, M. M. Stages of motor skill learning. *Mol Neurobiol* **32**, 205–216 (2005).
8. Dahms, C., Brodoehl, S., Witte, O. W. & Klingner, C. M. The importance of different learning stages for motor sequence learning after stroke. *Human Brain Mapping* **41**, 270–286 (2020).
9. Kelso, J. A. S. *Human Motor Behavior: An Introduction*. (Psychology Press, 1982).
10. Shadmehr, R. & Mussa-Ivaldi, F. A. Adaptive representation of dynamics during learning of a motor task. *J. Neurosci.* **14**, 3208–3224 (1994).
11. Tseng Yw, Diedrichsen J, Krakauer Jw, Shadmehr R, & Bastian Aj. Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *Journal of neurophysiology* vol. 98 https://pubmed.ncbi.nlm.nih.gov/17507504/ (2007).
12. Shmuelof, L., Krakauer, J. W. & Mazzoni, P. How is a motor skill learned? Change and invariance at the levels of task success and trajectory control. *J Neurophysiol* **108**, 578–594 (2012).

13. Hikosaka O. *et al.* Long-term retention of motor skill in macaque monkeys and humans. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale* **147**, 494–504 (2003).
14. Karni, A. & Sagi, D. The time course of learning a visual skill. *Nature* **365**, 250–252 (1993).
15. Nezafat, R., Shadmehr, R. & Holcomb, H. H. Long-term adaptation to dynamics of reaching movements: a PET study. *Exp Brain Res* **140**, 66–76 (2001).
16. Penhune, V. B. & Doyon, J. Dynamic Cortical and Subcortical Networks in Learning and Delayed Recall of Timed Motor Sequences. *J. Neurosci.* **22**, 1397–1406 (2002).
17. Shadmehr, R. & Brashers-Krug, T. Functional Stages in the Formation of Human Long-Term Motor Memory. *J. Neurosci.* **17**, 409–419 (1997).
18. Beldarrain, M. G., Grafman, J., Pascual-Leone, A. & Garcia-Monco, J. C. Procedural learning is impaired in patients with prefrontal lesions. *Neurology* **52**, 1853–1853 (1999).
19. Pohl, P. S., McDowd, J. M., Filion, D. L., Richards, L. G. & Stiers, W. Implicit Learning of a Perceptual-Motor Skill After Stroke. *Phys Ther* **81**, 1780–1789 (2001).
20. Boyd, L. A. *et al.* Motor sequence chunking is impaired by basal ganglia stroke. *Neurobiology of learning and memory* **92**, 35–44 (2009).
21. Bidet-Ildei, C., Pollak, P., Kandel, S., Fraix, V. & Orliaguet, J.-P. Handwriting in patients with Parkinson disease: effect of L-dopa and stimulation of the sub-thalamic nucleus on motor anticipation. *Hum Mov Sci* **30**, 783–791 (2011).
22. Mazzoni, P., Hristova, A. & Krakauer, J. W. Why Don't We Move Faster? Parkinson's Disease, Movement Vigor, and Implicit Motivation. *J. Neurosci.* **27**, 7105–7116 (2007).
23. Rohrer, B. *et al.* Movement Smoothness Changes during Stroke Recovery. *The Journal of Neuroscience* **22**, 8297 (2002).
24. Rohrer, B. *et al.* Submovements Grow Larger, Fewer, and More Blended during Stroke Recovery. *Motor Control* **8**, 472–483 (2004).

25. Sosnik, R., Hauptmann, B., Karni, A. & Flash, T. When practice leads to co-articulation: the evolution of geometrically defined movement primitives. *Exp Brain Res* **156**, 422–438 (2004).

26. Sosnik, R., Chaim, E. & Flash, T. Stopping is not an option: the evolution of unstoppable motion elements (primitives). *Journal of Neurophysiology* **114**, 846 (2015).

27. Sosnik, R., Flash, T., Hauptmann, B. & Karni, A. The acquisition and implementation of the smoothness maximization motion strategy is dependent on spatial accuracy demands. *Exp Brain Res* **176**, 311–331 (2007).

28. Todorov, E. & Jordan, M. I. Smoothness Maximization Along a Predefined Path Accurately Predicts the Speed Profiles of Complex Arm Movements. *Journal of Neurophysiology* **80**, 696–714 (1998).

29. Lange, K. W. *et al.* Brain dopamine and kinematics of graphomotor functions. *Hum Mov Sci* **25**, 492–509 (2006).

30. Shadmehr, R., Brandt, J. & Corkin, S. Time-Dependent Motor Memory Processes in Amnesic Subjects. *Journal of Neurophysiology* **80**, 1590–1597 (1998).

31. Ashby, F. G. & Ell, S. W. Single versus Multiple Systems of Learning and Memory. in *Stevens' Handbook of Experimental Psychology* (American Cancer Society, 2002). doi:10.1002/0471214426.pas0416.

32. Kleynen, M. *et al.* Using a Delphi Technique to Seek Consensus Regarding Definitions, Descriptions and Classification of Terms Related to Implicit and Explicit Forms of Motor Learning. *PLOS ONE* **9**, e100227 (2014).

33. Fletcher, P. C. *et al.* On the Benefits of not Trying: Brain Activity and Connectivity Reflecting the Interactions of Explicit and Implicit Sequence Learning. *Cereb Cortex* **15**, 1002–1015 (2005).

34. Song, S., Marks, B., Howard, J. H. & Howard, D. V. Evidence for parallel explicit and implicit sequence learning systems in older adults. *Behavioural Brain Research* **196**, 328–332 (2009).

