Technische Universität München
Fakultät für Sport- und Gesundheitswissenschaften

Seeing the future is a fundamental skill: the influence of time-to-target in visuomotor control

Justinas Česonis

Vollständiger Abdruck der von der Fakultät für Sport- und Gesundheitswissenschaften der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Philosophie (Dr. phil.)

genehmigten Dissertation.

Vorsitzender:	Prof. Dr. Henning Wackerhage
Prüfende der Dissertation:	1. Prof. Dr. David W. Franklin
	2. Prof. Dr. Fabrice R. Sarlegna
	3. Prof. Dr. Jeroen B. J. Smeets

Die Dissertation wurde am 13.06.2022 bei der Technischen Universität München eingereicht und durch die Fakultät für Sport- und Gesundheitswissenschaften am 11.01.2023 angenommen.

Abstract

Humans utilise vision in various aspects of their daily interactions with the surrounding environment. One such aspect is using vision as one of the inputs to produce physical movements, e.g. by orienting in space, finding a movement goal before performing an action towards it, or reacting to various changes in the surroundings. However, in order to efficiently apply visual information to the control of movement, human brain needs to solve a feedback control task to map the observed noisy, delayed visual information to an appropriate motor output. The work presented in thesis focusses on these feedback control principles and the computational mechanisms that drive them, in order to better understand the algorithmic processes implemented in human brain. In five presented studies I focus on the human control behaviour in goal directed reaching (studies I-III), as well as in balancing (studies IV and V). The first assesses the optimality of human reaching movements in presence of visual perturbations, verifying earlier evidence that optimal feedback control can well describe such movements. In addition, the results of study I demonstrate that the feedback responses generated by such control are primarily regulated by time-to-target via a characteristic relationship, and not by other kinematic variables such as position, speed or acceleration. As such control requires an accurate estimate of the future in form of time-to-target, study II proposes a computational implementation–a mixed-horizon optimal feedback controller–that provides access to the estimation of this variable. In turn, the mixed-horizon OFC then allows for computational modelling of perturbed goal-directed reaching movements without the need of external information of time-to-target, enabling novel methods for post-hoc evaluations of experimental results, as well as raising new hypothesis. In study III we test our participants in two separate tasks that require different feedback controllers, showing that human behaviour well matches with the predictions of the mixed-horizon model. Moreover, here we also demonstrate that participants can flexibly switch between these feedback controllers on-cue, similarly as they do in feedforward control.

In studies IV and V we explore the mechanisms of visuomotor feedback control in balancing, by first introducing a novel experimental setup of simulated inverted pendulum on the cart, which allows us to flexibly induce various perturbations in experimental design. More specifically, in study IV we then evaluate human control behaviour of such a pendulum of various lengths, showing that it is consistent with the equivalent balance of real, fully mechanical pendulum. Furthermore, we also demonstrate that a proportional-derivative controller with time-lag matching that of the visuomotor

system can well describe the human control behaviour. Finally, in study V we distort the visual feedback, present in a similar balancing task, demonstrating that our participants produced the most stable control when the visual feedback best matched the dynamical behaviour of the pendulum.

The contributions of work, presented in this thesis is novel understanding that the time-to-target is an important control input in visuomotor control, and more generally – optimal feedback control. In addition to this fundamental result, here I also present a new computational model that allows for the accurate modelling of reaching movements, as well as the formulation of hypotheses, which helped to revise some long standing phenomena (such as the temporal evolution of feedback intensities). Finally, this work also contributes towards the accumulating evidence of the relationship between feedforward and feedback control, specifically that the two systems are likely not independent.

Contents

"With great power comes high feedback gains"

-Unknown

"You know what's cooler than magic? Math!"

-Peter Parker

1 Introduction

Human sensorimotor system has to solve a range of different problems in order to produce skilful movements that are usually taken for granted. For example, throughout the course of PhD, a typical doctoral student will take about 70000 sips of coffee by reaching to a mug, grasping it, lifting and bringing it towards the lips, and putting it back down. However, for every such action to be successful, the sensorimotor system needs to perform numerous smaller tasks, such as plan the movement, send the feedforward motor command and then correct for errors using feedback information. In addition, these tasks are further complicated by an infinite number of joint orientations and muscle activations that are possible when reaching for the cup, change in motor commands required as the cup gets lighter over time, or a possible knock on the arm when navigating around a crowded café. Finally, all this is executed amid uncertainty due to sensory noise, motor noise or even delays within feedback systems, yet spectacularly not one of the 70000 sips missed the target.

While different aspects of human movements, such as planning, feedforward control, feedback control or others can be studied independently, current research increasingly shows that these modalities are not independent from each other. For example, recent neurophysiological studies now support parallel processing of action selection, and sensorimotor control [1, 2], thus suggesting that motor planning and motor control are taking place simultaneously and interactively, as opposed to serially one after the other. In addition, behavioural research examining relations between feedforward and feedback control have demonstrated that adaptation of feedforward control affects the performance of feedback control [3–7]. Thus, even though the research, presented in this thesis primarily focuses on the aspects of feedback control during the goal-directed reaching movements of humans, implications of the results are also discussed more holistically.

1.1 Feedforward and feedback control

Feedforward and feedback control are continuously present in motor systems of humans, acting together to produce effortless and coordinated movements. The main difference between the two control schemes, is that feedback controller generates movements by utilising live feedback, such as the sensory information, while feedforward controller does not. Instead, the feedforward controller relies on understanding the internal dynamics of the controlled system and produces a signal that will achieve the intended output [8]. For example, in order to shoot the basketball well, a player has to instinctively understand the dynamics of his own body to know that a particular pattern of muscle activation will bring the ball to a set point, and then will launch the shot. In addition, the external dynamics such as gravity and the weight of the ball also have to be accounted for in order to shoot the ball precisely to the basket. However, after the ball has left the contact with the player's hand, the motor action is over and the shot can not be influenced anymore by any feedback signal. While adjustments will be made based on the outcome of the shot (hit or miss), they will update the internal models of the body (e.g. in case of fatigue) or the world (e.g. if the wind is consistently blowing from a single direction), but not the outcome of the current shot.

On the other hand, feedback control typically utilises the sensory information throughout the duration of the movement. In a similar basketball example, a defending player will attempt to rebound a missed shot by catching the ball after it bounces from the basket. Here the player will first observe that the ball did not hit the basket, triggering a motion towards it. In addition, as the ball falls to the ground, the player will continuously track its location while simultaneously reaching towards it. Even if the ball is further displaced by teammates or opponents, the player will be able to track it down, as he can continuously produce an appropriate motor action that eventually reduces the distance between him and the ball.

While the two control systems can be isolated in these examples to emphasise their contributions, in human motor control they are constantly interacting. First, the attacking player can still utilize haptic feedback when launching the shot and slightly tweak his feedforward command if he did not catch the ball precisely. Second, a great rebounder will use the feedforward control to get into the perfect position even before the ball hits the rim of the basket. Such interactions are critical for a successful control system, as the two controllers have their individual strengths and limitations. For example, the feedback controller can always produce an appropriate action towards the goal, however it is limited by the delays that are present in the feedback loop due to signal propagation and cortical processing. While this is not always an issue, particularly when the goal of the movement is stationary, even physiological delays in human motor system have been shown to significantly increase

the uncertainly of the feedback signal [9], while slightly longer delays may render balancing tasks completely impossible [10]. In contrast, the feedforward commands are executed in anticipation of the system dynamics, and are therefore produced without any delay [8]. In case of a systematic perturbations, e.g. gravity, the feedforward system produces motor commands that directly accounts for such stimulus. However, without feedback the feedforward controller can not account for any unforeseen disturbances such as a gust of wind or a knock on the arm, and would produce an erroneous action. As a result, while in our work we primarily focussed on feedback control, the mutual relationship between two systems also allows us to discuss the mechanisms of the feedforward control.

1.2 Visuomotor control

Central to this doctoral thesis and different studies presented here is the feedback system of visuomotor control. In broad terms, such control happens when motor output is produced as a result of some visual stimulus. This chapter presents the background about such control, with particular emphasis to the phenomenon of the visuomotor feedback response.

1.2.1 Rapid onset of visuomotor feedback responses

When producing hand movements from one point to another, for example in reaching, humans make use of visual information of the environment, as well as the hand. Specifically, in addition to observing where the reach goal is positioned in relation to the hand when planning the movement, visual information is used throughout the reaching action in a feedback process that guides the movement amid various sources of noise [11, 12]. While early studies entertained a hypothesis that these visual feedback processes are only recruited near the end of the movement in order to adjust the accuracy [13], it is now well known that the use of vision is continuous, and produces corrections to various visual perturbations as rapidly as within 150 ms after the perturbation [11, 12, 14]. In addition, such corrections are largely automatic [15] and cannot be entirely suppressed [14], even if they actively impede the performance of the task [14, 16], leading to discussions whether or not these responses should be characterised as reflexes [17].

Numerous studies have now investigated properties of visuomotor feedback responses via visuomotor perturbation tasks, where a reaching effector (e.g. hand) or a target is perturbed by a spacial shift or a jump during an ongoing movement. A typical visuomotor feedback response is produced as a response, and in a direction to correct for these perturbations. Even when instructed to move in the opposite direction, that would effectively further increase the error, participants initially produce a corrective response, only later reversing the movement direction to complete the instruction [14, 16].

As a result, such paradigms allowed for measuring of the voluntary timing of visuomotor control, with earliest conventionally accepted voluntary responses produced no earlier as 230 ms after the perturbation onset [16]. Note, that this voluntary timing is much later than the initial responses recorded at 150 ms, leading to conclusions that visuomotor responses between 150 ms - 230 ms are entirely involuntary. In turn, some recent studies now consider these responses in two separate stages: early, with onsets before 230 ms, and late, with onsets after 230 ms [5, 18–20].

1.2.2 Modulation of visuomotor feedback responses

The key feature that makes visuomotor feedback responses fascinating, is that in addition to very rapid onsets, these responses also manifest complex modulation. Early work by [14] demonstrated that with practice human participants reduced their response intensities, when instructed to produce movements that oppose the natural direction of visuomotor corrections. Similar findings by [16] also demonstrated that such responses get reduced if the perturbations are task-irrelevant (e.g. if they "self-correct" after a short duration), hence showing the capability of visuomotor responses to gradually adapt to the environment, similar to what is known for feedforward controllers. However, the visuomotor feedback responses are also regulated by the properties of the environment in another, much faster way, where different intensity responses could be produced in two consecutive movements if the perturbations occurred at different timings [21], positions [22], or if otherwise identical movements were performed to targets of differing width [23].

Two similar studies of visuomotor feedback responses demonstrated spectacular variability of these responses by perturbing a target [21] or a cursor, representing hand location [22] at various different onsets throughout the movements. While the requirements for forward motion were kept fixed within a given condition – reaching to the same target with the same kinematic requirements – the perturbations were spontaneously cued at one of many possible onset locations, spaced throughout the movement distance. Interestingly, both studies demonstrated that even on timescales of a single trial, the visuomotor feedback system can produce vastly different responses that are modulated by an inverse of time to movement end after the perturbation [21], or in a non-monotonic fashion related to the perturbation onset location [22]. While the two control schemes seem to contradict each other on a surface level (but can be integrated together in light of the results of this thesis), they both demonstrate elaborate variability of visuomotor feedback control, raising further questions about how such responses can be both so flexible and so fast.

Further studies have analysed properties of visuomotor feedback responses. Target uncertainty seems to modulate the visuomotor feedback responses, with higher uncertainty reducing the intensities [24]. In a similar fashion, visual gaze location directly at the target produces stronger

responses than in otherwise identical movements where the visual fixation is away from the target [25, 26]. Moreover, while background force-load does not seem to affect the visuomotor feedback responses, adapting to a viscous force field increased the overall visuomotor feedback responses for stronger force fields [5].

1.2.3 Effectors for visuomotor control

Various research has demonstrated that visuomotor feedback responses can be invoked by perturbing different effectors, such as the cursor (i.e. the visual representation of the hand, [12, 16, 18, 22, 23, 27–30]), the target [11, 14, 18, 21, 28, 30–33], external tools (e.g. computer mouse [34]), or the background motion [33, 35]. Our own work, presented in this thesis (Studies I and III), as well as the general overview of all the aforementioned studies suggest that all these feedback responses behave qualitatively similarly, independent of the actual effector. For example, [18] contained both, perturbations of only the target, as well as the perturbations of only the cursor, the regulation of which were qualitatively similar. Moreover, while not a direct comparison, but the regulation of responses to cursor jumps in our Study I and target jumps in Study III both matched very well with mixed-horizon OFC predictions in Study II, which is in general agnostic to the effector. However, despite these qualitative similarities, it is now well known that visuomotor responses to cursor and target perturbations are processed at least partially via separate feedback loops [18, 28, 30]. Specifically, a study where both responses to cursor and to target jumps were compared [28], observed that responses to cursor jumps were generally weaker, likely because additional sensory information is available about the hand position through proprioception, while vision is the only source in estimating the target location. Further research demonstrated the existence of a dedicated visuomotor binding mechanism, that links the visual position of the hand with its internal motor representation, making visuomotor feedback control of the hand much less susceptible to environmental distractors, unlike in the case of target perturbations [30]. Finally, [18] proposed and experimentally tested a model for the integration of target and cursor information in visuomotor control, showing that only the integration of the two separate pathways that are combined in a later visuomotor processing stage can explain the observed human behaviours.

1.2.4 Optimality of visuomotor feedback responses

Visuomotor feedback responses, and goal directed reaching movements in general have been shown to demonstrate some behaviours that are optimal. Various early models that assume optimality, such as minimum jerk [36], minimum torque-change [37] or minimum end-point variance [38] have all been able to mimic human-like movement trajectories or velocity profiles with increasing success, suggesting that similar computational principles could be happening in humans. However,

many of these models inherently implied that the control is open-loop, and that the entire movement trajectory is pre-planned, limiting the ability to correct for any deviation from the set path. In contrast, the mere presence of visuomotor feedback responses suggests that feedback components are critical to overall control of trajectory. Moreover, many studies have now demonstrated regulation of these responses that can be considered optimal, such as reduced corrections to wide targets [23] (which follows the minimal intervention principle [39]), trajectory variability in via-point moving task (i.e. utilising the redundancies of the biomechanical system [39]), weaker responses to task-irrelevant disturbances [16], stronger responses for more "urgent" perturbations [21], and again reduced responses when corrections are "too late" to be made [22, 31]. Finally, similar movement control has been observed in other feedback systems, for example when responding to mechanical perturbations [17, 40–44], suggesting that similar algorithms are used among multiple control systems in humans. As a result, optimal feedback control has recently taken the forefront as the primary theory for the control of various movements, but particularly in the visuomotor system.

1.2.5 Visuomotor control in balance

Another contribution of visuomotor control in humans could be to stabilise unstable objects. For example, a child trying to balance a broomstick vertically on her fingers would inevitably perform better with her eyes open than closed. Such system, also known as controlling an inverted pendulum above the cart, could be stabilised either through increased stiffness (i.e. grasping the bottom of the broomstick and not allowing it to tip), or dynamically, by always correcting for the support to be under the centre of mass. While proprioception and tactile feedback play a significant role in such control, in order to achieve dynamic stability, visuomotor feedback is invaluable, as the observed visual errors can be rapidly corrected by adjusting the support. However, it is still unclear how exactly the visual feedback is used on the computational level to produce the responses in human motor system. Therefore, we have developed an experimental setup where we can distort or completely remove different feedback modalities, feedback delays or dynamics of the system in order to test the importance of visuomotor feedback in dynamic stability (Study IV). Furthermore, we tested multiple control models in order to gain better insights into the computations happening in such system (Studies IV-V, [45]).

1.3 Optimal control

1.3.1 Definition of optimal control

While much research in human motor control operates under the premise that humans have evolved to produce movements that are optimal in some aspect, it is still widely debated what precisely “optimal” means for the human motor system. As briefly discussed in an earlier chapter, early

control models simply assumed that humans aim to produce smooth movements by minimising joint torque-change [37], hand acceleration or hand jerk (the first derivative of hand acceleration, [36]). Later optimal control models also considered minimising the effect of the noise [46], reducing end-point variability [38] or minimising the effort while simultaneously achieving the reward of the goal [47–49]. Mathematically all these models are formulated through defining a "cost" (i.e. total noise over the movement, end point variability, effort) that a particular action will cost to perform and aim to minimise it. Therefore, an optimal model is not necessarily the "best" solution in a colloquial sense, but rather the one that produces an action with the minimum cost for a given cost function.

1.3.2 Optimal feedback control frameworks

A particular subset of optimal control models are the optimal feedback controllers (OFC). OFC is particularly prominent in this thesis, as the variability of visuomotor feedback responses (and also other feedback responses such as stretch or oculomotor) have conventionally been explained using this framework. Multiple phenomena in human motor control have been explained using OFC, particularly the minimum intervention principle, redundancy, movement variability patterns (all summarised in [39])) or undershooting when moving to the target [31]. However, the implementation of the vast majority of these models is based on a linear-quadratic regulators (LQR) or linear-quadratic Gaussians (LQG) – two optimal feedback controllers with convoluted cost functions [50, 51]. Typically, when applied to goal-directed reaching movements, such cost functions consider information about position, velocity and acceleration of the hand, as well as positions of target and any obstacles, and thus require seemingly heavy computations, raising a question how this process is executed so fast by humans.

To explain this, recent work suggested that feedback control gains could be precomputed in the movement planning stage and recalled if needed [52], however that would require storing an infinite number of discrete response mappings. In the first publication of this thesis we instead propose that a characteristic function, unique for the movement goal and independent of the movement kinematics could directly map the time left in the movement (time-to-target) to the optimal feedback intensities (gains). Specifically, through our experimental design we uncoupled the natural relationship between movement kinematics and the time-to-target, showing that the time-to-target, and not any of the kinematic variables are the independent variable that controls the responses.

Another distinct general framework that is relevant to the studies in this thesis is the proportional-integral-derivative (PID) control, or its special case proportional-derivative (PD) control, where any corrections related to the integral component are ignored. Particularly, PD and PID controllers aim to reduce the error between the target point in the system, and the current state, and have previously

been shown to well capture human behaviour in upright posture balancing [53, 54]. In essence, such controllers aim to minimise the sway of posture (or any other balanced object in a more general case) by correcting the error (e.g. the tilt angle of the body) by a fixed "proportion", as well as issuing additional corrections based on how fast the error is changing (e.g. proportional correction based on the angular velocity, or the derivative of error), and on the history (integral) of these errors. While our work has shown that LQG can very well explain the human behaviour when balancing an inverted pendulum on the cart [45], in studies IV and V we also test the performance of PD controllers for such balancing tasks.

1.4 Normative modelling

In addition to experimental work, we also relied on computational modelling to explain our results or to motivate hypotheses for these and future studies. One standard approach in such modelling is to find a best-fit model to a given dataset by minimising the error between the data and the model predictions. Such models by definition well describe the available data, however may not always result in a clear underlying mechanism of why the data behaves in such way. Instead, in our work we primarily adopted a normative modelling approach, where we built a computational model from the bottom-up and evaluated its behaviour in comparison to experimental results [55]. Such approach is useful for multiple reasons, and has therefore been used in multiple studies. First, by building a model (or in our case a control system) we know exactly the computational principles involved in producing its behaviour. Consequently, if this behaviour matches that observed in the experimental results, a parallel can be drawn between the known computational mechanism, and the unknown mechanism governing the human sensorimotor control. On the other hand, in case the behaviour between the model and the data does not match, additional or different mechanisms may be recruited in real-life behaviour. Second, such models can raise clear and testable hypotheses for new experimental studies, particularly if multiple alternative models can be compared. While we did demonstrate the influence of the time-to-target using the model simulations in our first study, such modelling was key to our second and third publications. Here, we could not only revisit the results of previous studies in context of model predictions that suggest alternative explanations, but also clearly formulate a new hypothesis for an experimental study.

2 Methods

2.1 Experimental setup

All studies presented in this thesis rely on the experimental results in some capacity. Particularly, for every study with the exception of [56] a separate experimental dataset was collected. For [56] we instead used previously collected datasets that followed similar experimental procedures as described in this chapter.

2.1.1 Participants

Participants in all presented studies were aged 18-45 years, had no known neurological diseases and were of normal, or corrected to normal vision. In addition, all participants self-reported as being right-handed according to the Edinburgh Handedness Questionnaire [57]. Before each study participants signed an informed consent to the voluntary nature of the study, and were compensated with 8 euros/h for their time. All studies and informed consent formulations were approved by the Ethics Committee of the Medical Faculty of the Technical University of Munich.

2.1.2 vBot

All experimental data presented in this thesis was collected via the vBot robotic manipullandum [58]. Participants were seated in an adjustable chair and restrained using a four-point harness in order to limit the movement of their shoulder. A six-axis force transducer (ATI Nano 25; ATI Industrial Automation) measured the end-point forces applied by the participant on the handle. Position and force data were sampled at 1 kHz, while velocity information was obtained by differentiating the position over time. Visual feedback was provided via a computer monitor and mirror system, such that this system prevented direct vision of the hand and arm, and the virtual workspace appeared in the horizontal plane of the hand (Figure 1).