35. Reber, P. J. & Squire, L. R. Encapsulation of Implicit and Explicit Memory in Sequence Learning. *Journal of Cognitive Neuroscience* **10**, 248–263 (1998).

36. Keele, S. W., Ivry, R., Mayr, U., Hazeltine, E. & Heuer, H. The cognitive and neural architecture of sequence representation. *Psychological Review* **110**, 316–339 (2003).

37. Shanks, D. R., Rowland, L. A. & Ranger, M. S. Attentional load and implicit sequence learning. *Psychological Research* **69**, 369–382 (2005).

38. Penhune, V. B. & Steele, C. J. Parallel contributions of cerebellar, striatal and M1 mechanisms to motor sequence learning. *Behavioural Brain Research* **226**, 579–591 (2012).

39. Ghilardi, M. F., Moisello, C., Silvestri, G., Ghez, C. & Krakauer, J. W. Learning of a Sequential Motor Skill Comprises Explicit and Implicit Components That Consolidate Differently. *Journal of Neurophysiology* **101**, 2218–2229 (2009).

40. Doya K. Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current opinion in neurobiology* vol. 10 https://pubmed.ncbi.nlm.nih.gov/11240282/ (2000).

41. Hikosaka O *et al.* Parallel neural networks for learning sequential procedures. *Trends in neurosciences* vol. 22 https://pubmed.ncbi.nlm.nih.gov/10481194/ (1999).

42. Savion-Lemieux T & Penhune Vb. The effects of practice and delay on motor skill learning and retention. *Experimental brain research* vol. 161 https://pubmed.ncbi.nlm.nih.gov/15551084/ (2005).

43. Wilkinson, L., Khan, Z. & Jahanshahi, M. The role of the basal ganglia and its cortical connections in sequence learning: Evidence from implicit and explicit sequence learning in Parkinson's disease. *Neuropsychologia* **47**, 2564–2573 (2009).

44. Orrell, A. J., Eves, F. F., Masters, R. S. W. & Macmahon, K. M. M. Implicit sequence learning processes after unilateral stroke. *Neuropsychological Rehabilitation* **17**, 335–354 (2007).

45. Kantak, S. S., Mummidisetty, C. K. & Stinear, J. W. Primary motor and premotor cortex in implicit sequence learning – evidence for competition between implicit and explicit human motor memory systems. *European Journal of Neuroscience* **36**, 2710–2715 (2012).

46. Fitts, P. M. & Posner, M. I. *Human performance*. (Brooks/Cole Pub. Co., 1967).

47. Censor N, Sagi D, & Cohen Lg. Common mechanisms of human perceptual and motor learning. *Nature reviews. Neuroscience* vol. 13 https://pubmed.ncbi.nlm.nih.gov/22903222/ (2012).

48. Doyon J, Penhune V, & Ungerleider Lg. Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia* vol. 41 https://pubmed.ncbi.nlm.nih.gov/12457751/ (2003).

49. Karni, A. *et al.* The acquisition of skilled motor performance: Fast and slow experience-driven changes in primary motor cortex. *PNAS* **95**, 861–868 (1998).

50. Hikosaka O, Nakamura K, Sakai K, & Nakahara H. Central mechanisms of motor skill learning. *Current opinion in neurobiology* vol. 12 https://pubmed.ncbi.nlm.nih.gov/12015240/ (2002).

51. Eliassen Jc, Souza T, & Sanes Jn. Human brain activation accompanying explicitly directed movement sequence learning. *Experimental brain research* vol. 141 https://pubmed.ncbi.nlm.nih.gov/11715072/ (2001).

52. Lewis Pa & Miall Rc. Distinct systems for automatic and cognitively controlled time measurement: evidence from neuroimaging. *Current opinion in neurobiology* vol. 13 https://pubmed.ncbi.nlm.nih.gov/12744981/ (2003).

53. Collins, P., Roberts, A. C., Dias, R., Everitt, B. J. & Robbins, T. W. Perseveration and strategy in a novel spatial self-ordered sequencing task for nonhuman primates: effects of excitotoxic lesions and dopamine depletions of the prefrontal cortex. *J Cogn Neurosci* **10**, 332–354 (1998).

54. Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E. & Ramnani, N. A probabilistic MR atlas of the human cerebellum. *Neuroimage* **46**, 39–46 (2009).

55. Miall, R. C. *et al.* Modulation of linguistic prediction by TDCS of the right lateral cerebellum. *Neuropsychologia* **86**, 103–109 (2016).

56. Ito, M. Mechanisms of motor learning in the cerebellum11Published on the World Wide Web on 24 November 2000. *Brain Research* **886**, 237–245 (2000).

57. Hardwick, R. M., Rottschy, C., Miall, R. C. & Eickhoff, S. B. A quantitative meta-analysis and review of motor learning in the human brain. *NeuroImage* **67**, 283–297 (2013).

58. Bastian, A. J. Learning to predict the future: the cerebellum adapts feedforward movement control. *Curr Opin Neurobiol* **16**, 645–649 (2006).

59. Doyon, J. & Benali, H. Reorganization and plasticity in the adult brain during learning of motor skills. *Curr Opin Neurobiol* **15**, 161–167 (2005).

60. Jin, X., Tecuapetla, F. & Costa, R. M. Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. *Nat Neurosci* **17**, 423–430 (2014).

61. Popp, N. J., Yokoi, A., Gribble, P. L. & Diedrichsen, J. The effect of instruction on motor skill learning. *Journal of Neurophysiology* **124**, 1449–1457 (2020).