2.2 Visuomotor feedback response studies

2.2.1 Experimental paradigm

In visuomotor feedback studies participants were required to perform goal directed reaching movements from a given start position to a target. Participants controlled a cursor (a small circle) by moving the robotic handle of vBot. A standard procedure would require participants to bring this

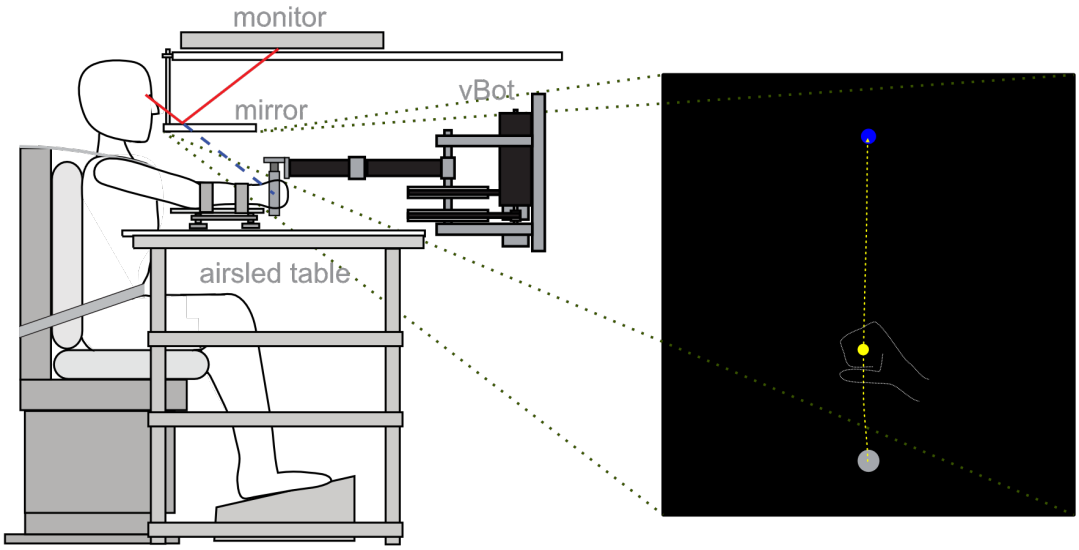


Figure 1: Schematic of vBot robotic system. Participants were seated with their arm supported on an airsled and grasping the handle of the robotic manipulandum. The direct feedback of the hand and arm was blocked by a mirror, which projected the computer screen to the plane of the hand. In this virtual workspace participants saw the hand represented as a small circular cursor (here presented in yellow), that was moved from a start position (grey circle) to the target (here presented as a blue circle).

cursor to a start position (typically a grey circle of a larger diameter than the cursor), then wait for a go cue and reach to the target. After each trial, feedback about the kinematics was presented to participants, indicating their speed in relation to the task requirements, as well as the temporal evolution of the kinematics. If participants produced a movement that fit the requirements of a successful trial, they were awarded with an increase in the total score to motivate their performance. During these reaching movements perturbations to the cursor or to the target could be applied by suddenly shifting them by a small distance, perpendicularly to the line joining the target and the start position (Figure 2A). This perturbation typically evoked sizeable responses from the participants, who produced the corrections in the task-relevant direction, and with onset times in sub-voluntary range (150 ms) (Figure 2B). Perturbation types and onset criteria varied across different studies, with specifics presented in the Methods of a given study.

2.2.2 Channel trials and maintained perturbations

In order to evaluate the intensity of feedback responses we measured the forces that our participants produced as a result of the perturbation. To do so, we utilised virtual mechanical channels [59] of stiffness K = 2 N/m and damping D = 4000 Ns/m, that were generated by the vBot. In these trials participants were constrained by lateral forces to move along a single line, connecting the start position and the target. Any attempt to move perpendicular to the required movement direction was met with the counteracting force, effectively constraining the movement to this channel. As a result,

any corrective force, produced by the participants could be recorded with high accuracy by the force transducer, mounted on the robotic handle.

While these channel trials allow us to accurately record participant responses, they also pose certain challenges. First, as any corrective movement after the perturbation is resisted by the channel, any perturbations in these trials are uncorrectable, and thus the movement can never converge to the target. As a result, in the channel trials we induced perturbations that are task irrelevant. Specifically, these perturbations are only maintained for a short duration of 250 ms, before they are removed, and the cursor or the target come back to their original path or location. In this way, the movements can be successfully completed and the responses can be measured. However, as these perturbations are task irrelevant, the most efficient behaviour is to completely ignore the perturbation, thus producing no responses. Indeed, when presented with only channel trials and task irrelevant perturbations participants significantly reduce their response magnitude, increasing signal-to-noise ratio and reducing the power of the study. One way to eliminate the decay of feedback responses due to the channel trials is by interspersing these trials with maintained perturbations. Here the movements are performed without any constraints, but the perturbations are maintained beyond the 250 ms duration, and are therefore task-relevant. To successfully complete these trials participants have to actively correct for the perturbations in order to bring the cursor to the target. The reason why such trials are effective in keeping the responses upregulated, is that the movement pre-perturbation is consciously indistinguishable between maintained perturbation trials and channel trials. As a result, participants are forced to produce the corrective responses, otherwise risking to fail the trial. However, as there is no force channel to push against, these trials provide less sensitive feedback data, such as the change in trajectory over time, which then needs to be differentiated twice to be even remotely comparable with force channel data. Hence, we typically do not use the maintained perturbation data for the analysis, but rather use it to facilitate stronger feedback responses.

Another reason why maintained perturbations can be tricky to use in a given design is that the perturbed movements in channel trials take shorter time to complete than the maintained perturbation trials. This is due to additional path that maintained perturbation trials take, often only left with the corrective motion after the hand is already at the target distance. In contrast, channel trials immediately end once the target distance is reached, as no lateral error is present by that time. While these effects were not considered in traditional studies, our work presented in this thesis suggests, that movement duration (more precisely the duration from the perturbation onset to the movement end, or time-to-target) may be an important control variable in regulating the feedback responses. Consequently, using maintained perturbations together with channel trials could affect

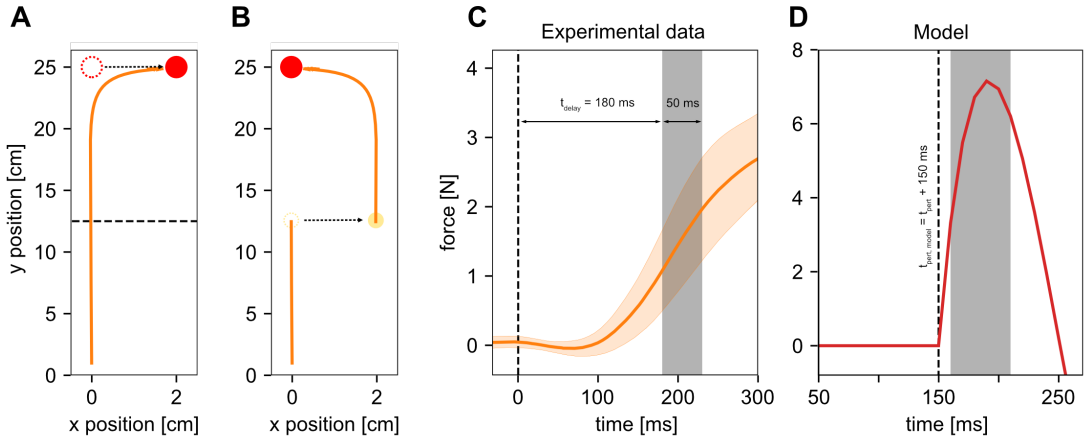


Figure 2: Schematic of visual perturbations and their induced feedback. **A.** Target perturbations. As participants moved their hand to the target, target perturbation was induced by shifting the target representation (here a red circle) laterally to the direction of movement. As a result, participants produced corrective movements to bring their hand to the target. Dashed horizontal line indicates the perturbation onset location. **B.** Cursor perturbations. Similarly to target perturbations, the cursor perturbations were induced by shifting the cursor in a direction, perpendicular to the movement. As a result, participants produced a corrective movement in the opposite direction. Red circle represents the target and yellow circle represents the cursor. **C.** Example feedback responses to target or cursor jumps, recorded experimentally. Zero-time indicates the perturbation onset. Force responses (orange line) are produced to correct for the perturbations, with onset times 150 ms. These responses, averaged over 180-230 ms window (grey area) constitute visuomotor feedback intensities. Shaded error represents 95% CI. **D.** Modelled feedback responses to target or cursor jumps. Simulated force (red line) is artificially delayed by a fixed delay (here 150 ms), representing the visuomotor delay in humans. After that, the corrective response is produced rapidly and immediately due to the modelled muscle being implemented as a first-order filter. Due to rapid onsets, the simulated feedback intensity time window (grey rectangle) is shifted towards earlier times than in humans (160-210 ms).

the response regulation in a non-trivial way, as participants would not be able to reliably estimate the time-to-target. As a result, in order to avoid uncontrolled effects that maintained perturbations could possibly induce, our most recent study [60] did not anymore contain maintained perturbations.

2.2.3 Visuomotor feedback intensity

In studies, investigating visuomotor feedback responses, it is necessary to first quantify the strength of the motor response to the visual perturbations in order to then study their modulation. To do so, a force response, produced as a result of visual perturbation was measured and averaged over a time window between 180 ms and 230 ms after the perturbation, to produce one value, usually termed visuomotor feedback gain [16, 18, 22, 25, 26, 43] (Figure 2B). While conventionally earlier studies call these force responses as visuomotor gains, in our work we refer to them as visuomotor intensities. The main reason for such choice is to avoid confusion with other gains that are present

in our computational models, specifically Kalman (observer) gains K or feedback (control) gains L.

The two extremes of the time window are selected between the typical force onset at about 150 ms after the perturbation [11], and earliest detectable voluntary produced force at 230 ms after the perturbation [16]. As a result, these gains represent early, sub-voluntary visuomotor feedback responses. Note that some studies [18, 19] also analyse the responses during later time windows, where voluntary corrections are also contributing to the total response, however in our work we primarily focussed on the early responses.

In a similar fashion, when evaluating the control performance of computational models we also estimated their simulated feedback intensities, equivalent to visuomotor feedback intensities. In order to mimic the visuomotor delay within the model, we induced the simulated perturbations 150 ms later than they would normally be induced for humans. Moreover, due to the difference in muscle properties between human muscle and simple first order filter, the responses to simulated perturbations typically ramped up much faster, requiring the further adjustment of the visuomotor feedback window in model evaluations. Here we used a time window between 10-60 ms after the perturbation onset, equivalent to 160-210 ms in humans (Figure 3C). Further details of the model implementation are presented in the following sections.

2.3 OFC models

In the three main studies, presented in this thesis, we relied on computational modelling to interpret the experimental results or to motivate new hypotheses. Our primary tools in simulating a human-like control behaviour were optimal feedback control (OFC) models. This section details the general implementation of different types of OFC models, as well as their features. However, as various models, that are presented in the studies of this thesis are implemented with different numerical values of parameters, for the exact implementations the reader should refer to the studies directly.

2.3.1 Optimality

In order to be precise when interpreting the results of our studies it is important to first define what we mean by optimality and optimal control in general. As an example, let's consider again reaching towards a cup of coffee that is placed out of our reach. First, in order to get to this cup we need to spend our energy to move our muscles – stand up, walk towards the cup, extend our arm and pick the cup up. Here we have unlimited choices in how we move, from not moving at all to jumping up and trying to get to the cup as fast as we physically can. Typically, faster, more vigorous movements will cost us more energy as well as increase motor noise, leading towards additional corrections to compensate for this noise. Second, depending on how much we want that particular cup of coffee

at this given moment, we incur an intrinsic “penalty” for not already having that cup in our hands with every passing second. These two factors, sometimes called movement cost and state cost are contributing to the total cost of the movement, that optimal control aims to minimise. As a result, optimality is defined in the context of the cost function for that movement, and is simply a solution for a given cost function. In turn, an optimal movement with one cost function will most likely be non-optimal, if instead another cost function were used. Hence, when we talk about optimal control, we simply solve for the most efficient control signals that drive movements with a given cost function, however the cost function itself is a model design problem, which might not always be the most appropriate for the given task.

2.3.2 State-space representation

In order to simulate a controller we first need to describe the system that we aim to control, as well as its properties. Here we defined the hand as a point of mass m and intrinsic damping b, to simulate viscoelastic behaviour. This point mass is controlled in two dimensions by two orthogonal force actuators that simulate muscles, and are regulated by a control signal u via a first-order filter with time-delay τ. At time t within the movement, such system could be described by a set of simultaneous equations:

$$dp = v \tag{1}$$

$$mdv = f - bv, \tag{2}$$

$$\tau df = u - f \tag{3}$$

The above equations in continuous time can be expressed as a matrix equation (also known as state-space representation):

$$\dot{x} = A_c x + B_c u \tag{4}$$

In order to simulate the system numerically, we can discretise the system using Euler’s approximation, and include random processes that represent noise:

$$x_{t+1} = Ax_t + B(I + C_t)u_t + \xi_t. \tag{5}$$

Here x_t is a state vector, containing variables of position p, velocity v, muscle force f and an augmented target position $p*$. u_t is an external control force that is driving the system. A and B are state-transition and control matrices respectively, and C_t is a 2x2 matrix, whose each element is sampled from a zero-mean Normal distribution, representing control dependent noise. As in all of our studies we aimed to simulate only the mean behaviour of our participants, we typically kept control-independent noise ξ_t to zero in our simulations.

2.3.3 Linear-quadratic regulator (LQR)

After we define the system dynamics, next step in simulating its behaviour is to define how such system is controlled. Specifically, at any given state (a set of values like position, velocity and acceleration) there is an infinite number of inputs that will propel such system. However, only a small subset of all these inputs will bring the system towards our desired goal (for example the target of reach), and even fewer will do that optimally (i.e. by minimising the movement cost). One way of defining a cost of movement and then finding its optimal solution is through a linear-quadratic regulator (LQR, [61]). LQR assumes linear state-space dynamics (as satisfied in Eq.5), as well a quadratic movement cost with respect to state variables:

$$J = \sum_{t=0}^{N} x_t^T Q_t x_t + u_t^T R_t u_t. \tag{6}$$

Here Q and R are known as state-dependent, and activation-dependent costs. As our state variable x includes states of position p, velocity v and force f, as well as information about target position $p*$, the above equation at time t can also be expressed as:

$$J = \sum_{t=0}^{N} \omega_{p,t}(\vec{p_t} - \vec{p^*})^2 + \omega_{v,t}||\vec{v}_t||^2 + \omega_{f,t}||\vec{f}_t||^2 + \omega_{r,t}||u_t||^2, \tag{7}$$

with $\omega_{p,t}$, $\omega_{v,t}$, $\omega_{f,t}$ and $\omega_{r,t}$ being numerical costs of position, velocity, force and motor input (activation). Finally, N is the time horizon of the movement (see Section 2.3.5 for further details),

The solution that minimises the cost function J is well known, and can be achieved by solving an algebraic Riccati equation. The resulting control matrix L then defines an optimal feedback control signal for any given state

$$u_t = -Lx_t. \tag{8}$$

2.3.4 Linear-quadratic Gaussian (LQG)

LQR implementation, presented in the previous section is useful when we are only concerned with bringing the system from one state to another. However, human motor system generally does not have exact information about its state, but infers it using noisy sensory inputs. First, this requires an optimum estimator, known as Kalman filter, to estimate the current state from sensory information and the internal knowledge of system dynamics. Second, we need to generate an optimal control solution, accounting for the fact that the state information x_t is not exact, but estimated. Such a problem, where optimal estimator (Kalman filter) and an optimal controller (LQR) are combined into one is known as linear-quadratic Gaussian (LQG).

In addition to the state transition equation (Eqn. 5), for the LQG we also define a state observer

$$y_t = Hx_t + \epsilon_t + D_t x_t. \tag{9}$$

Here y_t is sensory information, observed from the real state x_t via observation matrix H, and corrupted by stationary noise ϵ_t and state-dependent noise D_t.

In cases where multiplicative noise does not exist ($C_t = 0$ and $D_t = 0$ throughout the movement), and the system is only disturbed by additive noise, the optimum solutions for Kalman gain K and control gain L are separable. That is, both gains can be found independently, and then combined to produce the LQG [62]. In turn, Kalman gain and control gain can then be used to estimate the current state, and to produce the control signal:

$$\hat{x_{t+1}} = A\hat{x}_t + Bu_t + K_t(y_t - H\hat{x}_t) \tag{10}$$

$$u_t = -L\hat{x}_t. \tag{11}$$

When multiplicative noise is present, the two gains become inseparable, and have to be optimised iteratively, until the solution converges. An algorithm to achieve this has been proposed and demonstrated by [63, 64].

2.3.5 Control horizon

Another distinction between various optimal control models (both LQG and LQR) is by their control horizon. The three types of control horizons that are utilised in our studies are infinite [56, 65, 66], finite [31, 43, 60, 63, 67] and receding [56, 68]. Each of the three types have slight differences in implementation, which in turn leads to different predicted behaviours, uses and limitations.

Infinite horizon control assumes that the control action is infinite, and thus produces a control signal indefinitely, as long as the current state and target state are different. Such a definition implies that the time horizon N is infinite, and state and activation costs Q and R are stationary over time (Eqn 6). In turn, this results in a control gain L that is stationary as well.

Due to the infinite horizon, movements generated by such a controller always converge to their target, as the control gain L will always drive the system towards the goal. As a result, such models typically produce human-like kinematics, with bell-shaped velocity profiles, convergent trajectories, and, importantly, can well capture the variability of movement duration in presence of disturbances [56, 67]. On the other hand, due to the stationarity of the control gain L, infinite horizon control always produces a consistent control signal if the state error is the same. This means that a lateral target jump of the fixed size will always invoke identical control responses in the infinite horizon, which has been showed in numerous studies to not be true in human behaviour [16, 18, 22, 67].

In contrast to infinite horizon control, finite horizon OFC is described not only by costs Q and R, but also by a time horizon N, which is finite. In addition, such architecture allows the costs to also be

time-dependent, resulting in a different simulated behaviours compared to infinite-horizon control. First, for finite horizon control movement end is pre-defined and ends abruptly even if the target is not reached. Liu and Todorov [31] have demonstrated, that such control well describes kinematics of perturbed human movement in conditions where getting into the target is not compulsory. Specifically, such movements stereotypically undershoot and stop without reaching the target, both in OFC simulation as well as in human data. Importantly, finite horizon control also produces variable feedback gains, even if the state error is the same at two separate time points, which predicts temporal evolution of feedback intensities, as demonstrated in human studies [16, 18, 22, 60, 67]. However, finite horizon control is limited by the need of knowing the movement duration in advance, as otherwise the solution can not be computed.

Receding horizon control combines features of finite and infinite horizon. Mathematically the receding horizon OFC is implemented in the same way as the finite horizon control, however it does not assume the time horizon of the whole movement but only of the portion of it. Particularly, in finite horizon control, the cost function is computed over the whole duration from $t = 0$ to $t = N$. In receding horizon control, the cost function is only computed over the duration from time t to time $t + N_{rec}$, where $N_{rec} \ll N$ is the fixed length of the horizon that the movement is accounting for. As the time t advances, so does the planned end-time of the movement, so producing the control that always converges to the target.

The receding horizon models of human movement assume that humans do not plan movements from start to finish, but rather in stints through via-points. An analogy of receding horizon control could be following a GPS route between two locations: each portion of the movement is only planned until the next upcoming turn, even if the end goal is known before the journey. In spirit with this analogy, receding horizon control well encapsulates some human behaviours that are present in long, slow movements, for example oscillatory non-bell-shaped velocity profiles [68]. However, as with infinite horizon control, receding horizon also fails to capture the temporal evolution of feedback gains which are known to be present in humans.

2.4 Pendulum studies

In addition to studies where we analysed human behaviour in short, ballistic perturbed movements, we also studied the feedback control systems in situations of extended control. In this section I present the methods of studies where participants balanced a virtual inverted pendulum on a cart.

2.4.1 Experimental paradigm

Participants were seated in a vBot robotic manipulandum and grasped the robotic handle to control a cart (a red 3.0 cm x 1.5 cm rectangle) of mass M = 0.1 kg, which could be moved within the x-y plane. The pendulum was simulated as a massless pole of variable length L, extending from the middle of the cart in the positive y-direction (away from participant). At the tip of this pole, a point mass with m = 1 kg was attached, and the gravity with g = 9.8 m/s^2 was simulated in the negative y-direction. Such system was marginally stable, if and only if the cart was directly below the point mass. Otherwise, the system was pushed away from the equilibrium, with its dynamics described by a set of equations:

$$F_x = \ddot{x}(m\sin^2\theta + M) - mL\dot{\theta}^2\sin\theta + mg\sin\theta\cos\theta \tag{12}$$

$$\ddot{\theta} = (g\sin\theta - \ddot{x}\cos\theta)/L \tag{13}$$

where F_x is a lateral force applied by the pendulum on the cart, θ is the angle between the pendulum and the y-axis, and x is the position of the cart.

While the cart could be moved in all directions in the x-y plane, its movement in y-axis was constrained by a virtual mechanical channel with stiffness of 4000 N/m, damping of 2 Nm/s, and maximum applied force capped at 25 N, to encourage movements primarily along the x-axis. This channel was also framed by two thin yellow lines (Figure 3). In order to maximise the range of motion, the cart position did not exactly match the position of the hand, but was shifted in the negative y-direction (towards the participant) by 13.0 cm. The x-coordinate was always matched between the cart and the hand. In most studies, any force F_x, produced by the dynamics of the pendulum was applied to the robotic handle to provide haptic feedback, but was capped at 5 N and completely turned off beyond the point of recovery (θ>30^o) for safety reasons. However, in studies where this force interfered with desired measurements this feedback was suppressed.

The pendulum itself was displayed as a line of 3 mm thickness, extending between the cart and the point-mass for its actual length L. As this length was almost always longer than the screen height, this line was truncated at the top. To provide the information about the pendulum beyond screen size, a thin separated strip at the top of the screen contained a circle (d = 1.0 cm) that indicated precisely the actual lateral location of the tip of the pendulum. Due to the pendulum implementation as a point mass on a massless pole, this information also coincided with the centre of mass of the pendulum, unless otherwise specified in the design of a particular study.

Typically, the trials were self-paced in pendulum experiments, with each trial starting after participant brought the cart to the designated starting position. After that, the pendulum would be pushed from

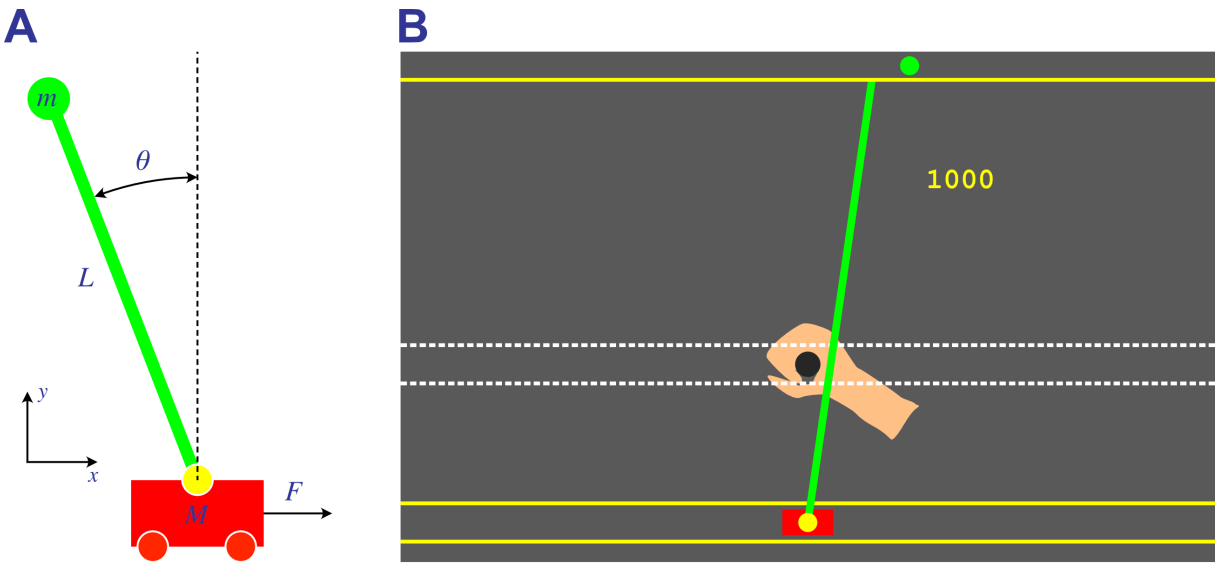


Figure 3: A. Schematic of the pendulum-above-cart system. The equilibrium of the pendulum (green) is achieved through moving the cart (red) in order to minimise the angular deviation θ. **B.** A sample screenshot of an experimental trial. The circular cursor at the top of the screen provides visual feedback of the centre of mass as the pendulum (green line) is truncated at the top of the screen. The y-coordinate of the physical hand location (not visible to subjects) is offset with respect to the cart position. The cart (red rectangle) is constrained between the two yellow rails to allow unconstrained movement along the x-axis.

its equilibrium in a random direction with angular velocity $\dot{\theta}$ = 0.01 rad/s. A trial was finished either after the pendulum was maintained upright for 5 s, or if it "fell" to the angle of $\theta = 90^o$. Participants were instructed to maintain pendulum upright and as stable as possible, and were rewarded with a score, indicating their performance, at the end of the trial.

2.4.2 PD controller

A proportional-derivative controller is a simple linear feedback controller that produces control signal based on the error and the derivative of this error. For a linear system without a time-delay (also sometimes called dead-time) a control input in continuous time is produced simply as:

$$u(t) = K_p e(t) + K_d \frac{de}{dt} \tag{14}$$

where K_p and K_d are proportional and derivative gains, set for the controller, and $e(t)$ is the feedback error signal at time t. In case of our pendulum system, the feedback error signal $e(t)$ is the pendulum angle θ, leading to the following specific form of the feedback controller:

$$u(t) = K_p \theta(t) + K_d \dot{\theta}(t) \tag{15}$$

As the control signal, produced by the PD controller scales linearly with feedback errors, such a controller is best suited for linear systems. However, in its nature the feedback control system

of the inverted pendulum is non-linear. First, even without feedback delays, the dynamics of the pendulum-cart system are non-linear (Eq. 12, 13), with pendulum falling faster and thus requiring disproportionally stronger control responses in order to stabilise it at larger angular deviations. In an ideal case, even if such system is non-linear, a PD controller is capable to produce an adequate control, as we can linearise the pendulum behaviour around the equilibrium point, resulting in close-to-linear behaviour, provided the pendulum does not deviate far from the equilibrium point. However, in addition to the inherent non-linearity of the system, the feedback loop in humans is time-delayed, with a typical visuomotor feedback delay of $\approx$150 ms. This delay, in addition to making it difficult to keep the pendulum within small deviations, also produces a non-linearity by itself, and thus needs to be accounted for. Thus, in order to simulate the PD control for our pendulum system despite all the non-linearities, we used Padè approximation, linearise our entire system and prepare it for the linear PD control.

3 Studies

3.1 Overview of the publications

The primary results, presented in this thesis, consist of three journal publications, investigating the phenomenon of the visuomotor feedback response:

Study I: **Česonis, J.,** and Franklin, DW., (2020) "Time-to-Target Simplifies Optimal Control of Visuomotor Feedback Responses", *eNeuro*, 7(2)

Study II: **Česonis, J.,** and Franklin, DW., (2021) "Mixed-horizon optimal feedback control as a model of human movement", *Neurons, Behavior, Data analysis, and Theory*

Study III: **Česonis, J.,** and Franklin, DW., (2022) "Task-dependent switching of feedback controllers", *accepted for publication at PLoS Computational Biology*

In addition to these journal publications, two additional studies resulted in two peer-reviewed conference papers, where the author of this thesis is the first author:

Study IV: **Česonis, J.,** Franklin, S., and Franklin, DW., (2018) "A Simulated Inverted Pendulum to Investigate Human Sensorimotor Control", *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 5166-5169, © 2018 IEEE

Study V: **Česonis, J.,** Leib, R., Franklin, S., and Franklin, DW., (2019) "Controller Gains of an Inverted Pendulum are Influenced by the Visual Feedback Position", *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 5068-5071, © 2019 IEEE. – also presented as a Poster presentation

Finally, in addition to the above publications, which are the foundation for this thesis, further work has been presented at international conferences and/or resulted in peer reviewed articles, where the author of this thesis was not the first author:

Česonis, J., and Franklin, DW., (2017) "Visuomotor feedback gains: optimal temporal evolution or fixed relation to movement velocity?", *Poster presentation at NCM Annual Meeting 2017, Dublin, Ireland*

Česonis, J., and Franklin, DW., (2017) "Visuomotor feedback gains: optimal temporal evolution

or fixed relation to movement velocity?", *Oral presentation at 23. Sportwissenschaftlicher Hochschultag, München, Germany*

Franklin, S., **Česonis, J.,** and Franklin, DW., (2018) "Influence of visual feedback on the sensorimotor control of an inverted pendulum", *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 5170-5173

Česonis, J., and Franklin, DW., (2019) "Retinal eccentricity modulates visuomotor feedback gains along the movement", *Poster presentation at NCM Annual Meeting 2019, Toyama, Japan*

Franklin, S., **Česonis, J.,** Leib, R., and Franklin, DW., (2019) "Feedback Delay Changes the Control of an Inverted Pendulum", *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1517-1520

Franklin, DW., **Česonis, J.,** Franklin, S., and Leib, R., (2019) "A Technique for Measuring Visuomotor Feedback Contributions to the Control of an Inverted Pendulum", *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1513-1516

Leib, R., **Česonis, J.,** Franklin, S., and Franklin, DW., (2019) "LQG framework explains performance of balancing inverted pendulum with incongruent visual feedback", *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1940-1943 – also presented as Oral presentation

3.2 Study aims

The general aim of all the work presented in this thesis is to better understand the computational mechanisms that underlie the visuomotor feedback control in humans. It has been known that short movements, simulated by OFC can achieve velocity profiles and trajectories of movements similar to those of humans [31] by producing feedback gains that vary over the movement. Similarly, other research has shown that visuomotor feedback intensities can vary over the movement with very fast response times (<150 ms) [16, 18, 19, 22, 23, 43], drawing parallels between the feedback gains and feedback intensities. However, it was still unclear whether the visuomotor feedback intensities are really an outcome of the optimal control, as opposed to other sources. For one, OFC computations are expensive – the system needs to account for various sources of the cost function, such as position, velocity, force, as well as external states – but the control signal is produced at extremely short timescales. In addition, studies that have mapped the feedback intensity profiles over the course of the movement [22] showed that these profiles well track the velocity profiles, raising the hypotheses that visuomotor intensities could simply be a simple scaling of velocity. Results of Study I show that the visuomotor feedback intensities can indeed be explained using finite-horizon OFC, and also demonstrate that only one input parameter is required for such control – time-to-target. Study II then proposes a new control architecture of mixed-horizon OFC by combining the infinite horizon

Table 1: Overview of the design in first-author studies, presented in this thesis

Study	Task	Participants	Outcome measures	Analysis tools
I	Goal directed reaching with cursor jumps	11 (5 females), right-handed, 27.3 $\pm$ 4.5 y.o.	Reach kinematics, visuomotor feedback intensities	Normative OFC modelling, BIC for model comparison, ANOVA, Bayesian factors, descriptive evaluations
II	Re-evaluation of earlier reaching studies	N/A	Descriptive evaluation of model behaviour	Normative OFC modelling, model fitting, descriptive evaluations
III	Goal directed reaching with target jumps	14 (5 females), right-handed, 21-29 y.o.	Reach kinematics, visuomotor feedback intensities	Normative OFC modelling, descriptive evaluations, ANOVA, ANCOVA, Bayesian factors.
IV	Balance of an inverted pendulum	6 (1 female), right-handed, mean age 29 y.o.	Balance score, balance time, kinematics	Descriptive analysis, PD controller modelling, model fitting.
V	Balance of an inverted pendulum	6 (1 female), right-handed, mean age 24.7 y.o.	Balance score, balance time, kinematics	ANOVA, Least-squares model comparison, descriptive evaluation.

OFC to reliably estimate the upcoming movement duration, and then using finite-horizon control to generate the movement. In Study II we also revise the results of earlier studies, showing how we can use the mixed-horizon control to uncouple control behaviours that could be masking one another. Finally, Study III directly tests the predictions of time-to-target control by testing participants in two tasks with different task requirements, strengthening the evidence for time-to-target as the primary control input. Moreover, Study III also demonstrates that humans are able to utilise the contextual cues to switch between two different feedback controllers, similar to numerous evidence of such switching in feedforward control [69–75].

Studies IV and V study visuomotor control in terms of stabilising an unstable object, rather than simply guiding short movements from one point to another. While control of an inverted pendulum

has been a classic approach in studying the processes of balance and dealing with delays, traditionally such research involves balancing a real mechanical pendulum [76–78]. Often such designs are not only inconvenient, but also limit experimenter's ability to control the experimental design, for example by suppressing or altering a particular source of feedback. In Study IV we propose a fully virtual inverted pendulum setup, simulated in vBot, which allows us to selectively remove or tweak a particular feedback modality (visual, haptic or temporal). In addition, we also demonstrate that the control of such a pendulum can be well explained by a proportional-derivative (PD) controller with time delay Δt=150 ms – roughly coinciding with a visuomotor delay, and thus consistent with a control of a mechanical pendulum. Finally, in Study V we expand on earlier work [79] that studied the balance of inverted pendulum in presence of imprecise visual feedback. The results of this study show that the human control behaviour is the most efficient when the visual feedback is directly matching the haptic feedback, and decays away from this matching point.

3.3 Study design

The Table 1 presents the design of each study, including experimental setup, tasks, participant information, outcome measures as well as analysis techniques. Note that all studies (including earlier, re-evaluated studies) were performed in a vBot robotic manipulandum.

3.4 Study I

Time-to-Target Simplifies Optimal Control of Visuomotor Feedback Responses

This study, authored by Justinas Česonis and David W. Franklin was published in *eNeuro* in March, 2020. The study analyses the temporal evolution of visuomotor feedback intensities in neurologically healthy, right-handed adults under five different movement kinematics. The methods of this study are carefully designed to dissociate the bell-shaped temporal evolution of velocity profiles during the reaching movements from the temporal evolution of visuomotor feedback intensities that also followed similar shape in earlier research [22]. The results demonstrate, that time-to-target (i.e. time left in the movement at the time of the perturbation) is a strong predictor for the visuomotor feedback intensities, produced by human participants.

Contributions

Justinas Česonis was the primary contributor and lead author in this research. In addition, Justinas Česonis implemented the experimental design, collected and analysed the data, built the computational models and wrote the first draft. Both authors designed the experimental paradigm, selected the analysis methods and revised the final version of the manuscript.

Abstract

Visuomotor feedback responses vary in intensity throughout a reach, commonly explained by optimal control. Here we show that the optimal control for a range of movements with the same goal can be simplified to a time-to-target dependent control scheme. We measure our human participants' visuomotor responses in five reaching conditions, each with different hand or cursor kinematics. Participants only produced different feedback responses when these kinematic changes resulted in different times-to-target. We complement our experimental data with a range of finite and non-finite horizon optimal feedback control models, finding that the model with time-to-target as one of the input parameters best replicates the experimental data. Overall, this suggests that time-to-target is a critical control parameter in online feedback control. Moreover, we propose that for a specific task and known dynamics, humans can instantly produce a control signal without any additional online computation allowing rapid response onset and close to optimal control.

eNeuro

Research Article: New Research

Sensory and Motor Systems

Time-to-Target Simplifies Optimal Control of Visuomotor Feedback Responses

Justinas Česonis and David W. Franklin

https://doi.org/10.1523/ENEURO.0514-19.2020

Department of Sport and Health Sciences, Technical University of Munich, 80992 Munich, Germany

Abstract

Visuomotor feedback responses vary in intensity throughout a reach, commonly explained by optimal control. Here, we show that the optimal control for a range of movements with the same goal can be simplified to a time-to-target dependent control scheme. We measure our human participants' visuomotor responses in five reaching conditions, each with different hand or cursor kinematics. Participants only produced different feedback responses when these kinematic changes resulted in different times-to-target. We complement our experimental data with a range of finite and non-finite horizon optimal feedback control (OFC) models, finding that the model with time-to-target as one of the input parameters best replicates the experimental data. Overall, this suggests that time-to-target is a critical control parameter in online feedback control. Moreover, we propose that for a specific task and known dynamics, humans can instantly produce a control signal without any additional online computation allowing rapid response onset and close to optimal control.

Key words: motor control; optimal feedback control; reaching; time-to-target; visuomotor control; visuomotor feedback response

Significance Statement

Human behavior has widely been explained using stochastic optimal feedback control (OFC), formulating movement control as a set of time-dependent feedback and control gains. However, OFC is computationally expensive leading to questions about whether such a theory could be implemented in real time. Here, we show that OFC could be approximated by a simple relationship between feedback gains and the time-to-target over a variety of movement kinematics, matching the evolution of visuomotor feedback gains of our human participants during reaching. As this relationship to time-to-target is similar across a wide range of kinematics, this suggests that early stages of the OFC controlled movement could be approximated by a time-to-target control, saving computational costs and allowing for rapid execution.

Introduction

From intercepting a basketball pass between opponents to catching a vase accidentally knocked off the shelf, visuomotor feedback responses play a familiar role in human motor behavior. Previous research has extensively analyzed these responses in human reaching movements (Day and Lyon, 2000; Saunders and Knill, 2003, 2004, 2005; Sarlegna et al., 2003; Knill et al., 2011; Reichenbach et al., 2014; de Brouwer et al., 2017, 2018), and showed an interesting combination of task-dependent variability on the timescale of a single movement (Dimitriou et al.,

Received December 9, 2019; accepted March 1, 2020; First published March 20, 2020.

The authors declare no competing financial interests.

Author contributions: J.Č. and D.W.F. designed research; J.Č. and D.W.F. performed research; D.W.F. contributed unpublished reagents/analytic tools; J.Č. analyzed data; J.Č. and D.W.F. wrote the paper.

Acknowledgements: We thank Matthew Millard, Michael Dimitriou, Sae Franklin, Raz Leib, and Marion Forano for their comments on an earlier version of this manuscript.

Correspondence should be addressed to David W. Franklin at david.franklin@tum.de.

https://doi.org/10.1523/ENEURO.0514-19.2020

2013; Franklin et al., 2014, 2017; Cross et al., 2019), as well as sub-voluntary feedback onset times (Prablanc and Martin, 1992; Day and Lyon, 2000; Franklin and Wolpert, 2008; Oostwoud Wijdenes et al., 2011; Zhang et al., 2018). These visuomotor feedback responses have been shown to modulate throughout a movement depending on the perturbation onset location (Dimitriou et al., 2013). This observation was explained through optimality principles, however such control was modeled only indirectly, by replicating velocity profiles and trajectories of visually perturbed movements (Liu and Todorov, 2007; Rigoux and Guigon, 2012). In this study, we test to what degree optimal feedback control (OFC) can be used to model the visuomotor feedback responses directly.

Optimal control as a theory of human movement has normally been compared against other theories in terms of prediction of kinematics and dynamics (Todorov and Jordan, 2002; Izawa et al., 2008; Guigon et al., 2007, 2008; Nagengast et al., 2009; Yeo et al., 2016). Nevertheless, OFC has been used to motivate extensive studies investigating the control and task-dependent modulation of feedback responses (Knill et al., 2011; Pruszynski and Scott, 2012; Nashed et al., 2012, 2014). The results of these and other studies have highlighted the flexibility of the modulation of these feedback responses. While a few studies have compared the predictions of the controller feedback gains against the feedback responses in human subjects (Knill et al., 2011), such predictions have not been made about the temporal evolution of these feedback responses during reaching. For example, Dimitriou et al., 2013) show temporal evolution of feedback response intensity throughout a reaching movement, suggesting that this is similar to the feedback gain predictions of Liu and Todorov (2007). However, a direct comparison of these feedback intensities has not been made. Here, we directly compare the temporal evolution of visuomotor feedback response intensities in human participants with the prediction of these intensities in an OFC model.

Visuomotor feedback response intensity over a goal directed reaching movement follows a roughly bell-shaped profile, with peak intensity in the middle and decay toward the beginning and the end of the movement (Dimitriou et al., 2013). The results of Liu and Todorov (2007) suggest that such modulation is a combination of gains related to movement position, velocity and acceleration. However, we do not yet know whether these gains would more strongly depend on the visual kinematics or haptic kinematics. In addition, models of ball catching were shown to produce systematic errors in the prediction of the hand kinematics when using only velocity or acceleration based gains (Dessing et al., 2002), suggesting an integration of multiple state variables to produce the feedback response. Evidence of such integration then raises two important questions. First, could there be other states than position and its derivatives that also contribute to such control? Second, how can these responses be produced so rapidly, when multiple inputs need to be integrated into one solution?

One method to solve these two problems would be a controller based on time-to-target. Within a state-space system, all state variables are constantly changing with time with a fixed relationship to one another as described by the state transition and control matrices. Such a system can then be re-imagined as a system with time as its input, and these physical states as the hidden states. Such mapping simplifies the multiple input system where the inputs are state variables, to a one-input (time) system. Indeed, the expected time-to-target (or time-to-contact) has been shown to be related to the control in finger pointing (Oostwoud Wijdenes et al., 2011) and catching tasks (Dessing et al., 2002). Therefore, we test whether a simple relation to the time-to-target can explain the temporal profile of visuomotor feedback responses in humans. To test our hypotheses, we devised an experimental paradigm where we offset the usual bell-shaped velocity profile in the aim to separate the effect of the times-to-target from the effect of kinematics (both, of the hand and of the cursor) on the visuomotor feedback responses. Finally, we compare these results with a normative OFC model of visuomotor feedback responses to better understand how and whether these responses can be the result of optimality and still maintain rapid onset times.

Materials and Methods

Code availability

The code and the experimental data described in the paper is freely available online at https://doi.org/10.6084/m9.figshare.11323289. The code is available as Extended Data 1.

Participants

Eleven right-handed (Oldfield, 1971) human participants (five females; 27.3 ± 4.5 years of age) with no known neurologic diseases took part in the experiment. All participants provided written informed consent before participating. All participants except one were naive to the purpose of the study. Each participant took part in five separate experimental sessions, each of which took ~3 h. One participant was removed from analysis as their kinematic profiles under the five experimental sessions overlapped. The study was approved by the Ethics Committee of the Medical Faculty of the Technical University of Munich.