62. Hardwick, R. M., Forrence, A. D., Krakauer, J. W. & Haith, A. M. Time-dependent competition between goal-directed and habitual response preparation. *Nature Human Behaviour* **3**, 1252–1262 (2019).

63. Ostlund, S. B. & Balleine, B. W. On habits and addiction: An associative analysis of compulsive drug seeking. *Drug Discov Today Dis Models* **5**, 235–245 (2008).

64. Ariani, G. & Diedrichsen, J. Sequence learning is driven by improvements in motor planning. *J. Neurophysiol.* **121**, 2088–2100 (2019).

65. Fitts, P. M. The information capacity of the human motor system in controlling the amplitude of movement. *J Exp Psychol* **47**, 381–391 (1954).

66. Heitz, R. P. The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Front. Neurosci.* **8**, (2014).

67. Yarrow, K., Brown, P. & Krakauer, J. W. Inside the brain of an elite athlete: the neural processes that support high achievement in sports. *Nat Rev Neurosci* **10**, 585–596 (2009).

68. Sanger, T. D. Arm trajectories in dyskinetic cerebral palsy have increased random variability. *J Child Neurol* **21**, 551–557 (2006).

69. Reis, J. *et al.* Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. *Proc Natl Acad Sci U S A* **106**, 1590–1595 (2009).

70. Lefebvre, S., Dricot, L., Gradkowski, W., Laloux, P. & Vandermeeren, Y. Brain activations underlying different patterns of performance improvement during early motor skill learning. *NeuroImage* **62**, 290–299 (2012).
71. Dayan, E. & Cohen, L. G. Neuroplasticity Subserving Motor Skill Learning. *Neuron* **72**, 443–454 (2011).
72. Kawagoe, R., Takikawa, Y. & Hikosaka, O. Reward-Predicting Activity of Dopamine and Caudate Neurons—A Possible Mechanism of Motivational Control of Saccadic Eye Movement. *Journal of Neurophysiology* **91**, 1013–1024 (2004).
73. Manohar, S. G. *et al.* Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control. *Current Biology* **25**, 1707 (2015).
74. Summerside, E. M., Shadmehr, R. & Ahmed, A. A. Control of Movement: Vigor of reaching movements: reward discounts the cost of effort. *Journal of Neurophysiology* **119**, 2347 (2018).
75. Codol, O., Holland, P. J., Manohar, S. G. & Galea, J. M. Reward-based improvements in motor control are driven by multiple error-reducing mechanisms. *J. Neurosci.* (2020) doi:10.1523/JNEUROSCI.2646-19.2020.
76. Robertson, E. M. The Serial Reaction Time Task: Implicit Motor Skill Learning? *J. Neurosci.* **27**, 10073–10075 (2007).
77. Craddock, P., Molet, M. & Miller, R. R. Reaction time as a measure of human associative learning. *Behavioural Processes* **90**, 189–197 (2012).
78. Haith, A. M., Pakpoor, J. & Krakauer, J. W. Independence of Movement Preparation and Movement Initiation. *J Neurosci* **36**, 3007–3015 (2016).
79. Mawase, F., Lopez, D., Celnik, P. A. & Haith, A. M. Movement Repetition Facilitates Response Preparation. *Cell Reports* **24**, 801–808 (2018).
80. Wong, A. L., Lindquist, M. A., Haith, A. M. & Krakauer, J. W. Explicit knowledge enhances motor vigor and performance: motivation versus practice in sequence tasks. *Journal of Neurophysiology* **114**, 219–232 (2015).

81. Haith, A. M. & Krakauer, J. W. The multiple effects of practice: skill, habit and reduced cognitive load. *Current Opinion in Behavioral Sciences* **20**, 196–201 (2018).

82. Dhawale, A. K., Smith, M. A. & Ölveczky, B. P. The Role of Variability in Motor Learning. *Annu Rev Neurosci* **40**, 479–498 (2017).

83. Diedrichsen, J. & Kornysheva, K. Motor skill learning between selection and execution. *Trends Cogn. Sci. (Regul. Ed.)* **19**, 227–233 (2015).

84. Wu, H. G., Miyamoto, Y. R., Gonzalez Castro, L. N., Ölveczky, B. P. & Smith, M. A. Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nat. Neurosci.* **17**, 312–321 (2014).

85. Balasubramanian, S., Melendez-Calderon, A., Roby-Brami, A. & Burdet, E. On the analysis of movement smoothness. *Journal of NeuroEngineering and Rehabilitation* **12**, 112 (2015).

86. Sejnowski, T. J. Making smooth moves. *Nature* **394**, 725–726 (1998).

87. Gulde, P. & Hermsdörfer, J. Smoothness Metrics in Complex Movement Tasks. *Front. Neurol.* **9**, (2018).

88. Cohen, R. G. & Sternad, D. Variability In Motor Learning: Relocating, Channeling and Reducing Noise. *Exp Brain Res* **193**, 69–83 (2009).

89. Faisal, A. A., Selen, L. P. J. & Wolpert, D. M. Noise in the nervous system. *Nat Rev Neurosci* **9**, 292–303 (2008).

90. Harris, C. M. & Wolpert, D. M. Signal-dependent noise determines motor planning. *Nature* **394**, 780–784 (1998).

91. Sporn, S., Hein, T. & Herrojo Ruiz, M. Alterations in the amplitude and burst rate of beta oscillations impair reward-dependent motor learning in anxiety. *eLife* **9**, e50654 (2020).