Experimental setup

Participants performed forward reaching movements to a target while grasping the handle of a robotic manipulandum with their right hand. Participants were seated in an adjustable chair and restrained using a four-point harness. The right arm of participants was supported on an air sled while grasping the handle of a planar robotic interface (vBOT; Howard et al., 2009). A six-axis force transducer (ATI Nano 25; ATI Industrial Automation) measured the end-point forces applied by the participant on the handle. Position and force data were sampled at 1 kHz. Visual feedback was provided in the plane of the hand via

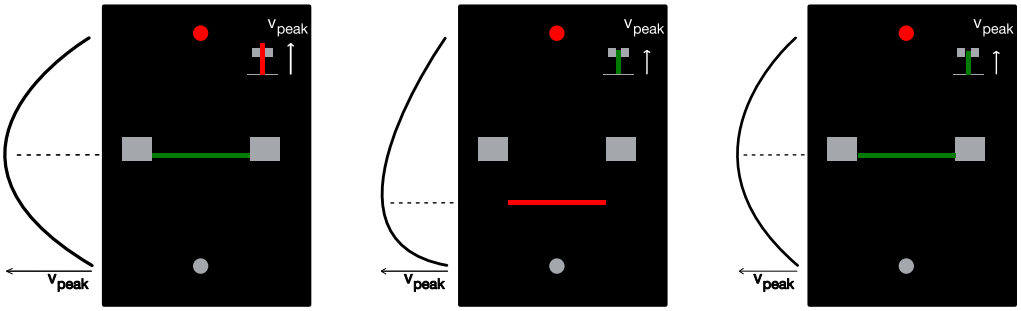


Figure 1. Examples of feedback presented to the participants. Feedback regarding the peak velocity and the timing of the peak velocity was provided after each trial. Large gray blocks indicate the velocity peak location target, while the bar chart at the top-right corner indicates peak *y*-velocity magnitude. Feedback was provided on the modality (cursor or hand) that matched the baseline, where the horizontal line indicated the location of the peak velocity in this modality. Left, Velocity peak location is within the target, but the movement was too fast (unsuccessful trial). Middle, Velocity peak location is too early, but the movement speed is within the target (unsuccessful trial). Right, Successful trial.

a computer monitor and a mirror system, such that this system prevented direct visual feedback of the hand and arm. The exact onset time of any visual stimulus presented to the participant was determined from the graphics card refresh signal.

Participants initiated each trial by moving the cursor (yellow circle of 1.0 cm diameter) into the start position (gray circle of 1.6 cm diameter) located ~25 cm in front of the participant, centered with their body. This start position turned from gray to white once the cursor was within the start position. Once the hand was within the start position for a random delay drawn from a truncated exponential distribution (1.0–2.0 s, mean 1.43 s), a go cue (short beep) was provided signaling participants to initiate a straight reaching movement to the target (red circle of 1.2 cm diameter, located 25.0 cm directly in front of the start position). If participants failed to initiate the movement within 1000 ms, the trial was aborted and restarted. Once the cursor was within 0.6 cm of the center of the target, participants were notified by the target changing color to white. The movement was considered complete when the participants maintained the cursor continuously within this 0.6 cm region for 600 ms. If participants did not complete the movement within 4 s from first arriving at the start position (e.g., by undershooting or overshooting the target), the movement timed-out and had to be repeated. Otherwise, as long as participants arrived at the target within 4 s, the trial was considered to have been completed. After each trial, the participant's hand was passively returned by the robot to the start position while visual feedback regarding the success of the previous trial was provided (Fig. 1). Movements were self-paced, and short breaks were enforced after every 100 trials.

Experimental paradigm

Participants performed the experiment under five different conditions, each performed in a separate session. In the baseline condition the cursor matched the forward movement of the hand, with a peak velocity in the middle of the movement. In the other four conditions, the cursor location was scaled relative to the hand location in the forward direction only (with no change in the lateral direction), such that the cursor and the hand location matched only at the start and end of the movements (Fig. 2). In two of the conditions (matched-hand velocity), the hand velocity matched the baseline condition throughout the movement (with the peak in the middle of the movement) but the cursor velocity peaked either earlier (33% of movement distance) or later (66% of movement distance). In the other two conditions (matched-cursor velocity), the cursor velocity was matched to the baseline condition throughout the movement (with the peak in the middle of the movement) but the hand velocity peaked either earlier (33% of movement distance) or later (66% of movement distance). The difference between the cursor velocity and the hand velocity was produced through a linear scaling of the cursor velocity as a function of the forward position (Fig. 2*A*). Specifically, for the two conditions where the position of the peak cursor velocity is earlier than the position of the peak hand velocity (Fig. 2, top), this scaling was implemented as:

$$\frac{v_c}{v_h} = -0.012d + 1.6, \quad (1)$$

where v_c and v_h are cursor and hand velocities, respectively, and d is the distance along the movement direction in %. The cursor velocity was therefore manipulated by a linear scaling function such that its velocity is 160% of the hand velocity at the beginning of the movement, linearly decreasing to 40% at the target location (Fig. 2, top). For the two conditions where the position of the peak cursor velocity is later than the position of the peak hand velocity (Fig. 2, bottom), this scaling was implemented as:

$$\frac{v_c}{v_h} = 0.012d + 0.4, \quad (2)$$

such that the velocity gain function linearly increased from 40% hand velocity at the start of the movement to 160% at the end of the movement (Fig. 2, bottom). Desired velocity profiles of both the hand and the cursor are shown in Figure 2*B* for each condition.

Introducing the differences in velocity profiles across five experimental conditions allows us to look at the effect of the time-to-target separately from the kinematics of

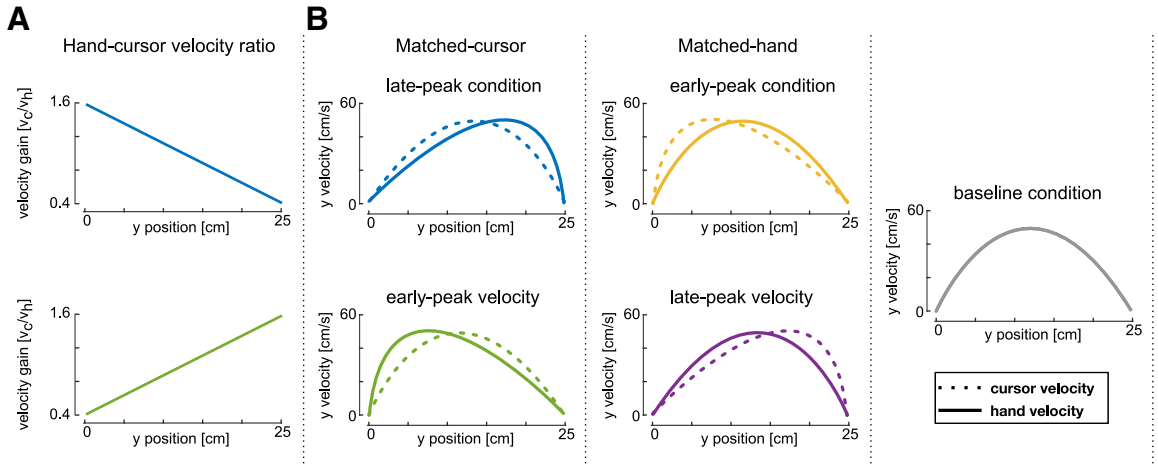


Figure 2. Experimental design. ***A***, top, Hand-cursor velocity scaling for conditions where the cursor position leads the hand position in *y*-axis (matched-cursor late-peak hand velocity condition, blue, and matched-hand early-peak cursor velocity condition, yellow). Bottom, Hand-cursor velocity scaling for conditions where the cursor position lags the hand position in *y*-axis (matched-cursor early-peak hand velocity condition, green, and matched-hand late-peak cursor velocity condition, purple). ***B***, Hand and cursor velocity-position profiles required to achieve the ideal movement to the target. Left, Matched-cursor velocity conditions. Middle, Baseline condition, where cursor position and hand position are consistent. Right, Matched-hand velocity conditions.

physical movement. First, as the visual perturbations always occurred at the five preset hand positions, our design allows us to maintain the effect of the distance to the target and distance in the movement constant across all five conditions. Second, two perturbation locations (one-third and two-thirds of movement distance) were chosen so that velocities at those locations matched across multiple conditions (early-peak condition velocity equals baseline at one-third, and late-peak velocity equals baseline at two-thirds), allowing for matching velocity contributions across conditions as well. However, across the three different physical kinematics the time-to-target is varied, allowing us to investigate whether time-to-target has an effect. Finally, the relative scaling between cursor velocity and hand velocity in the forward direction also separates the relative contributions of these two inputs, allowing us to examine the relative contributions of visual and physical kinematics in modulating the feedback responses.

Feedback regarding movement kinematics

In all conditions, one of the velocity modalities (cursor or hand) was required to be similar to the baseline velocity profile. Feedback was always provided about the velocity modality that matched the baseline. Ideal trials were defined as trials in which this peak velocity was between 42 and 58 cm/s with the peak location between 45% and 55% of the movement distance with no target overshoot. Participants were credited one point for achieving an ideal trial and zero points otherwise, however all the trials were included in the analysis. After each trial, visual feedback about the peak velocity and the location at which this peak occurred was provided to the participants graphically (Fig. 1). The peak velocity was indicated on the right-hand side of the screen with the length of a bar and the velocity target. This bar changed color from red to green if the velocity was within the ideal range. The location of the peak velocity was indicated as a horizontal line between home and target positions at the exact location it was achieved, along with the ideal range. This line was green when the location of the peak velocity was within the ideal range, and red otherwise. Overshooting the target was defined as the position of the cursor exceeding the center of the target in the forward direction by $>$0.9 cm. If participants reached the target while overshooting during the movement, a message indicating the overshot was shown, no points were scored and an error tone was played to discourage further overshooting movements.

Probe trials

During each session, probe trials were used to measure the visuomotor feedback intensity, the average strength of corrective motor response to a change in the visual feedback of hand position. To elicit these feedback responses (further visuomotor feedback responses), visual perturbations were initiated laterally ($\pm$2.0 cm) at five different hand distances (4.2, 8.3, 12.5, 16.7, and 20.8 cm) from the start (Fig. 3*A*). In addition, a zero-amplitude perturbation (cursor matched to the lateral position of the hand) was included, resulting in eleven different probe trials. On these trials the visual perturbations lasted 250 ms, after which the cursor was returned to the lateral location of the hand. The lateral hand position was constrained in these trials in a simulated mechanical channel throughout the movement, thereby requiring no correction to reach the target. The simulated mechanical channel was implemented with a stiffness of 4000 N/m and damping of 2 Ns/m acting perpendicularly to the line connecting the start position and the target (Scheidt et al., 2000; Milner and Franklin, 2005), allowing measurement of any lateral forces in response to a visual perturbation.

In previous experiments, feedback response intensity gradually decreased during the course of the experiment (Franklin and Wolpert, 2008; Franklin et al., 2012). However, it has been shown that including perturbation trials where

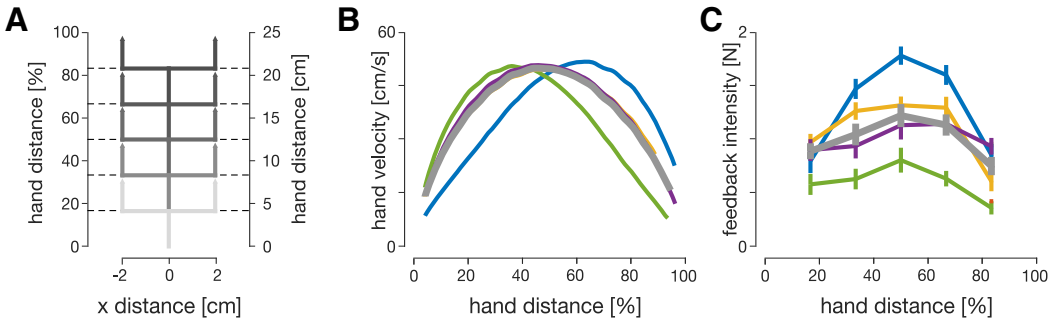


Figure 3. Human visuomotor feedback responses are modulated across the five experimental conditions. ***A***, Lateral perturbations of the cursor were applied in all five conditions. Perturbations were introduced as 2-cm cursor jumps perpendicular to the movement direction. The perturbation onset occurred at one of five equally spaced hand locations. ***B***, Mean velocity profiles of the hand in five experimental conditions: matched-cursor early-peak (green), matched-cursor late-peak (blue), matched-hand early-peak (yellow), matched-hand late-peak (purple), and baseline (gray). Participants successfully modulated forward movement kinematics to meet task demands, velocity profiles are skewed for matched-cursor conditions, and are similar to the baseline for matched-hand conditions. ***C***, Mean visuomotor feedback intensities (mean lateral force from 180 to 230 ms after perturbation onset) across all participants to cursor perturbations as a function of the hand distance in the movement. Error bars represent 1 standard error of the mean (SEM). Significant regulation is observed for matched-cursor early-peak and matched-cursor late-peak conditions (blue and green), but no significant regulation is seen for matched-hand conditions (yellow and purple), relative to the baseline.

the perturbations were maintained until the end of the movement, and where participants had to actively correct for the perturbation to reach the target, prevents this decrease in the feedback intensity (Franklin et al., 2016). Therefore, half of the trials contained the same range of perturbations as the probe trials but where these perturbations were maintained throughout the rest of the trial and participants had to correct for this perturbation. The hand movement was not constrained in these maintained perturbation trials. These maintained perturbations have now been used in several studies (Franklin et al., 2016, 2017; de Brouwer et al., 2017).

Session design

Before each session, participants performed 100–300 training trials to learn the specific velocity profiles of the reaching movements. All training trials contained no visual perturbations and were performed in the null force field. The training trials were stopped early once participants achieved an accuracy of 75% over the last 20 trials and were not used for the analysis.

Each session consisted of 40 blocks, where each block consisted of 22 trials performed in a randomized order. Eleven of these 22 trials were probe trials (five perturbation locations $\times$ two perturbation directions $+$ zero perturbation condition) performed in the mechanical channel. The other eleven trials consisted of the same perturbations but maintained throughout the trial and performed in the null field. Therefore, in each of the five sessions, participants performed a total 880 trials (440 probe trials). The order of the five different conditions (sessions) was pseudo-randomized and counterbalanced across participants. Participants were not told about the physical implementation of the different mappings, but were provided feedback after every trial and knew that each session was different from previous sessions.

Data analysis

Data were analyzed in MATLAB R2017b and JASP 0.8.2. Force and kinematic time series were low-pass filtered with a tenth-order zero-phase-lag Butterworth filter (40-Hz cutoff). The cursor velocity was calculated by multiplying the hand velocity by the appropriate scaling function. The visuomotor feedback response was measured for each perturbation location as the difference between the force responses to the leftward and rightward perturbations within a block. To measure the visuomotor feedback response intensity (mean force, produced as a response to a fixed-size visual perturbation) this response was averaged over a time window of 180–230 ms, a commonly used time interval for the involuntary visuomotor feedback response (Franklin and Wolpert, 2008; Franklin et al., 2012, 2016; Dimitriou et al., 2013). In order to compare any differences across the conditions a two-way repeated-measures ANOVA was performed with main effects of condition (five levels) and perturbation location (five levels). As a secondary method to frequentist analysis we also used the Bayesian factor analysis (Raftery and Kass, 1995) to verify our statistical results. Bayesian factor analysis is a method that in addition to the conventional hypothesis testing (evaluating evidence in favor of the alternative hypothesis) allows us to evaluate evidence in favor of the null hypothesis, therefore distinguishing between the rejection of the alternative hypothesis and not enough evidence to accept the alternative hypothesis.

Although we used the time window of 180–230 ms to estimate visuomotor feedback intensity, we also verified whether the onset of the visuomotor feedback response in our data are consistent with previously reported values. To estimate this onset time, we first estimated individual onset times for each participant at each perturbation location and movement condition. To do so, we used the receiver operator characteristic (ROC) to estimate where the force reaction to leftwards cursor perturbations deviated from the reaction to rightwards cursor perturbations (Pruszynski et al., 2008). For each type of trials, we built the ROC curve for the two signals at 1 ms intervals, starting from 50 ms before the perturbation, and calculated the area under this curve (aROC) for each of these points

until the aROC exceeded 0.75 for 10 consecutive milliseconds. In order to find where the force traces start deviating from each other, we then fit a function of the form $max(0.5, k \times (t - \tau)$ to the aROC curve. The time point where the linear component of this function first overtakes the constant component was taken as the threshold value. Overall, the mean onset times across all conditions and perturbation locations were 138 ± 7 ms (mean ± SD), with onset times consistent among movement conditions ($F_{(4,36)}$ = 1.410, p = 0.25, and BF_{10} = 0.105), perturbation locations ($F_{(4,36)}$ = 1.582, p = 0.20, BF_{10} = 0.252), and their interactions ($F_{(16,144)}$ = 1.350, p = 0.176, and BF_{10} = 0.005)

Modeling

OFC

In addition to our linear models we implemented two different OFC models: the classical model (Liu and Todorov, 2007) and the time-to-target model. The only comparison between the output of the optimal control models and the experimental results is via the feedback gains. For each movement we define time-to-target as the duration between the onset of the perturbation and the cursor first intercepting the target. In both models we modeled the hand as a point mass of m = 1.1 kg and the intrinsic muscle damping as a viscosity b = 7 Ns/m. This point mass was controlled in a horizontal plane by two orthogonal force actuators to simulate muscles. These actuators were controlled by the control signal u_t via a first order low-pass filter with a time constant τ = 0.05 s. The state-space representation of the dynamic system used to simulate the reaching movements can be expressed as

$$x_{t+1} = Ax_t + B(I + C)u_t + \xi_t, \tag{3}$$

where A is a state transition matrix, B is a control matrix, and C is a 2×2 matrix, whose each element is a zero-mean normal distribution representing control-dependent noise. Variables x_t and u_t are state and control at time t, respectively. State x_t exists in the Cartesian plane and consists of position $\vec{p}$ (2 dimensions), velocity $\vec{v}$ (2), force $\vec{f}$ (2), and target position $\vec{p^*}$ (2). The presence of these four states within the state vector means that the information about all of these states is eventually used for the control. For our simulation purposes we treat the control-independent noise ξ_t as zero.

The state of the plant is not directly observable, but has to be estimated from noisy sensory information. We model the observer as:

$$y_t = Hx_t + D_t, \tag{4}$$

where $H = \mathrm{diag}[1, 1, 1, 1, 1, 1, 0, 0]$ is the observation matrix, and D_t is a diagonal matrix of zero-mean normal distributions representing state-independent observation noise. Therefore, our observer can infer the state information of position, velocity and applied force of the plant, consistent with human participants.

The simulated movements were guided by the LQG controller with a state-dependent cost Q, an activation cost R, a reaching time N, and a time step t = 0.01 s. However, due to the presence of the control-dependent noise, the estimation and control processes are not anymore separable as in the classic LQG theory. In order to obtain optimal control and Kalman gain matrices we used the algorithm proposed by Todorov and Li (2005), where control and Kalman gain matrices are iteratively updated until convergence.

For both the classical and time-to-target models we simulated three different movement kinematics representing three different conditions in our experiment, the baseline and the two matched-cursor conditions. The state-dependent cost Q was identical for all three kinematics:

$$Q(t) = \begin{cases} 0, \text{ for } t \neq N \\ (\vec{\omega_p}(\vec{p}(t) - \vec{p^*}(t)))^2 + \omega_v ||\vec{v}(t)||^2 + \omega_f ||\vec{f}(t)||^2, \text{ for } t = N \end{cases}, \tag{5}$$

where $\vec{\omega_p} = [0.5, 1]$, $\omega_v = 0.02$, and ω_f = 2. The activation cost R(t) = 0.00001 was constant throughout the movement for the baseline condition, but was modulated for the two matched-cursor conditions by multiplying it elementwise by a scaling function:

$$R'(t) = \frac{\exp(p\frac{t+q}{r})}{mean(R')}, \tag{6}$$

where p, q, and r are constants.

Thus, each movement condition only differed from the other two by the profile of this activation cost R, but not by its magnitude. These modified activation costs shift the timing of the peak velocity toward either the beginning or the end of the movement by penalizing higher activations at either the end or beginning of the movements, respectively.

The mean activation cost is kept constant across the conditions resulting in each condition being equally "effortful." All other simulation parameters were kept constant across the three conditions.

Although LQG is a fixed time horizon problem, we did not predefine the movement duration N. Instead, we obtained the N, and constants p, q, and r using Bayesian adaptive direct search (BADS; Acerbi and Ma, 2017) to maximize the log-likelihood of the desired peak velocity location and magnitude. We did not fit any other parameters beyond this point. Rather, we analyzed our models' qualitative behavior compared with human participant data.

The classical and the time-to-target models only differed in the way the perturbations were handled. For the classical model, we simulated perturbation trials at every time step t_p by shifting the target x-coordinate by 2 cm at the time t_p + 120 ms. This 120 ms delay was used to mimic the visuomotor delay in human participants, and was taken from Liu and Todorov (2007). We then averaged the force response of the controller over the time window [t_p+130, t_p+180] as an estimate of the simulated feedback responses, equivalent of visuomotor feedback responses in our participants. This means that our simulated feedback responses arise due to separate contributions from the controller position, velocity and acceleration gains. For perturbations occurring at times where the movement is over

before the end of this time window, the intensity of this simulated feedback response is set to zero.

For the time-to-target model we introduced an extension in the time-to-target after the onset of any perturbation similar to that observed in our participants. Simulated feedback intensities were modeled at five locations, matching the perturbation locations in our experiment to obtain the appropriate increase in time-to-target after each perturbation. In order to simulate the response to perturbations we first extracted the perturbation onset times from movement kinematics by performing an unperturbed movement and recording the time point t_p at which this movement passed the perturbation onset location. We then simulated the post-perturbation portion of the movement as a new LQG movement with an initial state matching the state at t_p + 120 ms of the unperturbed movement, and movement duration matching the time-to-target recorded in our participants for the particular perturbation. Therefore, our time-to-target model can only simulate the feedback intensities at the five perturbation locations in the movement. Together, this keeps our simulated reaches "naive" to the perturbation before its onset and allows the time-to-target of the simulated reaches to match the respective time-to-target of our human participants. Finally, we calculated the simulated feedback intensities as described previously, using a time window [10, 60 ms] of the postperturbation movement. As in the previous simulations, these simulated feedback responses arise due to separate contributions from the controller position, velocity and acceleration gains.

Time-to-target tuning function

In order to understand the mechanisms that might underlie the consistent relationship between the simulated feedback intensities and the time-to-target, we fit a mathematical expression to the simulated feedback intensities. We modeled the relationship as the minimum of a squared-hyperbolic function and a logistic function:

$$G(t) = \min\left(\frac{\beta}{(t-t_1)^2}; \frac{\alpha}{1+\exp(-\frac{t-t_0}{\tau})}\right), \tag{7}$$

and used BADS to fit this function to our time-to-target-simulated feedback intensity data by optimizing the log-likelihood of this fit.

While the logistic function was chosen simply as it provided a good fit to the data, the squared-hyperbolic arises from the physics of the system. Specifically, from the kinematic equations of motion for a point mass (m) traveling a distance (d) under the influence of force F, the distance can be expressed as:

$$d = \frac{Ft^2}{2m} + v_0 t, \tag{8}$$

where $v_0 = 0$ is the lateral velocity at the start of perturbation correction. Rearranging gives:

$$F = \frac{2md}{t^2} \propto \frac{1}{t^2}. \tag{9}$$

Hence the lateral force necessary to bring a point mass to the target is proportional to $1/t^2$.

Receding horizon OFC

In addition to our finite horizon control we also implemented a receding horizon controller (Guigon et al., 2019). Irrespectively of the current state of the movement X_t, the receding horizon controller is defined to aim to arrive at the target at time $t + T_h$. In essence, such controller is therefore not different from the finite horizon controller in its implementation for a single state of the movement. We implemented the receding horizon controller by iterating a finite horizon controller described previously, but with the T_h = 500 ms, and Q and R costs scaled from the finite horizon model to fit the movement duration. For each iteration we recorded the next movement state (10 ms away from the initial state), and used that as the initial state for the next iteration. This process was repeated until the cursor was within the distance of 0.4 cm from the target position, and remained there without overshooting for 600 ms.

Simulating differently skewed velocity profiles within the framework of receding-horizon control is non-trivial. As a result, we chose to only model one, the baseline, experimental condition, where the activation cost R is constant within the movement. Therefore, we chose the costs

$$Q(t) = \begin{cases} 0, \text{ for } t \neq T_h \\ \vec{\omega_p}(\vec{p}(t) - \vec{p^*}(t))^2 + \omega_v ||\vec{v}(t)||^2 + \omega_f ||\vec{f}(t)||^2, \text{ for } t = T_h \end{cases}, \tag{10}$$

where $\vec{\omega_p} = [5, 5]$, $\omega_v = 0.05$, and $\omega_f = 5$. and the activation cost R = 0.000003. The values were selected so that the movement durations, produced by the receding-horizon model would match the experimental durations for the baseline condition. However, the resultant velocity profiles of this model more closely resembled those of the early-peak velocity condition, than those of the baseline. To account for any effects of the velocity profile we also fit the costs so the model prediction of movement durations matched the durations of the early-peak velocity condition. For this simulation we selected $\vec{\omega_p} = [0.7, 0.7]$, $\omega_v = 0.007$, and $\omega_f = 0.7$, while the activation cost remained unchanged.

In this model we introduced the simulated perturbation by shifting the target position by 2 cm at 120 ms after the y-coordinate of the movement passed the perturbation onset location. We only simulated the perturbations matching our experimental conditions, lateral 2 cm cursor jumps, with the onset at five evenly distributed forward distances. We calculated simulated feedback intensities the same way as for the classical and time-to-target models.

Infinite horizon OFC

We implemented the infinite horizon OFC to control our simulated hand based on the previous work of Qian et al. (2013). Specifically, we calculated the control gain matrix L, and Kalman gain matrix K to control the same system as in the previous models. We chose the state-dependent costs $\vec{\omega_p} = [1, 1]$, $\omega_v = 0.02$, and $\omega_f = 0$ for the baseline condition simulation, and $\vec{\omega_p} = [0.35, 0.35]$, $\omega_v = 0.007$,

and ω_f = 0 for the early-peak condition simulation. For both conditions, the activation cost R = 0.002 was kept the same. The protocol of simulating the mean trajectories, feedback responses and their intensities was otherwise identical to the receding horizon simulations.

Model comparison

We compared the simulated feedback intensities from each of the models with the experimental feedback responses intensities to evaluate our models. We do not evaluate models in terms of kinematics or any other variables. As the predictive simulated feedback intensities for each of the four models provided very different patterns, the important comparison is qualitative. However, we supplemented this qualitative comparison with a quantitative model comparison using the Bayesian information criterion (BIC). BIC is a conventional method for model comparison which evaluates the log-likelihood of the model fitting to the data while controlling for over-fitting by penalizing additional model parameters (Schwarz, 1978). A BIC difference of 10 is very strong evidence for the model with the lower BIC. Overall, we used individual participant mean feedback intensities for baseline, early-peak hand velocity and late-peak hand velocity conditions, providing us with a total of 150 data points (10 participants × three conditions × five perturbations) to determine the fit. Moreover, to compare the captured variance of the data between our OFC models and the time-to-target tuning curve we also calculated the sum of squared-residuals (SSRs) between the models and the data.

Results

Experimental results

In this study, we examine the relation between time-to-target (the time difference between the perturbation onset and the cursor intercepting the target) and the visuomotor feedback responses. To do so, we devised an experiment consisting of five different kinematic conditions. The baseline condition required movements with a natural, bell-shaped velocity profile, while the velocity profiles were modified for the four other conditions. In these four conditions, we introduced a manipulation between the hand velocity and the cursor velocity in the forward direction, such that the cursor and hand had different velocity profiles, but their positions matched at the start and end of the movement (Fig. 2). Two of these four conditions (matched-cursor conditions) required different kinematics of the physical movement to successfully complete the task, but the cursor velocity profiles matched the baseline. This manipulation of hand velocity profiles also resulted in different times-to-target at the same distance in the movement. The two other conditions (matched-hand conditions) required the same hand movement as for the baseline condition, but as a result the cursor moved with different velocity profiles (see Materials and Methods). This manipulation of the cursor velocity profiles separates the relative contributions of physical and visual hand information in regulating the feedback responses. For each condition we measured the visuomotor feedback intensities (mean corrective force applied during the 180 to 230 ms time window after a visual perturbation) at five different locations in the movement (Fig. 3*A*). Overall, our paradigm allowed us to modulate the times-to-target across conditions, as well as separate proprioceptive (hand) and visual (cursor) kinematics to examine their individual contribution to visuomotor feedback responses.

Different movement conditions exhibited differences in visuomotor feedback intensities (Fig. 3). Two-way repeated-measures ANOVA (both frequentist and Bayesian; Materials and Methods) showed significant main effects for both condition ($F_{(4,36)}$ = 10.807, $p < 0.001$, and $BF_{10} = 9.136 \times 10^{12}$), and perturbation location ($F_{(4,36)}$ = 33.928, $p < 0.001$, and $BF_{10} = 6.870 \times 10^{9}$). *Post hoc* analysis on movement conditions revealed significant differences between baseline (gray line) and matched-cursor late-peak hand velocity condition (blue line; $t_{(9)}$ = 4.262, $p_{bonf} < 0.001$ and BF_{10} = 247.868), and between baseline and matched-cursor early-peak hand velocity condition (green line; $t_{(9)}$ = –8.287, $p_{bonf} < 0.001$ and BF_{10} = 1.425 × 108). However, no significant differences were found between the baseline and the two matched hand velocity conditions ($t_{(9)}$ = 1.342, p_{bonf} = 1.0 and BF_{10} = 0.357 for early-peak cursor velocity, yellow; $t_{(9)}$ = 0.025, p_{bonf} = 1.0 and BF_{10} = 0.154 for late-peak cursor velocity, purple). Our results show that different kinematics of the hand movement have a significant effect on visuomotor feedback response regulation, but that different kinematics of the cursor movement do not.

One possible explanation for differences between the two matched-cursor conditions (Fig. 3*C*, blue and green) and the baseline condition (gray) might arise from a different mapping between cursor and hand velocities (Fig. 2*A*) that had to be learned. Alternatively, the incongruency between the vision and proprioception might be another explanation. However, the two matched-hand conditions (yellow and purple) had the identical mappings (and incongruencies) as the two matched-cursor conditions (blue and green, respectively) and yet no differences were found in these conditions. Instead, the only conditions in which differences in the feedback gains were found, were conditions in which the timing of the peak hand velocity was shifted.

In order to test whether a simple relationship between movement kinematics and visuomotor feedback intensities exists, we mapped visuomotor feedback intensity magnitudes as a linear function of the hand velocity and the cursor velocity. For each experimental condition, we find a different regression slope between the velocity and the feedback intensities regardless of whether this is the cursor or the hand velocity (Fig. 4*A*,*B*). Consistent with our previous results, this difference in slopes is significant for conditions where the hand, but not cursor, movement was different (Fig. 4*C*,*D*). Although feedback intensities increase with increasing velocity in both cursor and hand coordinates, no one coordinate modality could predict the changes in the feedback intensity.

To successfully complete each trial, participants were required to reach the target. However, the distance to reach the target is affected by the perturbation onset,

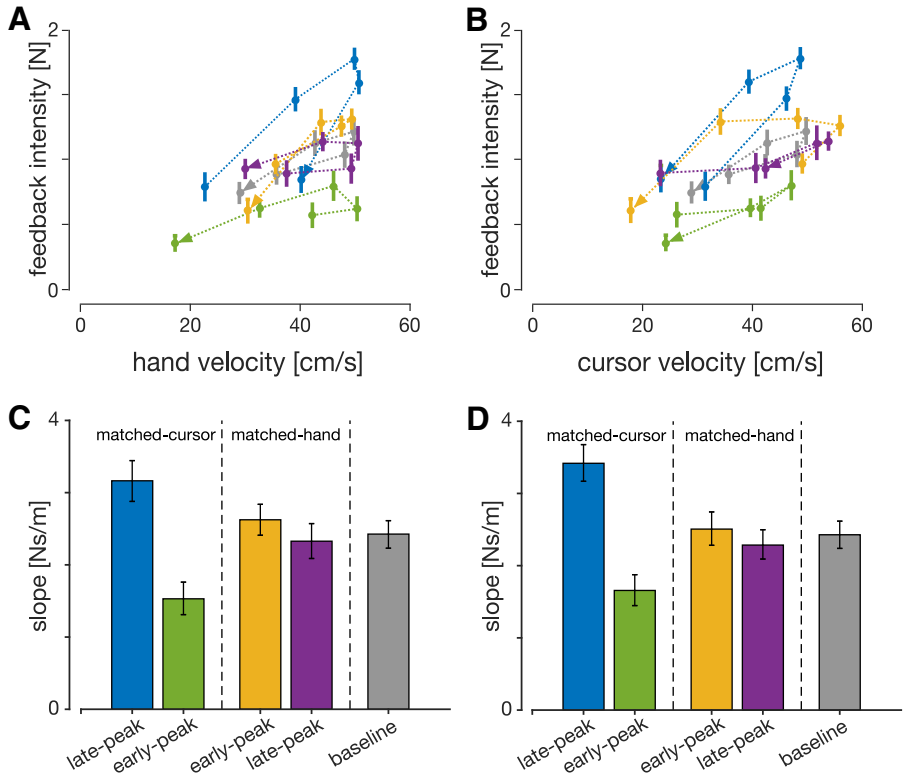


Figure 4. Visuomotor feedback intensities as a function of (***A***) hand velocity and (***B***) cursor velocity at the time of perturbation for all experimental conditions. Error bars represent 1 SEM, and the arrowheads represent the order of the perturbation locations. ***C***, ***D***, Regression slopes of feedback intensities for each condition as a function of hand and cursor velocities, respectively. Error bars represent 95% confidence intervals of the slopes. The slopes for the two matched-cursor conditions were significantly different (based on the confidence intervals) than for the baseline condition.

later perturbation locations lead to larger correction angles (Fig. 5*A*) and thus longer movement distances (Fig. 5*B*). That is, an earlier correction means that the trajectory can go directly toward the target, whereas a later correction would require a new corrective movement and therefore further distance. This effect is clearly seen where the extension of movement distance is enhanced for the perturbations closest to the target, with movement distance extended by almost 0.5 cm compared with less than 1 mm for the closest perturbations. Any extension of the movement distance requires an appropriate increase in movement duration. Consequently, participants extended their movement time, with longest durations for perturbations close to the target (Fig. 6*A*). This increase in movement duration increases the time-to-target for these late perturbations (Fig. 6*B*), and now allows sufficient time for the controller to issue any corrective commands.

Finite horizon OFC

As optimal control has been suggested to predict the temporal evolution of feedback intensities (Liu and Todorov, 2007; Dimitriou et al., 2013), we built two finite-horizon OFC models: the classical model (Liu and Todorov, 2007), and a time-to-target model. For the classical model we implemented an OFC (Todorov, 2005) to simulate movements with different velocity profiles, similar to the experiments performed by our participants. We extended this classical model to the time-to-target model, by increasing the movement duration after each perturbation onset according to experimental results (Fig. 6). For both models we only simulated different hand kinematics for computational ease and as our participants showed little effect of cursor kinematics on their feedback intensities.

For both models we controlled the activation cost R to simulate three conditions in which the location of the peak velocity was shifted to match the experimental hand kinematics (Fig. 7*A*). Specifically, we solved for the activation cost R and movement duration N by optimizing the log-likelihood of our model's peak velocity location and magnitude using BADS (Acerbi and Ma, 2017). The optimized movement durations (mean ± SEM) were N = 930 ± 0 ms for the baseline condition, N = 1050 ± 10 ms for the late-peak condition and N = 1130 ± 20 ms for the early-peak condition (10 optimization runs per condition). In comparison, experimental movement durations were N = 932 ± 30 ms for the baseline condition, N = 1048 ± 47 ms for the late-peak condition and 1201 ± 59 ms for the early-peak condition, matching well with the OFC predictions. Overall, this shows that specific constraints on the magnitude and location of peak velocity that we imposed on our participants resulted in a modulation of reaching times that matched OFC predictions under the same constraints.

For the classical model we estimated simulated feedback intensities by shifting the movement target at each time point in the movement and measuring the mean

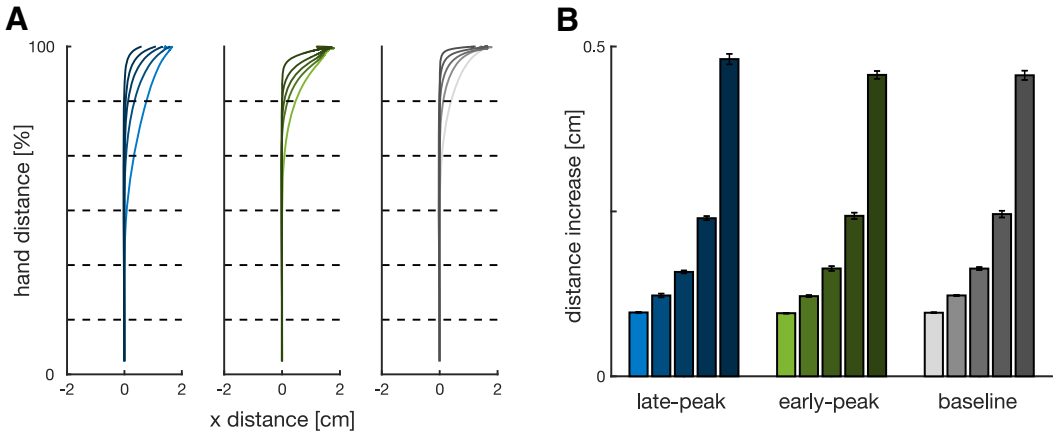


Figure 5. ***A***, Mean hand movement trajectories for matched-cursor late-peak (left), matched-cursor early-peak (middle), and baseline (right) conditions recorded in our participants, with perturbation onset at five locations [color light to dark: 4.2 cm (16.7%), 8.3 cm (33.3%), 12.5 cm (50%), 16.7 cm (66.7%), and 20.8 cm (83.4%) from the start position; dashed lines]. Corrections to rightward perturbations were flipped and combined with leftward corrections. ***B***, Distance increase for each perturbation location recorded in our participants. Perturbation locations closest to the target required the largest increases in movement distance. Error bars represent 1 SEM.

magnitude of the simulated force response over a 130- to 180-ms time window in the direction of this shift. The simulated feedback intensity profiles follow the same general shape as in human participants, intensity increases from the beginning of the movement and then falls off at the end (Fig. 7*B*). However, the overall profile of these simulated feedback intensities is very different for each of the kinematic conditions. For the early-peak velocity condition, the simulated feedback intensity peaks toward the end of the movement (green line), whereas for the late-peak velocity condition the simulated feedback intensity profile peaks early in the movement (blue line). These simulated feedback intensities do not appropriately capture the modulation of visuomotor feedback intensities in our experimental results. Specifically, they predict a temporal shift in the peak intensity that is not present in our participants data, and predict similar peak levels of feedback intensities across all three conditions. While the simulated feedback intensities are qualitatively similar to the experimental results within each condition; overall, this model cannot appropriately capture the modulation of visuomotor feedback responses across the conditions.

For the time-to-target OFC model, we extended the classical model to account for the different movement durations for each perturbation location (and movement condition) that is seen in the experimental results. After a

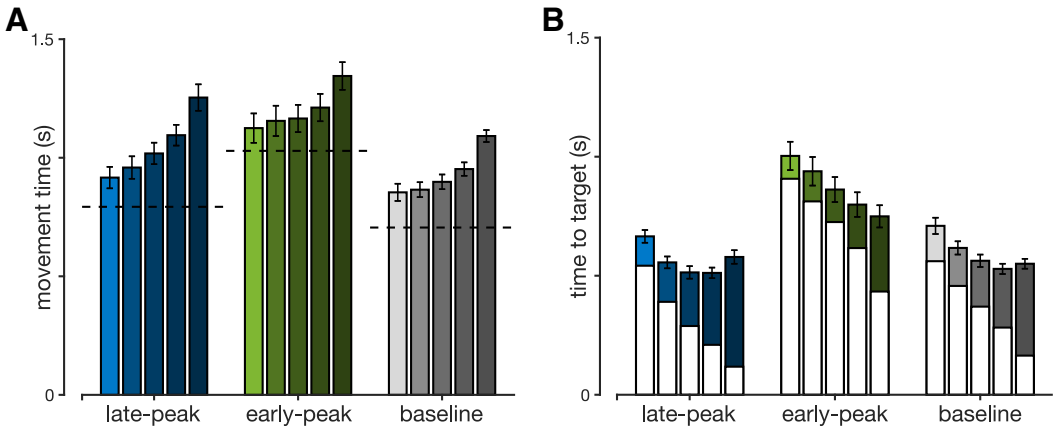


Figure 6. ***A***, Movement durations in maintained perturbation trials recorded by our participants in late-peak, early-peak and baseline conditions. Separate bars within the same color block represent different perturbation onset locations (left to right: 4.2, 8.3, 12.5, 16.7, and 20.8 cm from the start position). Error bars represent 1 SEM while the horizontal dashed lines represent movement durations in the same movement condition for non-perturbed movements. ***B***, Full bars represent times-to-target (time between a perturbation onset and target interception) in maintained perturbation trials in our participants for late-peak, early-peak, and baseline conditions. White bars represent the time-to-target for a respective non-perturbed movement, at the time when the perturbation would have happened. The colored part of the bars represents the extension in times-to-target due to the perturbation in a non-constrained movement. This shows that the perturbation during the movement evokes an extension in the time-to-target and subsequently in movement duration Each of the five bars represents a different perturbation onset location, as in ***A***. Error bars represent 1 SEM.

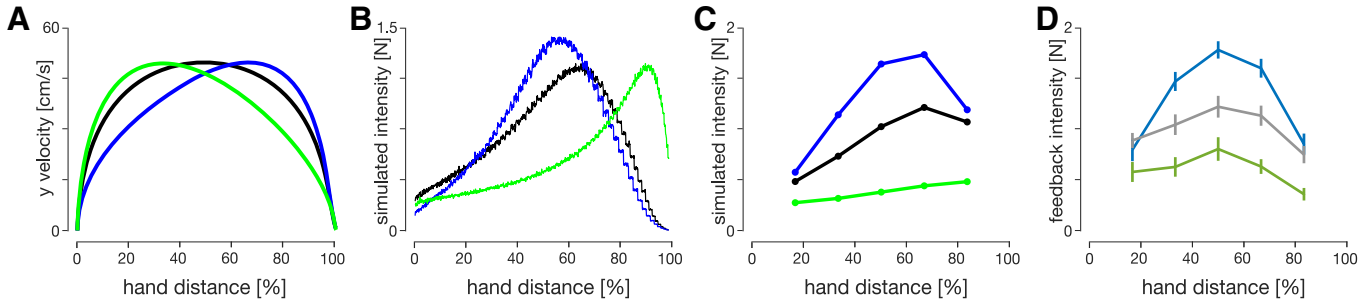


Figure 7. Comparison of feedback intensities between the two OFC models and experimental data. Simulated velocity profiles (***A***) and simulated feedback intensity profiles (***B***) of baseline (black), early-peak (green), and late-peak (blue) velocity condition simulations for the classical OFC model. Velocity profiles were obtained by constraining the velocity peak location and magnitude and optimizing for movement duration and activation cost function. Simulated feedback intensity profiles were obtained by applying virtual target jumps perpendicular to the movement direction during these movements and calculating the force exerted by the controller in the direction of the target jumps. The jagged appearance of the intensity traces is simply an outcome due to the simulation time step. ***C***, Simulated feedback intensities obtained via the time-to-target OFC model. Preperturbation movements were simulated as if no perturbation would occur, to keep the controller naive to an upcoming perturbation. At the perturbation onset the remaining movement duration is adjusted to match the mean time-to-target for a similar perturbation onset in human participants (Fig. 6*B*). Therefore, this model only simulates the feedback intensities at the five perturbation locations in the movement. The velocity profiles for the time-to-target model match the velocity profiles of the classical model, shown in ***A***. ***D***, Visuomotor feedback intensities recorded in human participants.

perturbation, the remaining time-to-target was adjusted to match the experimentally recorded times-to-target for this specific movement, while before the perturbation both the classical model and the time-to-target model were identical. After adjusting for the individual durations of each perturbation condition we are now able to qualitatively replicate the general regulation of feedback intensity profiles for different kinematics using OFC (Fig. 7*C*). In the late-velocity peak condition we predict a general increase in the feedback responses throughout the movement compared with the baseline condition, whereas in the early velocity peak condition we predict a general decrease in these feedback responses compared with the baseline condition. Thus, we show that within the OFC the time-to-target is critical for the regulation of feedback responses, and when we take this into account, we are able to replicate the feedback intensity modulation of our participants.