92. Oelveczky, B. P., Otchy, T. M., Goldberg, J. H., Aronov, D. & Fee, M. S. Changes in the neural control of a complex motor sequence during learning. *J Neurophysiol* **106**, 386–397 (2011).

93. Chen, X., Mohr, K. & Galea, J. M. Predicting explorative motor learning using decision-making and motor noise. *PLoS Comput Biol* **13**, e1005503 (2017).

94. Sternad, D. It's not (only) the mean that matters: variability, noise and exploration in skill learning. *Current Opinion in Behavioral Sciences* **20**, 183–195 (2018).

95. Berthier, N. E. & Keen, R. Development of reaching in infancy. *Exp Brain Res* **169**, 507–518 (2006).

96. Balasubramanian, S., Melendez-Calderon, A. & Burdet, E. A robust and sensitive metric for quantifying movement smoothness. *IEEE Trans Biomed Eng* **59**, 2126–2136 (2012).

97. Franklin, D. W. *et al.* CNS Learns Stable, Accurate, and Efficient Movements Using a Simple Algorithm. *J. Neurosci.* **28**, 11165–11173 (2008).

98. Hogan, N. & Sternad, D. Sensitivity of Smoothness Measures to Movement Duration, Amplitude and Arrests. *J Mot Behav* **41**, 529–534 (2009).

99. Flash, T. & Hogan, N. The coordination of arm movements: an experimentally confirmed mathematical model. *J. Neurosci.* **5**, 1688–1703 (1985).

100. Viviani, P. & Flash, T. Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning. *Journal of Experimental Psychology: Human Perception and Performance* **21**, 32–53 (1995).

101. Franklin, D. W. & Wolpert, D. M. Computational Mechanisms of Sensorimotor Control. *Neuron* **72**, 425–442 (2011).

102. Guigon, E., Baraduc, P. & Desmurget, M. Computational Motor Control: Redundancy and Invariance. *Journal of Neurophysiology* **97**, 331–347 (2007).

103. Hamilton, A. F. de C. & Wolpert, D. M. Controlling the statistics of action: obstacle avoidance. *J Neurophysiol* **87**, 2434–2440 (2002).

104. Jones, K. E., Hamilton, A. F. de C. & Wolpert, D. M. Sources of Signal-Dependent Noise During Isometric Force Production. *Journal of Neurophysiology* **88**, 1533–1544 (2002).

105. Burdet, E., Osu, R., Franklin, D. W., Milner, T. E. & Kawato, M. The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature* **414**, 446–449 (2001).

106. Takahashi, C. D., Scheidt, R. A. & Reinkensmeyer, D. J. Impedance Control and Internal Model Formation When Reaching in a Randomly Varying Dynamical Environment. *Journal of Neurophysiology* **86**, 1047–1051 (2001).

107. Dorfman, L. J. & Bosley, T. M. Age-related changes in peripheral and central nerve conduction in man. *Neurology* **29**, 38–38 (1979).

108. Lindle, R. S. *et al.* Age and gender comparisons of muscle strength in 654 women and men aged 20–93 yr. *Journal of Applied Physiology* **83**, 1581–1587 (1997).

109. Owsley, C., Sekuler, R. & Siemsen, D. Contrast sensitivity throughout adulthood. *Vision Research* **23**, 689–699 (1983).

110. Kay, B. A. The dimensionality of movement trajectories and the degrees of freedom problem: A tutorial. *Human Movement Science* **7**, 343–364 (1988).

111. Breteler, M. D. K., Hondzinski, J. M. & Flanders, M. Drawing Sequences of Segments in 3D: Kinetic Influences on Arm Configuration. *Journal of Neurophysiology* (2003) doi:10.1152/jn.01062.2002.

112. Klein Breteler, M. D., Meulenbroek, R. G. J. & Gielen, S. C. A. M. An evaluation of the minimum-jerk and minimum torque-change principles at the path, trajectory, and movement-cost levels. *Motor Control* **6**, 69–83 (2002).

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

114. Todorov, E. Optimality principles in sensorimotor control. *Nature Neuroscience* **7**, 907–915 (2004).

115. Haruno, M. & Wolpert, D. M. Optimal Control of Redundant Muscles in Step-Tracking Wrist Movements. *Journal of Neurophysiology* **94**, 4244–4255 (2005).

116. Chen-Harris, H., Joiner, W. M., Ethier, V., Zee, D. S. & Shadmehr, R. Adaptive Control of Saccades via Internal Feedback. *J. Neurosci.* **28**, 2804–2813 (2008).

117. Harris, C. M. & Wolpert, D. M. The Main Sequence of Saccades Optimizes Speed-accuracy Trade-off. *Biol Cybern* **95**, 21–29 (2006).

118. Braun, D. A., Aertsen, A., Wolpert, D. M. & Mehring, C. Learning Optimal Adaptation Strategies in Unpredictable Motor Tasks. *J. Neurosci.* **29**, 6472–6478 (2009).

119. Fagg, A. H., Shah, A. & Barto, A. G. A Computational Model of Muscle Recruitment for Wrist Movements. *Journal of Neurophysiology* **88**, 3348–3358 (2002).

120. Pandy, M. G., Zajac, F. E., Sim, E. & Levine, W. S. An optimal control model for maximum-height human jumping. *Journal of Biomechanics* **23**, 1185–1198 (1990).

121. Liu, D. & Todorov, E. Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. *J Neurosci* **27**, 9354–9368 (2007).

122. Pruszynski, J. A., Kurtzer, I. & Scott, S. H. Rapid motor responses are appropriately tuned to the metrics of a visuospatial task. *J Neurophysiol* **100**, 224–238 (2008).

123. Kurtzer, I. L., Pruszynski, J. A. & Scott, S. H. Long-Latency Reflexes of the Human Arm Reflect an Internal Model of Limb Dynamics. *Current Biology* **18**, 449–453 (2008).

124. Sutton, R. S. & Barto, A. G. Reinforcement Learning: An Introduction. 352.

125. Thrun, S. B. *Efficient Exploration In Reinforcement Learning*. (1992).

126. Dam, G. & Körding, K. Exploration and Exploitation During Sequential Search. *Cognitive Science* **33**, 530–541 (2009).

127. Gläscher, J., Daw, N., Dayan, P. & O'Doherty, J. P. States versus Rewards: Dissociable Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement Learning. *Neuron* **66**, 585–595 (2010).

128. Daw, N. D. & Tobler, P. N. Chapter 15 - Value Learning through Reinforcement: The Basics of Dopamine and Reinforcement Learning. in *Neuroeconomics (Second Edition)* (eds. Glimcher, P. W. & Fehr, E.) 283–298 (Academic Press, 2014). doi:10.1016/B978-0-12-416008-8.00015-2.

129. Gershman, S. J. The Successor Representation: Its Computational Logic and Neural Substrates. *J. Neurosci.* **38**, 7193–7200 (2018).

130. Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J. & Daw, N. D. Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLOS Computational Biology* **13**, e1005768 (2017).

131. Otto, A. R., Gershman, S. J., Markman, A. B. & Daw, N. D. The Curse of Planning: Dissecting Multiple Reinforcement-Learning Systems by Taxing the Central Executive. *Psychol Sci* **24**, 751–761 (2013).

132. Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A. & Daw, N. D. Working-memory capacity protects model-based learning from stress. *PNAS* **110**, 20941–20946 (2013).

133. Schad, D. J. *et al.* Processing speed enhances model-based over model-free reinforcement learning in the presence of high working memory functioning. *Front. Psychol.* **5**, (2014).

134. Dayan, P. & Balleine, B. W. Reward, motivation, and reinforcement learning. *Neuron* **36**, 285–298 (2002).

135. Glimcher, P. W. Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *PNAS* **108**, 15647–15654 (2011).

136. Bayer, H. M. & Glimcher, P. W. Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal. *Neuron* **47**, 129–141 (2005).

137. Wise, R. A. Dopamine, learning and motivation. *Nature Reviews Neuroscience* **5**, 483–494 (2004).

138. Schultz, W., Dayan, P. & Montague, P. R. A Neural Substrate of Prediction and Reward. *Science* **275**, 1593–1599 (1997).

139. Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J. & Frith, C. D. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* **442**, 1042–1045 (2006).

140. Chowdhury, R. *et al.* Dopamine restores reward prediction errors in old age. *Nature Neuroscience* **16**, 648–653 (2013).

141. Montague, P. R., Dayan, P. & Sejnowski, T. J. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* **16**, 1936–1947 (1996).

142. Schultz, W. Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *Journal of Neurophysiology* **56**, 1439–1461 (1986).

143. Bayer, H. M., Lau, B. & Glimcher, P. W. Statistics of Midbrain Dopamine Neuron Spike Trains in the Awake Primate. *Journal of Neurophysiology* **98**, 1428–1439 (2007).

144. Menegas, W. *et al.* Dopamine neurons projecting to the posterior striatum form an anatomically distinct subclass. *eLife* **4**, e10032.

145. Utter, A. A. & Basso, M. A. The basal ganglia: An overview of circuits and function. *Neuroscience & Biobehavioral Reviews* **32**, 333–342 (2008).

146. Calabresi, P. *et al.* Synaptic transmission in the striatum: from plasticity to neurodegeneration. *Progress in Neurobiology* **61**, 231–265 (2000).

147. Seamans, J. K., Durstewitz, D., Christie, B. R., Stevens, C. F. & Sejnowski, T. J. Dopamine D1/D5 receptor modulation of excitatory synaptic inputs to layer V prefrontal cortex neurons. *PNAS* **98**, 301–306 (2001).

148. Frank, M. J. & O’Reilly, R. C. A mechanistic account of striatal dopamine function in human cognition: Psychopharmacological studies with cabergoline and haloperidol. *Behavioral Neuroscience* **120**, 497–517 (2006).

149. Grogan, J. P. *et al.* Effects of dopamine on reinforcement learning and consolidation in Parkinson’s disease. *Elife* **6**, (2017).

150. Coulthard, E. J. *et al.* Distinct roles of dopamine and subthalamic nucleus in learning and probabilistic decision making. *Brain* **135**, 3721–3734 (2012).

151. Smittenaar, P. *et al.* Decomposing effects of dopaminergic medication in Parkinson’s disease on probabilistic action selection – learning or performance? *European Journal of Neuroscience* **35**, 1144–1151 (2012).

152. Choi, J. E. S., Vaswani, P. A. & Shadmehr, R. Vigor of movements and the cost of time in decision making. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **34**, 1212–1223 (2014).

153. Reppert, T. R., Lempert, K. M., Glimcher, P. W. & Shadmehr, R. Modulation of Saccade Vigor during Value-Based Decision Making. *J. Neurosci.* **35**, 15369–15378 (2015).

154. Xu-Wilson, M., Zee, D. S. & Shadmehr, R. The intrinsic value of visual information affects saccade velocities. *Experimental Brain Research* **196**, 475–481 (2009).

155. Manohar, S. G., Finzi, R. D., Drew, D. & Husain, M. Distinct Motivational Effects of Contingent and Noncontingent Rewards. *Psychol Sci* **28**, 1016–1026 (2017).