While in our experiment we manipulated the time-to-target through skewing the velocity profiles, time-to-target is naturally modified through changing the peak velocity. Therefore, we can further analyze the effect of the time-to-target by calculating the feedback intensities for movements with different peak velocities (Fig. 8*A*). The simulated feedback intensities vary widely across peak velocities, with a shift of peak feedback intensities toward the earlier locations for faster movements (Fig. 8*B*). However, when these distinct simulated feedback intensity profiles are re-mapped as a function of time-to-target, the simulated feedback intensities follow a consistent, albeit non-monotonic, relationship (Fig. 8*C*). This relationship is also consistent over a range of peak velocities across all three kinematic conditions and is well described by a combination of a square-hyperbolic and logistic function (Fig. 8*D*). The squared-hyperbolic arises from the physics of the system: the lateral force necessary to bring a point mass to a target is proportional to $1/t^2$ (Materials and Methods; Eq. 9). The logistic function simply provides a good fit to the data. Overall, our models show that the feedback intensity profiles under OFC are independent of the peak velocity or movement duration. Instead, our simulations suggest that time-to-target is a key variable in regulating visuomotor feedback responses.

It has been shown that the optimal controller gains (Liu and Todorov, 2007), as well as the visuomotor feedback intensities (Knill et al., 2011; de Brouwer et al., 2017) are influenced by task definition (e.g., instruction to hit the target or stop at the target). Here, we simulated the hit, fast hit and stop instructions for our classical model to test how it influenced the relation between simulated feedback intensity and time-to-target. Our previous simulations represent the stop instruction. We modified the ω_v and ω_f to simulate the baseline equivalent of hit and fast hit instructions. Specifically, we set $\omega_{v,hit} = \omega_v/4 = 0.05$, $\omega_{f,hit} = \omega_f/4 = 0.005$ for hit instruction, and $\omega_{v,fasthit} = \omega_v/10 = 0.02$, $\omega_{f,fasthit} = \omega_f/10 = 0.002$ for fast hit instruction. As changing the terminal costs also results in a change in peak velocity, we further reduced the desired movement times to N = 800 ms for the hit instruction and N = 750 ms for fast hit instruction, such that all three peak velocities match (Fig. 9*A*). According to our simulations, such modification of task demands produced different simulated feedback intensity profiles (Fig. 9*B*). However, the intensity relationship with time-to-target maintained the same structural profile independent of the task demand (Fig. 9*C*). Specifically, both the squared-hyperbolic and logistic segments of the control are still present, although we observe the shift in the temporal location of the crossover point. While each task requires a different pattern of feedback gains (and will therefore produce different responses), variations of the kinematic requirements within a task do not change these gains and therefore do not require recalculation.

Receding horizon and infinite horizon control

A limitation of the finite-horizon implementation used in classical and time-to-target models is that the variable movement duration (Fig. 6) is the model input rather than

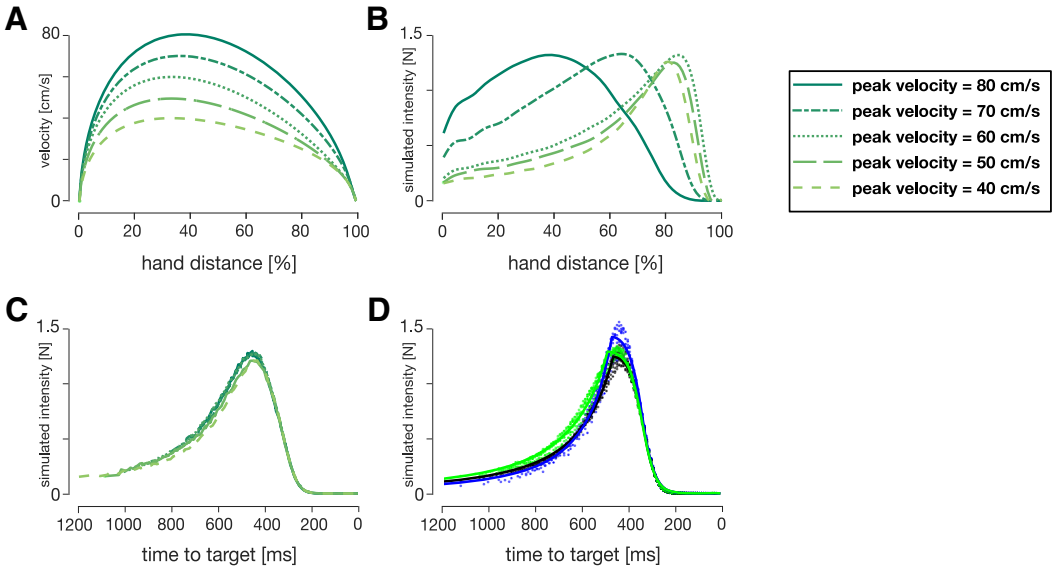


Figure 8. OFC simulations of (***A***) velocity profiles and (***B***) simulated feedback intensity profiles for different desired peak velocities (in order from light to dark line colors: 40, 50, 60, 70, and 80 cm/s). ***C***, Simulated feedback intensities of (***B***) re-mapped as a function of time-to-target at the time of perturbation. ***D***, Simulated feedback intensities vs time-to-target for the three kinematic conditions over the five peak velocities simulated by OFC (colored dots). Solid lines represent the tuning curves (Eq. 7) fit to the data. Both the tuning curves and the simulated feedback intensity profiles are similar across a variety of different kinematics when expressed as a function of time-to-target.

output. Therefore, in addition to finite-horizon models we also modeled our task in receding and infinite horizon for a single movement condition. Specifically, for the infinite horizon model both state-dependent and regulator costs were kept constant throughout the simulated movement. For the receding horizon model, the regulator cost was kept constant, while the state-dependent cost was zero for all but last "foreseeable" state. Such models were expected to simulate the baseline experimental condition, however the resultant velocity profile better resembled the early-peak condition (Fig. 10*A*). As a result, we compared these simulations with both baseline and early-peak velocity condition data and with the time-to-target model simulations (Fig. 10*B–D*).

Both receding horizon and infinite horizon LQG models were able to successfully capture the nonlinear change in trial durations for different perturbation onsets (Fig. 10*B*) matching the experimental results. In addition, these models also predicted variable times-to-target for the five perturbation onset locations: 700, 660, 620, 600, and 580 ms for the infinite horizon and 690, 640, 610, 610, and 600 ms for the receding horizon. However, neither model showed variation of the simulated feedback intensities for different perturbation onset locations (Fig. 10*C*,*D*), a result that was present in the experimental data and captured by our time-to-target model. Instead both models predicted constant feedback intensities for all perturbations locations. Therefore, neither the receding nor the infinite horizon models are able to explain our experimental results. While both of the approaches can accurately capture the variability in movement duration, only the time-to-target model well describes the behavioral variation in visuomotor feedback responses.

Quantitative model comparison

Qualitatively, our results suggest that the time-to-target is an important variable when correcting for visual perturbations in a visuomotor task. In order to supplement these findings quantitatively, we also evaluated model fits

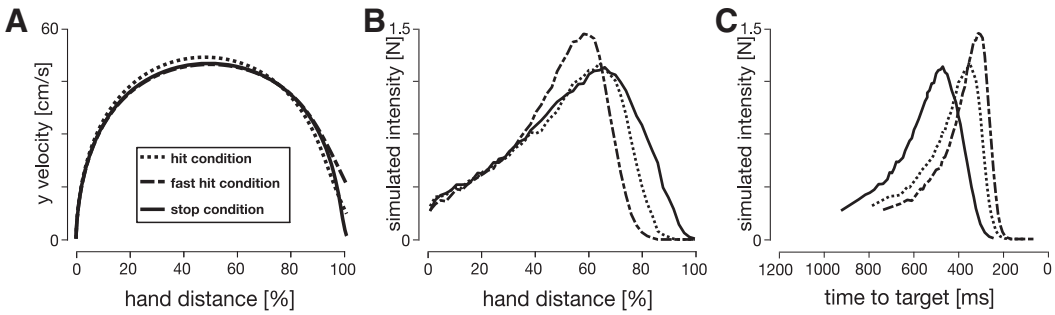


Figure 9. Comparisons between hit and stop instructions. ***A***, Velocity profiles for the stop, hit and fast-hit conditions. ***B***, Simulated feedback intensity profiles as a function of hand position. ***C***, Simulated feedback intensities of (***B***) re-mapped as a function of time-to-target at the time of target perturbation.

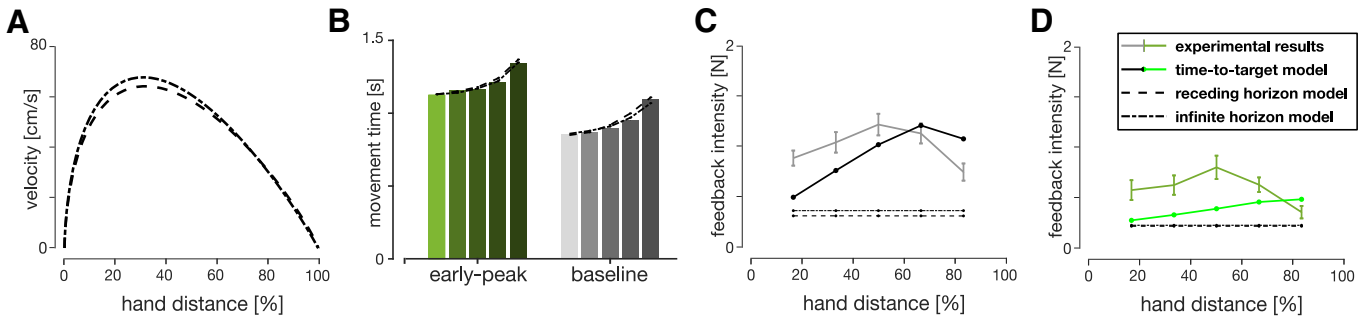


Figure 10. Receding horizon and infinite horizon model simulations. ***A***, Simulated velocity profiles of receding horizon (dashed) and infinite horizon (dot-dashed) models. Both models naturally produce positively skewed velocity profiles, more closely resembling early-peak velocity, rather than the baseline condition. ***B***, Mean experimental movement durations (bar chart) compared with the receding and infinite horizon model predictions. Both models accurately simulate the variations in the reach durations with perturbation location. Baseline (***C***) and early-peak velocity condition (***D***) simulations for receding horizon, infinite horizon and time-to-target (dot-solid lines) models, compared with the experimental data. Only the time-to-target model predicts different visuomotor feedback response intensities for different perturbation onset locations, while receding and infinite horizon models predict constant intensities. Note that models were not fit to match the intensities, only to qualitatively demonstrate the behavior.

between the data and the models using BIC and calculating the SSR for each of the models. We compared the OFC-based models with respect to the classical, finite horizon OFC model as our baseline model. Consistent with our qualitative estimations, the time-to-target OFC model performed the best of all OFC-based models ($\Delta\, BIC_{ttt}$ = 38.2). In addition, both receding-horizon and infinite-horizon models provided bad fits to the data ($\Delta BIC_{rec} = -23.4, \Delta BIC_{inf} = -18.4$). We also compared, using SSR, the fit of our OFC models with the simpler time-to-target tuning curve (Eq. 7; Fig. 11*A*). While the tuning curve is only a simple approximation to the time-to-target adjusted OFC feedback predictions, it is still able to explain a similar amount of variance in the data (R^2 = 0.33; $SSR_{tuningcurve} = 28.5$, $SSR_{classical} = 33.3, SSR_{ttt} = 24.9, SSR_{inf} = SSR_{rec} = 38.9$). Overall, both the BIC and SSR comparison confirms that the time-to-target OFC model best explains the feedback modulation during human reaching and suggests that time-to-target is a critical variable in online control.

Overall, our simulations suggest that, independent of movement kinematics (different temporal position, velocity, and acceleration profiles), the visuomotor feedback intensities follow the same profile with respect to the time-to-target. We further verified how our time-to-target prediction matches our actual experimental results by plotting participants' visuomotor feedback intensities against the average time-to-target for the respective perturbation locations and movement conditions (Fig. 11*A*). Specifically, the intensities monotonically increase with decreasing time-to-target until the peak (following the squared-hyperbolic function) and then reduce (the logistic function range).

Validation of the time-to-target model

We also compared the prediction of the time-to-target model to independent results from an external data set (Dimitriou et al., 2013). In the article, the authors could not rigorously encapsulate both conditions within a simple relationship to movement distance, movement fraction or movement velocity. We plotted visuomotor feedback intensities against time-to-target for two experimental conditions: goal directed reach of 17.5 cm and of 25 cm (Fig. 11*B*,*C*). Two observations can be made from these results. First, the time-to-target model prediction and the experimental data follow the same qualitative features, independent of the target distance (experimental condition). Second, the feedback intensities for both conditions are well explained by a single relationship with time-to-target (Fig. 11*C*; R^2 = 0.56, $SSR_{tuning\ curve}$ = 2.3). Thus, the validation against an external dataset supports our results that visuomotor feedback intensities vary with the time-to-target.

Finally, we evaluated our optimal control models (the classical and the time-to-target) on this dataset. Similar to our original fitting, we fit both of our OFC models to match the kinematics of the human participants, and then simulated the virtual experiment to extract simulated feedback intensities for the models. We found that model parameters $\vec{\omega_p} = [0.5, 1]$, $\omega_v = 0.03$, and ω_f = 0.03 and R(t) = 0.00000235 provide the best fit of kinematics between OFC models and data. As with our data, the time-to-target OFC model provided a better fit of the simulated feedback gains than the classical OFC model (ΔBIC = 19.2, SSR_{ttt} = 8.6, $SSR_{classical}$ = 10.7). Together, both our data and Dimitriou et al. (2013) data strongly support our time-to-target model.

Discussion

Here, we examined how movement kinematics regulate visuomotor feedback responses. Participants extended their movement duration after perturbations to successfully reach the target. In addition, visuomotor feedback responses were modulated when the hand followed different kinematics, but not when the cursor followed different kinematics. In order to better understand this modulation we built four normative models using OFC: a classical finite-horizon OFC (Liu and Todorov, 2007), a finite-horizon time-to-target adjusted OFC, a receding-horizon OFC (Guigon et al., 2019), and an infinite-horizon OFC (Qian et al., 2013). While the classical, receding and infinite horizon models failed to predict the experimental visuomotor feedback

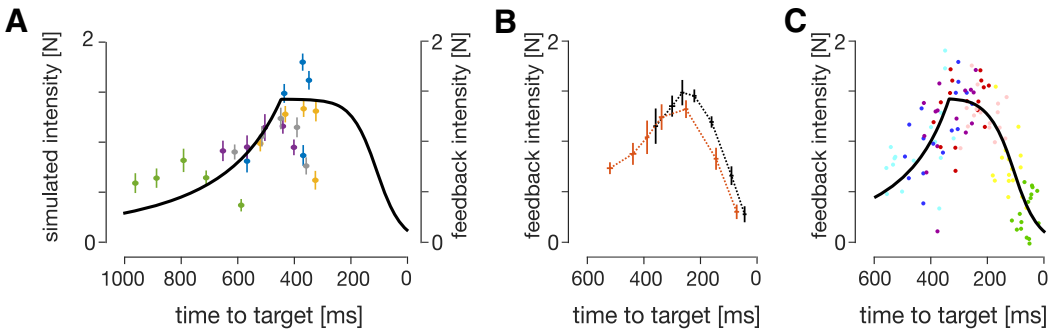


Figure 11. Validation of the time-to-target model. ***A***, Experimental visuomotor feedback intensities for all five experimental conditions (scatter plot) overlaid with the time-to-target tuning curve. The data and the tuning curve show similar qualitative features. Error bars represent 1 SEM. Marker colors indicate five experimental conditions as described in Figure 2*B*. ***B***, Experimental data of the visuomotor feedback intensities of Dimitriou et al. (2013), mapped against the time-to-target. Black and orange traces represent mean participant data for 17.5 and 25 cm movement conditions, respectively. ***C***, A scatter plot of individual subjects' data from ***B***, overlaid by the time-to-target tuning curve. Both, 17.5 and 25 cm movement conditions are combined to a single representation. Different colors represent different perturbation onset distances as in Dimitriou et al. (2013).

response intensities, the time-to-target model qualitatively replicated the visuomotor feedback intensity profile of our participants. Overall, OFC models suggested that feedback intensities for each perturbation location depended on the time-to-target (i.e., time between perturbation onset and target interception) rather than distance or velocity. Moreover, this explains why any mismatch between visual and haptic kinematics had no effect on the feedback intensities, as these manipulations did not affect the time-to-target. Simulated feedback intensities under all movements followed the same profile with respect to time-to-target, suggesting a critical role in the regulation of visuomotor feedback responses.

Experimentally, our participants exhibited a temporal evolution of visuomotor feedback intensities for each condition, confirming the findings of Dimitriou et al. (2013). In addition, we also showed the regulation of visuomotor feedback responses across conditions, allowing us to investigate the underlying mechanism of this temporal evolution. Specifically, our experimental results demonstrated strong regulation of visuomotor feedback intensity profiles with different hand kinematics, but not with different cursor kinematics (Fig. 3*C*). Compared with the baseline condition, in the matched-cursor early-peak velocity condition participants produced longer times-to-target at each perturbation location (Fig. 6*B*), resulting in weaker feedback responses based on the relationship between time-to-target and visuomotor feedback intensities (Fig. 11*A*). The opposite is true for the matched-cursor late-peak velocity condition. As the two matched-hand conditions produced similar times-to-target as the baseline due to similar hand kinematics, we did not observe a different regulation in feedback responses. Therefore, the condition-dependent visuomotor feedback response modulation exhibited by our participants meshes nicely with a control policy whereby the time-to-target regulates the feedback responses.

It has long been suggested that we select movements that minimize the noise or endpoint variability (Harris and Wolpert, 1998). Within the framework of optimal control, this idea has been expanded to the corrective movements, that is, optimality in reaching movements is achieved in part by minimizing the noise during any corrective response (Todorov and Jordan, 2002). As motor noise scales proportionally to muscle activation (Jones et al., 2002; Hamilton et al., 2004), one way of minimizing such noise is reducing the peak levels of muscle activation during the correction. Mathematically, the optimal solution to correct any perturbation approximates a constant activation, resulting in a constant force for the whole duration between perturbation onset and target interception. Such a solution assumes that the brain is capable of estimating the remaining duration of the movement (McIntyre et al., 2001; Benguigui et al., 2003; Zago et al., 2004) and that the force follows the squared-hyperbolic relationship to this duration (Eq. 9). The parallel can be drawn here between our results and the results of Oostwoud Wijdenes et al. (2011), where the authors showed a similar temporal evolution of peak acceleration against the time-to-target in a single forward velocity condition. Our results further show that time-to-target strongly modulates visuomotor feedback responses across a range of different kinematics, consistent with the idea that human participants aim to behave optimally. More specifically, we suggest that, among different optimality variables, the temporal evolution of visuomotor feedback response intensities serves to reduce effects of system noise.

Finite-horizon OFC predicts a time beyond which feedback responses are suppressed. Beyond this critical time, a logistic function well describes the relation between time-to-target and feedback responses, with response intensities reducing as the time-to-target decreases. The controller gains at this stage are the most sensitive to acceleration, suggesting a more "behavioral" outcome, the controller is trying to stop, rather than correct errors. The neural recordings in rhesus macaque monkeys' supplementary motor area and M1 (Russo et al., 2019) show that supplementary motor area can signal movement termination as far as 500 ms before the end of the movement. This further suggests that there may be multiple stages within a movement, where our control system might "care" more about error correction in one or movement termination in another. On the other hand, the suppression of responses

close to the target leads to undershooting the target. Our participants, however, had to bring the cursor to the target to advance to the next trial. As a result, they extended the movement durations postperturbation to return to the squared-hyperbolic range of control. The control performance of such behavior is well accounted for by our time-to-target model. Moreover, our time-to-target model also well explained the modulation of visuomotor feedback intensities from an external data set (Dimitriou et al., 2013). However, an important distinction from our study is that in Dimitriou et al. (2013), the suppression of feedback responses toward the end of movements would not interfere with reaching the target as perturbation trials were always in a mechanical channel so that no corrections were required. As a result, the times-to-target were shorter and the data clearly exhibits both logistic and squared-hyperbolic segments of the control.