156. Opris, I., Lebedev, M. & Nelson, R. J. Motor Planning under Unpredictable Reward: Modulations of Movement Vigor and Primate Striatum Activity. *Front. Neurosci.* **5**, (2011).

157. Pasquereau, B. *et al.* Shaping of motor responses by incentive values through the basal ganglia. *The Journal of Neuroscience* **27**, 1176–1183 (2007).

158. Wächter, T., Lungu, O. V., Liu, T., Willingham, D. T. & Ashe, J. Differential effect of reward and punishment on procedural learning. *J. Neurosci.* **29**, 436–443 (2009).

159. Steel, A., Silson, E. H., Stagg, C. J. & Baker, C. I. The impact of reward and punishment on skill learning depends on task demands. *Sci Rep* **6**, (2016).

160. Jin, X. & Costa, R. M. Shaping Action Sequences in Basal Ganglia Circuits. *Curr Opin Neurobiol* **33**, 188–196 (2015).

161. Shin, J. H., Kim, D. & Jung, M. W. Differential coding of reward and movement information in the dorsomedial striatal direct and indirect pathways. *Nature Communications* **9**, 404 (2018).

162. Shin, J. H., Song, M., Paik, S.-B. & Jung, M. W. Spatial organization of functional clusters representing reward and movement information in the striatal direct and indirect pathways. *PNAS* **117**, 27004–27015 (2020).

163. Tremblay, P.-L. *et al.* Motor sequence learning in primate: Role of the D2 receptor in movement chunking during consolidation. *Behavioural Brain Research* **198**, 231–239 (2009).

164. Tremblay, P.-L. *et al.* Movement chunking during sequence learning is a dopamine-dependant process: a study conducted in Parkinson's disease. *Exp Brain Res* **205**, 375–385 (2010).

165. Seo, M., Beigi, M., Jahanshahi, M. & Averbeck, B. B. Effects of Dopamine Medication on Sequence Learning with Stochastic Feedback in Parkinson's Disease. *Front. Syst. Neurosci.* **4**, (2010).

166. Lissek, S. *et al.* Opposing effects of dopamine antagonism in a motor sequence task—tiapride increases cortical excitability and impairs motor learning. *Front. Behav. Neurosci.* **8**, (2014).

167. Augustin, S. M., Loewinger, G. C., O'Neal, T. J., Kravitz, A. V. & Lovinger, D. M. Dopamine D2 receptor signaling on iMSNs is required for initiation and vigor of learned actions. *Neuropsychopharmacology* **45**, 2087–2097 (2020).

168. Grogan, J. P., Sandhu, T. R., Hu, M. T. & Manohar, S. G. Dopamine promotes instrumental motivation, but reduces reward-related vigour. *eLife* **9**, e58321 (2020).

169. Zénon, A., Devesse, S. & Olivier, E. Dopamine Manipulation Affects Response Vigor Independently of Opportunity Cost. *J. Neurosci.* **36**, 9516–9525 (2016).

170. Niv, Y. Cost, benefit, tonic, phasic: what do response rates tell us about dopamine and motivation? *Ann N Y Acad Sci* **1104**, 357–376 (2007).

171. Niv, Y., Daw, N. D., Joel, D. & Dayan, P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology* **191**, 507–520 (2007).

172. Heien, M. L. A. V. *et al.* Real-time measurement of dopamine fluctuations after cocaine in the brain of behaving rats. *PNAS* **102**, 10023–10028 (2005).

173. Hogan, N. & Sternad, D. On rhythmic and discrete movements: reflections, definitions and implications for motor control. *Exp Brain Res* **181**, 13–30 (2007).

174. Feldman, A. G. Superposition of motor programs—I. Rhythmic forearm movements in man. *Neuroscience* **5**, 81–90 (1980).

175. Jiménez, L., Vaquero, J. M. M. & Lupiáñez, J. Qualitative differences between implicit and explicit sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **32**, 475–490 (2006).

176. Godøy, R. I., Jensenius, A. R. & Nymoen, K. Chunking in Music by Coarticulation. *Acta Acustica united with Acustica* **96**, 690–700 (2010).

177. Abe, M. *et al.* Reward improves long-term retention of a motor memory through induction of offline memory gains. *Curr. Biol.* **21**, 557–562 (2011).

178. Takikawa, Y., Kawagoe, R., Itoh, H., Nakahara, H. & Hikosaka, O. Modulation of saccadic eye movements by predicted reward outcome. *Exp Brain Res* **142**, 284–291 (2002).

179. Codol, O., Galea, J. M., Jalali, R. & Holland, P. J. Reward-driven enhancements in motor control are robust to TMS manipulation. *Exp Brain Res* 1–13 (2020) doi:10.1007/s00221-020-05802-1.

180. Kent, R. D. & Minifie, F. D. Coarticulation in recent speech production models. *Journal of Phonetics* **5**, 115–133 (1977).

181. Fowler, C. A. Coarticulation and theories of extrinsic timing. 21.

182. Kuhnert, B. & Nolan, F. The origin of coarticulation. *FIPKM)* **35**, (1997).

183. Rosenbaum, D. A. 1 - INTRODUCTION. in *Human Motor Control* (ed. Rosenbaum, D. A.) 3–34 (Academic Press, 1991). doi:10.1016/B978-0-08-057108-9.50006-8.

184. Jerde, T. E., Soechting, J. F. & Flanders, M. Coarticulation in Fluent Fingerspelling. *The Journal of Neuroscience* **23**, 2383 (2003).