All of the variations of optimal control models are formulated as two controllers in *x*- and *y*-axes (coupled through control-dependent noise), with no modeling of the musculoskeletal dynamics. However, the experiments were performed using multijoint reaching movements of the arm. One possibility is that the presence of these musculoskeletal dynamics in the human participants could explain the differences in the feedback intensities, as the matched cursor conditions required different hand accelerations. One might therefore imagine that the condition with the fastest initial movements (early-peak velocity) requires the largest initial forces and could therefore produce larger initial feedback intensities as a default. However, several studies have shown that there is no scaling of visuomotor feedback gains with background loads or muscle activity (Franklin et al., 2012, 2017). More critically, this condition actually shows the lowest feedback gains early in the movement, whereas the condition with the slowest initial acceleration produces the highest feedback gains. Therefore, we suggest that the neuromuscular dynamics cannot explain the modulation of these feedback responses.

Both of the matched cursor conditions (early and late peak) require a change in the physical kinematics away from the naturally occurring bell-shaped velocity profile. One possibility is that this manipulation could have driven the changes in feedback intensity. We argue against this possibility for two reasons. First, all participants were able to fairly quickly learn this pattern of movement with training before the testing of the feedback intensities. Second, if this manipulation away from the naturally occurring bell-shaped profile affects these feedback intensities, we would expect the feedback gains in these two conditions to either both increase or both decrease. Instead we find that the feedback intensity profile for each condition changes in a manner that is explained by the change in the time-to-target.

A limitation of our time-to-target model is that it takes time-to-target as an input to generate feedback intensity predictions, rather than obtain the time-to-target as a model output. As a result, our time-to-target model does not describe exactly how the change in movement geometry after the perturbation influences this time-to-target, which in turn regulates the visuomotor feedback responses. On the other hand, both receding and infinite horizon models did predict the movement duration change after perturbations very well, but could not at all describe the changes in visuomotor response intensity. However, utility of movement has recently been used within optimal control to characterize reaching movements (Rigoux and Guigon, 2012; Shadmehr et al., 2016) in which optimal movement time falls out automatically from a trade-off between reward and effort. With respect to our models, this adds additional complexities to capturing the different movement conditions. Future approaches could attempt to model these results within the utility of movement framework.

In addition, our time-to-target model does not directly show the causality of the time-to-target as a control variable for the visuomotor feedback intensities. Particularly, the time-to-target relation to feedback intensity could be a by-product of a more sophisticated control scheme. Additional arguments for the time-to-target control scheme could be two-fold. First, there is evidence that humans are well capable of estimating the time-to-target of a moving stimulus, even if it is accelerating (McIntyre et al., 2001; Benguigui et al., 2003; Zago et al., 2004), indicating that time-to-target is at least an available input for such a controller. Second, while we have tested finite-horizon OFC and two other (receding and infinite horizon) OFCs, only the finite horizon controllers had any effect on the variation of simulated feedback intensities. Importantly, neither the receding nor infinite horizon models use time-to-target as an input to the controller. We posit that this time-to-target control input is the one key difference between the finite and non-finite models and is therefore the simplest explanation for our results.

Our results show that models incorporating time-to-target (the time-to-target OFC and the simple time-to-target tuning curve) better describe our experimental data and those of Dimitriou et al. (2013) than do other optimal control models. Specifically, the relative scaling of the conditions is explained using the time-to-target tuning curve (Fig. 11*A*): the times-to-target are longer for the early-peak velocity condition compared with the baseline, and therefore fall in the lower intensity range (and vice versa for the late-peak condition). However, there are still some qualitative differences between the experimental and model predictions. That is, our experimental results exhibited an inverted U-shape for the feedback intensity profiles, whereas the model predicts only a slow increase in some conditions (Fig. 8*C*, green curve). Our time-to-target model suggests that this inverted U-shape is not characteristic of the feedback intensity profile, but is simply an outcome of the experimental design (particularly the reaching duration). Indeed, the time-to-target model makes specific predictions about the feedback intensities for much faster movements, which should not show an inverted U-shape but instead decrease throughout the movement. Our model, therefore, makes strong predictions that can be tested in future studies.

Rapid feedback responses scale with the temporal urgency to correct for mechanical perturbations (Crevecoeur et al., 2013). Here, we have shown that visuomotor feedback

responses also follow a similar regulation, suggesting that these two systems share the same underlying control policy. Our work further extends this finding of Crevecoeur et al. (2013) by not just showing that temporal urgency affects feedback responses, but explaining the manner in which these responses are regulated with respect to urgency. That is, here, we have shown that for visual perturbations the feedback intensities scale with a squared-hyperbolic of the time-to-target, which is a direct measure of urgency. Moreover, the feedback intensities were rapidly adjusted due to the change in urgency as the task changed. Specifically, when the cursor jumps close to the target, the expected time-to-target is prolonged, and therefore the optimal visuomotor feedback response needs to be adjusted appropriately to this increase in time. Our results show that participants produce a visuomotor response consistent with the actual, postperturbation, time-to-target, as opposed to the expected time-to-target before the perturbation. Therefore, our results not only suggest that similar computations might occur for both stretch and visuomotor feedback response regulation, but also that this regulation originates from task-related OFC.

Our proposed time-to-target model is not meant to contradict the conventional OFC models, but rather show that the OFC could be approximated by a simple time-to-target control. Our work has shown that simulated feedback intensities from OFC exhibit the same underlying pattern as a function of time-to-target over a wide range of movement kinematics, matching well the feedback intensities of our human participants (Fig. 7). As expected, changes in the task goals (e.g., hit vs stop) changed the relation between feedback responses and time-to-target. However, the qualitative features, the squared-hyperbolic and logistic function, remained consistent across these tasks. These results suggest that, for a specific task and known dynamics, we do not need to recalculate the feedback gains before each movement, but instead can access the appropriate pattern as a function of the estimated time-to-target in each movement. Therefore, gain computation in reaching movements may not be a computationally expensive process, but instead could be part of an evolutionary control strategy that allows for rapid estimation of the appropriate feedback gains. Moreover, the fact that both stretch reflex and visuomotor feedback systems exhibit similar control policies despite different sensory inputs, perhaps only sharing the final output pathway, suggests that this simple feedback pathway may be an evolutionary old system. Indeed, several studies have suggested that visuomotor feedback is controlled via a pathway through the colliculus (Corneil et al., 2004; Reynolds and Day, 2012; Gu et al., 2019). Furthermore, it has been suggested that visuomotor feedback responses involve two different phases that are behaviorally different Cross et al. (2019). This might reflect two different pathways, the early through the colliculus and the later through cortex. The nature of our analysis only focuses on the earlier of the two phases, which shows limited sensitivity to environment, but is still sensitive to goal redundancy. We suggest that this limited sensitivity could be the outcome of the time-to-target model in action, providing simplified, yet still flexible control in the early phase of the visuomotor response. Such a system would then only need to be adapted as the dynamics or overall task goals change, allowing for fine tuning of the feedback gains according to changes in the environment (Franklin et al., 2017).

Our results have shown the connection between the visuomotor feedback response regulation and the time left to complete the movement. Specifically, in our human participants we recorded the increase in the time-to-target after the perturbation onset, which consequently increased the movement durations (Fig. 6). This increase was also longer for later perturbations, consistent with previous studies (Liu and Todorov, 2007). According to our normative time-to-target OFC model, the time-to-target alone is enough to successfully regulate visuomotor feedback responses as observed in humans. This result was independent of the physical kinematics of the movement or the onset times of the perturbations. This suggests that there is no recalculation of a control scheme for the rest of the movement after the perturbation, but rather a shift to a different state within the same control scheme. Such findings are consistent with the idea that visuomotor feedback gains are precomputed before the movement, allowing for faster than voluntary reaction times (Franklin, 2016). Moreover, through our results, we gain a deeper insight into how OFC governs these feedback gains, through a straightforward relationship to the estimated time-to-target, based on physics.

References

Acerbi L, Ma WJ (2017) Practical Bayesian optimization for model fitting with Bayesian adaptive direct search. Adv Neural Inf Process Syst 30:1836–1846.

Benguigui N, Ripoll H, Broderick MP (2003) Time-to-contact estimation of accelerated stimuli is based on first-order information. J Exp Psychol Hum Percept Perform 29:1083–1101.

Corneil BD, Olivier E, Munoz DP (2004) Visual responses on neck muscles reveal selective gating that prevents express saccades. Neuron 42:831–841.

Crevecoeur F, Kurtzer I, Bourke T, Scott SH (2013) Feedback responses rapidly scale with the urgency to correct for external perturbations. J Neurophysiol 110:1323–1332.

Cross KP, Cluff T, Takei T, Scott SH (2019) Visual feedback processing of the limb involves two distinct phases. J Neurosci 39:6751–6765.

Day BL, Lyon IN (2000) Voluntary modification of automatic arm movements evoked by motion of a visual target. Exp Brain Res 130:159–168.

de Brouwer AJ, Jarvis T, Gallivan JP, Flanagan JR (2017) Parallel specification of visuomotor feedback gains during bimanual reaching to independent goals. eNeuro 4:ENEURO.0026-17.2017.

de Brouwer AJ, Gallivan JP, Flanagan JR (2018) Visuomotor feedback gains are modulated by gaze position. J Neurophysiol 120:2522–2531.

Dessing JC, Bullock D, Peper CE, Beek PJ (2002) Prospective control of manual interceptive actions: comparative simulations of extant and new model constructs. Neural Netw 15:163–179.

Dimitriou M, Wolpert DM, Franklin DW (2013) The temporal evolution of feedback gains rapidly update to task demands. J Neurosci 33:10898–10909.

Franklin DW (2016) Rapid feedback responses arise from precomputed gains. Motor Control 20:171–176.

Franklin DW, Wolpert DM (2008) Specificity of reflex adaptation for task-relevant variability. J Neurosci 28:14165–14175.

Franklin DW, Franklin S, Wolpert DM (2014) Fractionation of the visuomotor feedback response to directions of movement and perturbation. J Neurophysiol 112:2218–2233.
Franklin DW, Reichenbach A, Franklin S, Diedrichsen J (2016) Temporal evolution of spatial computations for visuomotor control. J Neurosci 36:2329–2341.
Franklin S, Wolpert DM, Franklin DW (2012) Visuomotor feedback gains upregulate during the learning of novel dynamics. J Neurophysiol 108:467–478.
Franklin S, Wolpert DM, Franklin DW (2017) Rapid visuomotor feedback gains are tuned to the task dynamics. J Neurophysiol 118:2711–2726.
Gu C, Pruszynski JA, Gribble PL, Corneil BD (2019) A rapid visuomotor response on the human upper limb is selectively influenced by implicit motor learning. J Neurophysiol 121:85–95.
Guigon E, Baraduc P, Desmurget M (2007) Computational motor control: redundancy and invariance. J Neurophysiol 97:331–347.
Guigon E, Baraduc P, Desmurget M (2008) Optimality, stochasticity, and variability in motor behavior. J Comput Neurosci 24:57–68.
Guigon E, Chafik O, Jarrassé N, Roby-Brami A (2019) Experimental and theoretical study of velocity fluctuations during slow movements in humans. J Neurophysiol 121:715–727.
Hamilton AFDC, Jones KE, Wolpert DM (2004) The scaling of motor noise with muscle strength and motor unit number in humans. Exp Brain Res 157:417–430.
Harris CM, Wolpert DM (1998) Signal-dependent noise determines motor planning. Nature 394:780–784.
Howard IS, Ingram JN, Wolpert DM (2009) A modular planar robotic manipulandum with end-point torque control. J Neurosci Methods 181:199–211.
Izawa J, Rane T, Donchin O, Shadmehr R (2008) Motor adaptation as a process of reoptimization. J Neurosci 28:2883–2891.
Jones KE, Hamilton AF, Wolpert DM (2002) Sources of signal-dependent noise during isometric force production. J Neurophysiol 88:1533–1544.
Knill DC, Bondada A, Chhabra M (2011) Flexible, task-dependent use of sensory feedback to control hand movements. J Neurosci 31:1219–1237.
Liu D, Todorov E (2007) Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. J Neurosci 27:9354–9368.
McIntyre J, Zago M, Berthoz A, Lacquaniti F (2001) Does the brain model Newton's laws? Nat Neurosci 4:693–694.
Milner TE, Franklin DW (2005) Impedance control and internal model use during the initial stage of adaptation to novel dynamics in humans. J Physiol 567:651–664.
Nagengast AJ, Braun DA, Wolpert DM (2009) Optimal control predicts human performance on objects with internal degrees of freedom. PLoS Comput Biol 5:e1000419.
Nashed JY, Crevecoeur F, Scott SH (2012) Influence of the behavioral goal and environmental obstacles on rapid feedback responses. J Neurophysiol 108:999–1009.
Nashed JY, Crevecoeur F, Scott SH (2014) Rapid online selection between multiple motor plans. J Neurosci 34:1769–1780.
Oldfield R (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113.
Oostwoud Wijdenes L, Brenner E, Smeets JBJ (2011) Fast and fine-tuned corrections when the target of a hand movement is displaced. Exp Brain Res 214:453–462.
Prablanc C, Martin O (1992) Automatic control during hand reaching at undetected two-dimensional target displacements. J Neurophysiol 67:455–469.
Pruszynski JA, Scott SH (2012) Optimal feedback control and the long-latency stretch response. Exp Brain Res 218:341–359.
Pruszynski JA, Kurtzer I, Scott SH (2008) Rapid motor responses are appropriately tuned to the metrics of a visuospatial task. J Neurophysiol 100:224–238.
Qian N, Jiang Y, Jiang ZP, Mazzoni P (2013) Movement duration, Fitts's law, and an infinite-horizon optimal feedback control model for biological motor systems. Neural Comput 25:697–724.
Raftery AE, Kass RE (1995) Bayes factors. J Am Stat Assoc 90:773–795.
Reichenbach A, Franklin DW, Zatka-Haas P, Diedrichsen J (2014) A dedicated binding mechanism for the visual control of movement. Curr Biol 24:780–785.
Reynolds RF, Day BL (2012) Direct visuomotor mapping for fast visually-evoked arm movements. Neuropsychologia 50:3169–3173.
Rigoux L, Guigon E (2012) A model of reward- and effort-based optimal decision making and motor control. PLoS Comput Biol 8: e1002716.
Russo AA, Khajeh R, Bittner SR, Perkins SM, Cunningham JP, Abbott LF, Churchland MM (2019) Neural trajectories in the supplementary motor area and primary motor cortex exhibit distinct geometries, compatible with different classes of computation. bioRxiv 650002. doi: 10.1101/650002.
Sarlegna F, Blouin J, Bresciani JP, Bourdin C, Vercher JL, Gauthier GM (2003) Target and hand position information in the online control of goal-directed arm movements. Exp Brain Res 151:524–535.
Saunders JA, Knill DC (2003) Humans use continuous visual feedback from the hand to control fast reaching movements. Exp Brain Res 152:341–352.
Saunders JA, Knill DC (2004) Visual feedback control of hand movements. J Neurosci 24:3223–3234.
Saunders JA, Knill DC (2005) Humans use continuous visual feedback from the hand to control both the direction and distance of pointing movements. Exp Brain Res 162:458–473.
Scheidt RA, Reinkensmeyer DJ, Conditt MA, Rymer WZ, Mussa-Ivaldi FA (2000) Persistence of motor adaptation during constrained, multi-joint, arm movements. J Neurophysiol 84:853–862.
Schwarz G (1978) Estimating the dimension of a model. Ann Statist 6:461–464.
Shadmehr R, Huang HJ, Ahmed AA (2016) A representation of effort in decision-making and motor control. Curr Biol 26:1929–1934.
Todorov E (2005) Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. Neural Comput 17:1084–1108.
Todorov E, Jordan MI (2002) Optimal feedback control as a theory of motor coordination. Nat Neurosci 5:1226–1235.
Todorov E, Li W (2005). A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems. Proceedings of the 2005, American Control Conference, Portland, OR, USA, 2005, pp 300–306.
Yeo SH, Franklin DW, Wolpert DM (2016) When optimal feedback control is not enough: feedforward strategies are required for optimal control with active sensing. PLoS Comput Biol 12:e1005190.
Zago M, Bosco G, Maffei V, Iosa M, Ivanenko YP, Lacquaniti F (2004) Internal models of target motion: expected dynamics overrides measured kinematics in timing manual interceptions. J Neurophysiol 91:1620–1634.
Zhang Y, Brenner E, Duysens J, Verschueren S, Smeets JB (2018) Postural responses to target jumps and background motion in a fast pointing task. Exp Brain Res 236:1573–1581.

3.5 Study II

Mixed-horizon optimal feedback control as a model of human movement

This study, authored by Justinas Česonis and David W. Franklin was published in *Neurons, Behavior, Data Science and Theory* in September 2021. The study expands on results of Study I, suggesting that infinite-horizon OFC can be used as a source for the expected movement duration, which is a required input for finite-horizon control. The two controllers connected in series, here termed as mixed-horizon OFC are then evaluated against the results of earlier published studies and are able to explain some previously discussed, but not conclusively described phenomena of human visuomotor control. Overall, the results of Study II provide additional support for time-to-target based visuomotor control, as well as provide new powerful methods that can be used in modelling visuomotor control computationally.

Contributions

Justinas Česonis was the primary contributor and lead author in this research. In addition, Justinas Česonis designed, built and fit the computational models, as well as drafted the manuscript. Both authors designed the research, interpreted the model behaviour and revised the final version of the manuscript together.

Abstract

Computational optimal feedback control (OFC) models in the sensorimotor control literature span a vast range of different implementations. Among the popular algorithms, finite-horizon, receding-horizon or infinite-horizon linear-quadratic regulators (LQR) have been broadly used to model human reaching movements. While these different implementations have their unique merits, all three have limitations in simulating the temporal evolution of visuomotor feedback responses. Here we propose a novel approach – a mixed-horizon OFC – by combining the strengths of the traditional finite-horizon and the infinite-horizon controllers to address their individual limitations. Specifically, we use the infinite-horizon OFC to generate durations of the movements, which are then fed into the finite-horizon controller to generate control gains. We then demonstrate the stability of our model by performing extensive sensitivity analysis of both re-optimisation and different cost functions. Finally, we use our model to provide a fresh look to previously published studies by reinforcing the previous results, providing alternative explanations to previous studies, or generating new predictive results for prior experiments.

ORIGINAL ARTICLE
Journal Section

Mixed-horizon optimal feedback control as a model of human movement

Justinas Česonis[1] | **David W. Franklin**[1,2,3]

[1]Neuromuscular Diagnostics, Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany

[2]Munich Institute of Robotics and Machine Intelligence (MIRMI), Technical University of Munich, Germany

[3]Munich Data Science Institute (MDSI), Technical University of Munich, Munich, Germany

Correspondence
Department of Sport and Health Sciences, Technical University of Munich, Munich, 80992, Germany
Email: david.franklin@tum.de

Funding information

Computational optimal feedback control (OFC) models in the sensorimotor control literature span a vast range of different implementations. Among the popular algorithms, finite-horizon, receding-horizon or infinite-horizon linear-quadratic regulators (LQR) have been broadly used to model human reaching movements. While these different implementations have their unique merits, all three have limitations in simulating the temporal evolution of visuomotor feedback responses. Here we propose a novel approach – a mixed-horizon OFC – by combining the strengths of the traditional finite-horizon and the infinite-horizon controllers to address their individual limitations. Specifically, we use the infinite-horizon OFC to generate durations of the movements, which are then fed into the finite-horizon controller to generate control gains. We then demonstrate the stability of our model by performing extensive sensitivity analysis of both re-optimisation and different cost functions. Finally, we use our model to provide a fresh look to previously published studies by reinforcing the previous results, providing alternative explanations to previous studies, or generating new predictive results for prior experiments.

KEYWORDS
mixed-horizon, motor control, optimal feedback control, visuomotor responses, temporal evolution, feedback gains

1 | INTRODUCTION

Computational modelling has driven our understanding of human sensorimotor control by supplementing experimental results and motivating new hypotheses [1–4]. In particular, inspiration from control engineering theory has recently contributed to new ideas of how humans plan and execute movements [5]. For example, robust control inspired models have been considered in order to guarantee stability in the presence of noise [6–10]. Similarly, optimality principles have been proposed to explain human movement and solve issues of redundancy through trade-offs between different elements of the cost function, for example task goals and energy consumption [11–14]. Overall these computational approaches have been very successful at reproducing and explaining human-like behaviours [15–18]. Among numerous studies, different optimal control algorithms are applied to seemingly similar experimental paradigms. However, there are subtle differences in implementations of different optimal control paradigms that could result in meaningful behavioural differences, so the motivation of using any one specific algorithm is not always clear.

The majority of the control algorithms used for simulating optimal feedback control strategies rely on the linear-quadratic regulator (LQR) or linear-quadratic Gaussian (LQG) techniques. However, distinctions between the types of LQG or LQR implementations, such as the infinite, finite or receding time horizon, are still underdiscussed, which is particularly important as these different implementations result in different simulated behaviours. For example, in goal directed reaching movements, the finite-horizon control will accurately predict human-like dynamics such as feedback responses to a perturbation [13, 19], as well as some kinematics, like undershooting the target. However, in many experimental studies, where participants are required to finish each trial within the target, this undershooting also becomes an inconsistency, and a model limitation. In contrast, similar simulations with infinite or receding horizon models will produce appropriate kinematics [20, 21], but will fail to accurately model the feedback responses [19]. Here, we propose that in order to simulate more realistic movements using these paradigms we should utilise the strengths of multiple different implementations, rather than weighing the pros and cons of each algorithm to minimise the drawbacks.