185. Winges, S. A. & Furuya, S. Distinct digit kinematics by professional and amateur pianists. *Neuroscience* **0**, 643 (2015).

186. Soechting, J. F. & Flanders, M. Organization of sequential typing movements. *Journal of Neurophysiology* (1992) doi:10.1152/jn.1992.67.5.1275.

187. Soechting, J. F. & Flanders, M. Flexibility and Repeatability of Finger Movements During Typing: Analysis of Multiple Degrees of Freedom. *J Comput Neurosci* **4**, 29–46 (1997).

188. Rumelhart, D. E. & Norman, D. A. Simulating a Skilled Typist: A Study of Skilled Cognitive-Motor Performance. *Cognitive Science* **6**, 1–36 (1982).

189. Verwey, W. B. Evidence for a multistage model of practice in a sequential movement task. *J EXP PSYCHOL HUMAN* **25**, 1693–1708 (1999).

190. Verwey, W. B. & Dronkert, Y. Practicing a Structured Continuous Key-Pressing Task: Motor Chunking or Rhythm Consolidation? *J Mot Behav* **28**, 71–79 (1996).

191. Wymbs, N. F., Bassett, D. S., Mucha, P. J., Porter, M. A. & Grafton, S. T. Differential recruitment of the sensorimotor putamen and frontoparietal cortex during motor chunking in humans. *Neuron* **74**, 936 (2012).

192. Thompson, J. J., McColeman, C. M., Blair, M. R. & Henrey, A. J. Classic motor chunking theory fails to account for behavioural diversity and speed in a complex naturalistic task. *PLOS ONE* **14**, e0218251 (2019).

193. Friedman, J. & Korman, M. Observation of an expert model induces a skilled movement coordination pattern in a single session of intermittent practice. *Sci Rep* **9**, 1–15 (2019).

194. Ramkumar, P. *et al.* Chunking as the result of an efficiency computation trade-off. *Nat Commun* **7**, 12176 (2016).

195. Shah, A., Barto, A. & Fagg, A. A Dual Process Account of Coarticulation in Motor Skill Acquisition. *Journal of Motor Behavior* **45**, 531 (2013).

196. Sosnik, R., Flash, T., Sterkin, A., Hauptmann, B. & Karni, A. The activity in the contralateral primary motor cortex, dorsal premotor and supplementary motor area is modulated by performance gains. *Frontiers in Human Neuroscience* **8**, (2014).

197. https://www.mathworks.com/matlabcentral/mlc-downloads/downloads/submissions/58403/versions/7/previews/min_jerk_todorov.m/index.html.

198. siva82kb/smoothness. *GitHub* https://github.com/siva82kb/smoothness.

199. Sadnicka, A. *et al.* High motor variability in DYT1 dystonia is associated with impaired visuomotor adaptation. *Scientific Reports* **8**, 1–11 (2018).

200. Pekny, S. E., Izawa, J. & Shadmehr, R. Reward-Dependent Modulation of Movement Variability. *The Journal of Neuroscience* **35**, 4015 (2015).

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

202. Carroll, T. J., McNamee, D., Ingram, J. N. & Wolpert, D. M. Rapid Visuomotor Responses Reflect Value-Based Decisions. *The Journal of Neuroscience* **39**, 3906 (2019).

203. Willingham, D. B. A Neuropsychological Theory of Motor Skill Learning. 27.

204. Doeringer, J. A. & Hogan, N. Serial processing in human movement production. *Neural Networks* **11**, 1345–1356 (1998).

205. Hansen, E., Grimme, B., Reimann, H. & Schöner, G. Anticipatory coarticulation in non-speeded arm movements can be motor-equivalent, carry-over coarticulation always is. *Exp Brain Res* **236**, 1293–1307 (2018).

206. Verwey, W. B., Groen, E. C. & Wright, D. L. The stuff that motor chunks are made of: Spatial instead of motor representations? *Exp Brain Res* **234**, 353–366 (2016).

207. Revol, P. *et al.* Pointing errors in immediate and delayed conditions in unilateral optic ataxia. *Spat Vis* **16**, 347–364 (2003).

208. D’Errico, J. SLM - Shape Language Modeling. https://de.mathworks.com/matlabcentral/fileexchange/24443-slm-shape-language-modeling (2020).

209. Huang, H. J., Kram, R. & Ahmed, A. A. Reduction of Metabolic Cost during Motor Learning of Arm Reaching Dynamics. *Journal of Neuroscience* **32**, 2182–2190 (2012).

210. Gribble, P. L., Mullin, L. I., Cothros, N. & Mattar, A. Role of cocontraction in arm movement accuracy. *J. Neurophysiol.* **89**, 2396–2405 (2003).

211. Schmidt, R. A. & Lee, T. D. *Motor control and learning: A behavioral emphasis, 5th ed.* ix, 581 (Human Kinetics, 2011).

212. Huntley, J., Bor, D., Hampshire, A., Owen, A. & Howard, R. Working memory task performance and chunking in early Alzheimer's disease. *Br J Psychiatry* **198**, 398–403 (2011).

213. Kornysheva, K. *et al.* Neural Competitive Queuing of Ordinal Structure Underlies Skilled Sequential Action. *Neuron* **101**, 1166-1180.e3 (2019).

214. Dipietro, L., Krebs, H. I., Fasoli, S. E., Volpe, B. T. & Hogan, N. Submovement changes characterize generalization of motor recovery after stroke. *Cortex* **45**, 318–324 (2009).

215. Gulde, P., Hughes, C. M. L. & Hermsdörfer, J. Effects of Stroke on Ipsilesional End-Effector Kinematics in a Multi-Step Activity of Daily Living. *Front. Hum. Neurosci.* **11**, (2017).

216. Collins, A. L. *et al.* Dynamic mesolimbic dopamine signaling during action sequence learning and expectation violation. *Sci Rep* **6**, 20231 (2016).

217. Sporn, S., Chen, X. & Galea, J. M. *Reward-based invigoration of sequential reaching*. http://biorxiv.org/lookup/doi/10.1101/2020.06.15.152876 (2020) doi:10.1101/2020.06.15.152876.

218. Matuschek, H. & Kliegl, R. On the ambiguity of interaction and nonlinear main effects in a regime of dependent covariates. *Behav Res* **50**, 1882–1894 (2018).

219. He, K. *et al.* The Statistical Determinants of the Speed of Motor Learning. *PLOS Computational Biology* **12**, e1005023 (2016).

220. Singh, P., Jana, S., Ghosal, A. & Murthy, A. Exploration of joint redundancy but not task space variability facilitates supervised motor learning. *PNAS* **113**, 14414–14419 (2016).

221. Maxwell, J. P., Masters, R. S. W. & Eves, F. F. The role of working memory in motor learning and performance. *Consciousness and Cognition* **12**, 376–402 (2003).

222. Anguera, J. A., Reuter-Lorenz, P. A., Willingham, D. T. & Seidler, R. D. Contributions of Spatial Working Memory to Visuomotor Learning. *Journal of Cognitive Neuroscience* **22**, 1917–1930 (2009).

223. Cortese, A. & Rossi-Arnaud, C. Working memory for ballet moves and spatial locations in professional ballet dancers. *Applied Cognitive Psychology* **24**, 266–286 (2010).

224. Christou, A. I., Miall, R. C., McNab, F. & Galea, J. M. Individual differences in explicit and implicit visuomotor learning and working memory capacity. *Scientific Reports* **6**, 36633 (2016).

225. Holland, P. *et al.* Domain-Specific Working Memory, But Not Dopamine-Related Genetic Variability, Shapes Reward-Based Motor Learning. *J. Neurosci.* **39**, 9383–9396 (2019).

226. Sidarta, A., van Vugt, F. T. & Ostry, D. J. Somatosensory working memory in human reinforcement-based motor learning. *Journal of Neurophysiology* **120**, 3275–3286 (2018).

227. Osu, R. *et al.* Quantifying the quality of hand movement in stroke patients through three-dimensional curvature. *Journal of NeuroEngineering and Rehabilitation* **8**, 62 (2011).

228. BEPPU, H., SUDA, M. & TANAKA, R. ANALYSIS OF CEREBELLAR MOTOR DISORDERS BY VISUALLY GUIDED ELBOW TRACKING MOVEMENT. *Brain* **107**, 787–809 (1984).

229. Kahn, L. E., Zygman, M. L., Rymer, W. Z. & Reinkensmeyer, D. J. Effect of robot-assisted and unassisted exercise on functional reaching in chronic hemiparesis. in *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society* vol. 2 1344–1347 vol.2 (2001).

230. Fetters, L. & Todd, J. Quantitative Assessment of Infant Reaching Movements. *Journal of Motor Behavior* **19**, 147–166 (1987).

231. Walton, M. E. & Bouret, S. What Is the Relationship between Dopamine and Effort? *Trends Neurosci* **42**, 79–91 (2019).

232. Le Heron, C. *et al.* Distinct effects of apathy and dopamine on effort-based decision-making in Parkinson’s disease. *Brain* **141**, 1455–1469 (2018).

233. Tolman, E. C. Cognitive maps in rats and men. *Psychological Review* **55**, 189–208 (1948).

234. Friston, K. What Is Optimal about Motor Control? *Neuron* **72**, 488–498 (2011).

235. Shah, A. Biologically-Based Functional Mechanisms of Motor Skill Acquisition. (2008).

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

237. Beukema, P., Diedrichsen, J. & Verstynen, T. Binding during sequence learning does not alter cortical representations of individual actions. *J. Neurosci.* (2019) doi:10.1523/JNEUROSCI.2669-18.2019.

238. Song, S. & Cohen, L. Impact of conscious intent on chunking during motor learning. *Learn Mem* **21**, 449–451 (2014).

239. Jouen, A.-L. *et al.* Discrete sequence production with and without a pause: the role of cortex, basal ganglia, and cerebellum. *Front Hum Neurosci* **7**, (2013).

240. Shmuelof, L., Yang, J., Caffo, B., Mazzoni, P. & Krakauer, J. W. The neural correlates of learned motor acuity. *J. Neurophysiol.* **112**, 971–980 (2014).

241. Peters, J. & Schaal, S. Reinforcement learning of motor skills with policy gradients. *Neural Networks* **21**, 682–697 (2008).

242. Ward, N. S., Brander, F. & Kelly, K. Intensive upper limb neurorehabilitation in chronic stroke: outcomes from the Queen Square programme. *J Neurol Neurosurg Psychiatry* **90**, 498–506 (2019).

243. Zimnik, A. J. & Churchland, M. M. Independent generation of sequence elements by motor cortex. *Nature Neuroscience* 1–13 (2021) doi:10.1038/s41593-021-00798-5.

244. Dayan, E., Averbeck, B., Richmond, B. & Cohen, L. Stochastic reinforcement benefits skill acquisition. *Learning & memory (Cold Spring Harbor, N.Y.)* **21**, 140–2 (2014).