Previous research has shown that humans plan and execute movements in separate steps, which likely occur in different brain areas [22, 23]. As such, it is not unreasonable to assume that these steps could be modelled by different algorithmic implementations. Particularly, here we propose a new, mixed-horizon approach in modelling movement planning and execution where the planning stage is represented by an infinite-horizon optimal feedback controller, and the execution stage is represented by a finite-horizon OFC, with the infinite-horizon controller providing movement durations to the finite-horizon controller. While we often consider planning as occurring prior to the movement initiation, here the mixed-horizon model uses the infinite-horizon controller to re-plan the movement duration (or time-to-target) after any perturbation throughout the movement. That is, the planning and execution processes continue throughout the entire movement, allowing the model to respond to any unseen or unpredicted perturbations with appropriate human-like changes. As a result, the combined system allows us to benefit from the individual strengths of the two controllers while addressing each of their limitations.

2 | MATERIALS AND METHODS

2.1 | Mixed-horizon optimal feedback controller

In this article we propose a new optimal feedback control (OFC) framework for modelling human reaching movements, called the mixed-horizon OFC. We have termed this framework as mixed-horizon, as it combines the features of finite-horizon control [12, 24], and infinite-horizon control [20, 25]. Specifically, we have previously shown [19] that even though the infinite-horizon controllers produce reaching movements of an appropriate duration even after the movement is visually perturbed, the corrections to these perturbations do not vary in intensity in the same way as they do in human movements [19, 26–29]. On the other hand, the finite-horizon controller produces feedback responses with variable intensity that depend on the time-to-target [19], consistent with the data of human participants, but inherently requiring movement duration as an input variable. By combining the two frameworks into the single mixed-horizon controller we can overcome the limitations of each individual model and generate a more human-like control response. For completeness, we present two different types of models: a simplistic, separable model, implemented as a linear-quadratic regulator (LQR) and assuming a perfect observer, and a more advanced non-separable model implemented as iterative linear-quadratic Gaussian (iLQG) and with signal-dependent noise present.

A particular limitation of finite-horizon implementations for the modelling of perturbed goal directed movements is that they inherently require movement duration (or time-to-target) as an input variable. However, in most real-life and experimental cases this movement duration is not predefined. While it is possible to estimate this duration from the data, or set experimentally for non-perturbed movements, we have previously demonstrated that human participants non-trivially extend their movement times post-perturbation if the goal of the task is to reach the target [19]. Thus, for every perturbed movement, unless all movement variables (time, distance, perturbation onset, perturbation magnitude) perfectly match, the movement duration would need to be separately estimated in order to accurately apply the finite-horizon OFC. For tasks where model fitting is of interest such limitations may be addressed by individually assessing different types of movements within the task, however such implementation would still not generalise to unseen perturbations.

The limitation of the required input duration for the finite-horizon models can technically be addressed in a multitude of different ways: arbitrary choice, use-dependence, temporal discounting of reward [14], feed-forward learning, or feedback control (particularly infinite [20, 25] or receding horizon [21]). However, while there could definitely be feedforward effects in non-perturbed baseline movements, previous research indicates that feedback processes play a significant role in setting the movement duration post-perturbation [19, 30], as this duration is adjusted based on perturbation onset, type or magnitude. On the other hand, both infinite and receding horizon OFC implementations have been shown to reliably predict movement durations for both non-perturbed and perturbed movements, independent of the perturbation onset, and without changing control parameters. Therefore, there exists an infinite (or receding) horizon OFC with a fixed set of control gains, that could immediately and reliably compute the movement duration for the non-perturbed movement, and in case of a perturbation at any point during this movement – immediately recompute an appropriate, extended movement duration. In turn this duration can then be used to adjust the finite-horizon control policies post-perturbation. Finally, while we previously have shown that the infinite-horizon OFC and receding-horizon OFC can both predict the post-perturbation movement durations equally well, for this paper we utilised the infinite-horizon architecture due to fewer model parameters (i.e. no required length of horizon as an additional input).

2.2 | Experimental data

The focus of this article is on building the computational model using the OFC framework. However, to illustrate the usability of our model, we apply it to model the results of previous behavioural studies. Specifically, we model the findings of goal directed reaching movement studies with cursor and/or target perturbations [19, 28, 29]. We chose these particular studies, as their experimental findings show the regulation of feedback responses to cursor or target perturbations that traditional OFC models could not replicate. Moreover these studies demonstrate complex modulation of feedback gains across a variety of time points, perturbation magnitudes, and conditions that we can use to test our models.

2.3 | State space representation

For all models we used the same state space representation. The hand was modelled as a point mass with $m = 2$ kg. The intrinsic muscle damping was modelled as viscosity $b = 10$ Ns/m (consistent with [13] for movements of comparable speed). This point mass was controlled in a two dimensional x-y plane by two orthogonal force actuators that simulated muscles. These actuators were regulated by a control signal u_t via a first-order low-pass filter with a time constant $\tau = 0.06$ s. The generic state-space representation at time t, used to simulate the system, could be written as

$$x_{t+1} = Ax_t + Bu_t + \epsilon_t BCu_t + \xi_t, \tag{1}$$

where A is a state transition matrix (in discrete time, here only shown for one spatial dimension):

$$A = \begin{bmatrix} 1 & \delta t & 0 \\ 0 & 1 - b\delta t/m & \delta t/m \\ 0 & 0 & 1 - \delta t/\tau \end{bmatrix}$$

and B is a control matrix (in discrete time, here only shown for one spatial dimension):

$$B = \begin{bmatrix} 0 \\ \delta t/\tau \\ 0 \end{bmatrix}$$

ξ_t represents additive control noise, which for our simulation purposes was always set to zero. $\epsilon_t C$ represents signal-dependent multiplicative noise where ϵ_t is sampled at time t from zero mean and unit variance Gaussian noise, and C is a 2×2 scaling matrix defining the magnitude of this noise. The numeric values of C were adjusted to simulate low, medium or high noise levels for non-separable models, or set to zero for the separable model. State x_t exists in the Cartesian plane and consists of position $\vec{p}$, velocity $\vec{v}$ and force $\vec{f}$ (two dimensions each). For our model implementations in discrete time we used the sampling rate $\delta t = 0.01$ s

The state of the plant that is used for control, x_t, is not directly observable, but has to be estimated from the system's output, y_t via the output equation

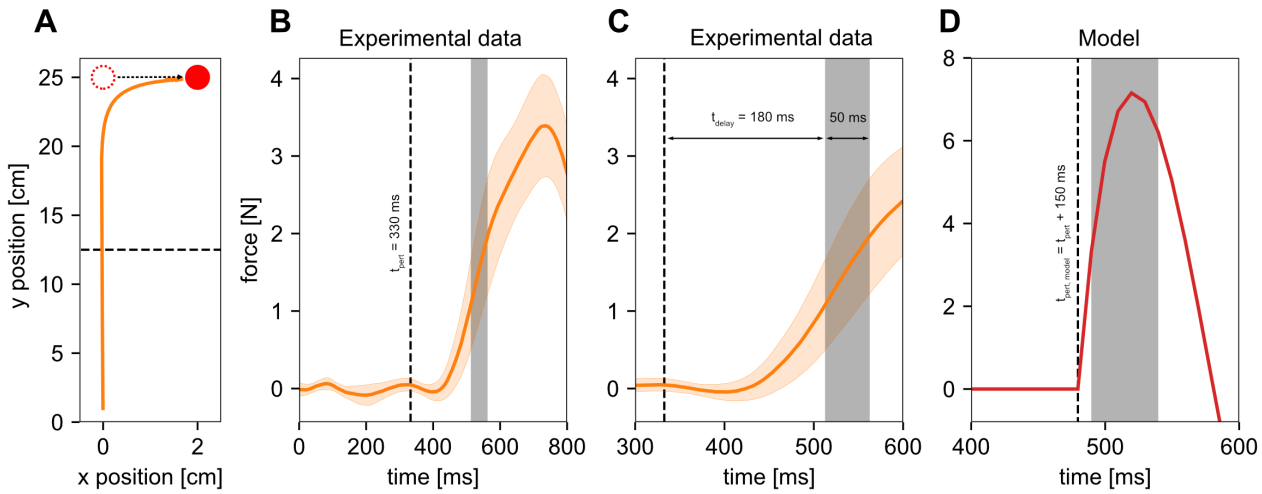


FIGURE 1 An example of visuomotor feedback responses and intensities. **A.** A typical movement trajectory to the visual perturbation of the target, perpendicular to the direction of movement. The perturbation is induced as a target jump from its original position (dashed red circle) to the new position (solid red circle) when the movement trajectory passes the perturbation onset location. An example here shows perturbation induced at the mid-distance along the movement. Note that x and y axes are not shown to scale for better visibility, and participants do not undershoot the target in experimental data. **B.** A typical motor response to the visual perturbation of the target, perpendicular to the direction of movement. No force responses are present in the perturbation direction until the perturbation is induced. An exemplar perturbation (same as in **A**) at t = 330 ms from the beginning of the movement produces a force response in the direction of the perturbation approximately 150 ms later, due to visuomotor delays. **C.** Same data as in **B**, zoomed to time of the perturbation and the response. Force response is averaged over the time window of [180 ms - 230 ms] from the onset of the perturbation (grey shaded area) to determine the response intensity. **D.** Equivalent perturbation induced in the model simulations. Due to no delays in the computational model, the perturbation is induced 150 ms later compared to the human participants. In addition, due to the faster ramp-up of the model responses, the feedback intensity is computed by averaging the response over [10 ms - 60 ms] after the perturbation (gray shaded area).

$$y_t = Hx_t + \sigma_t D. \tag{2}$$

For our models we set $H = I_6 = diag(1, 1, 1, 1, 1, 1)$, meaning that all our hidden state variables (position, velocity, force) are observable by the controller which is consistent with human physiology. σ_t is sampled at time t from zero-mean unit variance Gaussian noise and D is a diagonal matrix representing the intensity of this noise on all output states. In the non-separable models we set $D = 0.015\, diag(1, 1, 1, 1, 10, 10)$, consistent with values in [13]. Even though the same values could be used, in the separable models, for simplicity we set $D = 0$.

2.4 | Estimation of feedback responses

We apply all our models described in this paper to simulate visuomotor feedback responses [26, 31–33] by mimicking the target jump or cursor jump paradigms. In these studies, a visual perturbation of either a target or a cursor during a movement induces a feedback motor response that allows participants to bring the cursor to the target (Figure 1A). When measured in terms of corrective force, this response is typically delayed about 150 ms from this perturbation

onset until the corrective force is produced (Figure 1B). In order to quantify the magnitude of this response, the force is typically averaged over a time window of [180 - 230] ms after the perturbation [19, 26, 28, 29] to produce a visuomotor feedback intensity (Figure 1C). This intensity has previously been shown to vary based on the perturbation onset location [28], magnitude [29], time-to-target [19] or task dependency [26]. Note that in many previous studies the feedback intensity was referred to as the feedback gain, however as we use the term gain to describe the control gains in this article, we opted to use feedback intensities to define these averaged force responses.

In order to estimate the appropriate kinematic behaviour and simulated feedback responses we first trigger the perturbations based on the forward movement of the cursor. Specifically, based on the experimental design of the simulated study there exists a perturbation onset location along the forward movement. Once the cursor crosses this location, the perturbation timer of 150 ms is started to simulate the visuomotor feedback delay. After this delay, the perturbation is triggered by shifting the cursor or the target appropriately to the task design (Figure 1D). Thus, from the external observer's perspective any perturbations in our model are happening 150 ms later than in human experiments. However, considering the delays present in human visuomotor system, the motor responses in human participants would be observed at the similar time as in our model, for the perturbations with the same onset.

In our models in this study we induced the visuomotor delays to the feedback system of our model by simply delaying the perturbation by the fixed duration. This results in the similar behaviour between the model and human participants, as the responses are produced at the matching times. An alternative implementation for the delay could be a fixed intrinsic system delay. However, this implementation would require expanding the state space by 15 times (with our model sampling rate of 100 Hz and 150 ms visuomotor delay), and adding additional computational load. Our preliminary work showed little differences in these two implementations for simulation of these visuomotor feedback responses.

After the perturbation is physically induced we record the average response of the model to the perturbation over the time window of [10 - 60] ms to calculate the feedback intensity. This is equivalent to [160 - 210] ms time window in humans which is somewhat shorter than the conventional [180 - 230] ms time window. However, this is necessary as our models produce faster responses than humans and is also consistent with our previous work [19].

Finally, while there is a distinction between responses to cursor perturbations and target perturbations in human participants [34], in terms of the modelling there is no functional distinction between the target or cursor jumps, as our models only account for the difference vector between them to estimate an appropriate control signal. Therefore, the selection of the perturbed entity (cursor or target) has no effect on our final simulation results.

2.5 | Separable mixed-horizon OFC

Separable mixed-horizon OFC is a simplified model with no additive or multiplicative control noise and perfect state information. This model consists of two major building blocks: the infinite-horizon controller, implemented as an infinite-horizon linear quadratic regulator (LQR), and a finite-horizon LQR controller. First, we use the infinite-horizon LQR to simulate the movement from start to target position in order to estimate the appropriate movement duration for such a movement. This could be considered as a movement planning stage. Second, we use that movement duration as a parameter for the simulation using the finite-horizon controller. As a result, the two parts are connected linearly and can thus be implemented separately and then combined together.

Infinite-horizon LQR is a simplistic optimal controller that generates an optimal control solution to a given system with state cost $x_t^T Q x_t$ and control cost $u_t^T R u_t$. We define the generic form of the performance index J as

$$J = \sum_{t=0}^{\infty} x_t^T Q x_t + u_t^T R u_t = \sum_{t=0}^{\infty} \omega_p (\vec{p}_t - \vec{p^*})^2 + \omega_v ||\vec{v}_t||^2 + \omega_f ||\vec{f}_t||^2 + \omega_r ||u_t||^2 \quad (3)$$

where ω_p, ω_v and ω_f are position, velocity and force state cost parameters, ω_r is the activation cost parameter and $\vec{p^*}$ is a target position. For the finite-horizon controller this performance index is instead

$$J = \sum_{t=0}^{N} x_t^T Q_t x_t + u_t^T R_t u_t = \sum_{t=0}^{N} \omega_{p,t} (\vec{p}_t - \vec{p^*})^2 + \omega_{v,t} ||\vec{v}_t||^2 + \omega_{f,t} ||\vec{f}_t||^2 + \omega_{r,t} ||u_t||^2 \quad (4)$$

where N is the duration of the movement obtained via the infinite-horizon controller. Note here that the control parameters for the finite horizon can be non-stationary and thus be different for every time-point. As a general rule unless stated otherwise, for fast, goal-directed reaching movements we set $Q = 0$ for $t \neq N$, and kept R stationary, consistent with [12, 13].

2.5.1 | Model optimisation

For a given mechanical system, the behaviour of the LQR system is defined by the control parameters ω_p, ω_v, ω_f and ω_r. Thus, instead of selecting the values of control parameters arbitrarily (e.g. from previous literature or by qualitative inspection of final system kinematics or dynamics) we opted to quantitatively optimise our model for the best suited parameters. We executed the optimisation for the infinite-horizon part and the finite-horizon part separately due to their linear relation.

Within our mixed-horizon implementation we use the infinite-horizon part to estimate the durations of individual movements. As a result, we optimise the control parameters of the infinite-horizon part based on the fit between the experimental movement durations, and the movement durations produced by our model in an equivalent paradigm. Specifically, we optimise these parameters on the mean durations of all available perturbed movements simultaneously for the study that we are modelling. Thus, with a single set of optimal control parameters, the infinite-horizon controller produces the movement duration that matches the movement duration of any experimentally perturbed movement, when this matching perturbation is induced. In this way, instead of searching for a new set of control parameters (and thus deriving new controller gains) for every different perturbation, we ensure that for any perturbation our infinite-horizon controller always produces a movement duration that matches that of the human participants, without changing control parameters or control gains. We used the sum of squared-residuals as a goodness of fit measure between the real experimental movement durations and those produced by our models, and the Nelder-Mead algorithm to find the best fit. Furthermore, in order to reduce the chance of finding a local minimum we instantiated the optimisation 10 times with random initial conditions and selected the solution with the lowest sum of square-residuals (SSR).

We use the finite-horizon implementation in order to obtain the kinematics and dynamics for each model. As a result, we aim to fit the kinematics of the finite-horizon simulations to the kinematics of the movement from experimental data. In order to evaluate how well the model kinematics fit to the experimental kinematics we devised a kinematic cost function

$$\begin{aligned}\Gamma = \gamma_1 (v_{peak,req} - v_{peak,model})^2 + \gamma_2 (p_{peak,req} - p_{peak,model})^2 \\ + \gamma_3 (v_{end,req})^2 + \gamma_4 (p_{end,req} - p_{end,model})^2\end{aligned} \tag{5}$$

where the four components are squared-errors of the peak velocity magnitude, forward position of where the peak velocity was achieved, the end-point velocity and the end position (i.e. the distance of the under/overshoot). The parameters γ_{1-4} are the relative weights assigned to each of the components. The $v_{peak,req}$, $p_{peak,req}$ and $p_{end,req}$ are fixed quantities dependent on the experiment of interest and are measured in cm/s and cm respectively. Equation 5 is designed to specifically simulate the experimental conditions that participants in the experiments were instructed to follow. Instead of fitting the whole kinematics, we optimise the kinematics to the similar "instructions" as the participants are given. For example, if participants are instructed to "stop at the target" (meaning have zero velocity and zero error from the target), "try to produce movements that are the right speed" (indirectly instructing peak velocity via "too fast" and "too slow" feedback), or in some specific cases (i.e. experiments in study 1) providing a direct feedback of peak velocity location, our kinematic cost function provides similar constraints to our model via parameters $\gamma_1 - \gamma_4$.

2.5.2 | Model sensitivity to the kinematic cost function

The kinematic cost function influences the desired model behaviour. For example, based on Equation 5, a model where $\gamma_1 >> \gamma_2, \gamma_3, \gamma_4$ would reward kinematics that match the peak velocity requirement, while placing less emphasis on the peak velocity location, final position or final velocity. In contrast, a model with $\gamma_3 >> \gamma_1, \gamma_2, \gamma_4$ would prefer kinematics with no residual velocity while relaxing demands on the other three components. As a result, we conducted a sensitivity analysis on the kinematic cost function to estimate a range of different model behaviours when performing the same task. To do so, we first selected a range of different kinematic cost functions. For the baseline Γ we used relative weights $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ = (4, 4, 0.25, 25). For all other Γ we kept three of the relative weights at their baseline value, while varying the fourth one. We chose $\gamma_1 \in [0.125, 16]$, $\gamma_2 \in [0.125, 16]$, $\gamma_3 \in [0.125, 16]$ and $\gamma_4 \in [0.25, 400]$. For each Γ we then performed a full model optimisation to find the control parameters $\omega_p, \omega_v, \omega_f$ and ω_r that minimise this given Γ. As previously, we conducted a Nelder-Mead optimisation with 10 different instances and selected a solution with the minimum Γ as a single output. Finally, we analysed the resultant model dynamics qualitatively by comparing the simulated force responses (intensities) to the visuomotor feedback intensities in the experimental data.

2.5.3 | Model sensitivity to individual optimisations

We also tested how sensitive our model is to individual optimisation instances. Specifically, for every individual optimisation the best fit behaviour is achieved with a different set of parameters $(\omega_p, \omega_v, \omega_f, \omega_r)$. Thus, we also analysed the relative distribution of these best fit parameters of individual optimisations, and how this change influences the dynamics and kinematics of the control system. We performed optimisation sensitivity analysis 40 times on the baseline kinematic cost function.

2.6 | Non-separable mixed-horizon OFC

The implementation of non-separable mixed-horizon OFC is similar to the separable mixed-horizon OFC. The main difference between the two is the presence of multiplicative control noise in the finite and infinite-horizon blocks. For the infinite-horizon part, instead of using an infinite-horizon LQR implementation we adapted the implementation used by [25], where the control noise was transformed to a loss on control gains and iteratively optimised until convergence. For the finite-horizon part of the model we used the iterative LQG algorithm to obtain the control policy in presence of noise [12, 13].

In presence of multiplicative control noise both infinite and finite-horizon LQG become computationally more expensive due to the iteration until convergence when calculating control and observer gains. As a result, to make the optimisation manageable, for the models with multiplicative noise we initiated every individual optimisation three times instead of the 10 that we used for the separable models, and selected the best value. While this increased the chance that the performance of the model is sub-optimal, our models still behaved generally similar to the separable conditions.

2.6.1 | Noise parameters

In total we tested the optimisation and controller behaviour in three different noise conditions. We introduced these conditions via a noise scaling matrix C (Equation 1). The scaling matrix was defined as

$$C = k \begin{bmatrix} 0.15 & 0.05 \\ -0.05 & 0.15 \end{bmatrix} \tag{6}$$

with $k = 1$ for low noise condition, $k = 3$ for medium noise condition and $k = 5$ for high noise condition. The implementation of C, where noise is proportional not only on the activation of the actuator responsible for moving in a desired direction, but also to the activation of a perpendicular one, induces a coupling between the two Cartesian actuators, consistent with human muscles. The values of k were selected by trial and error: $k = 1$ produces a control behaviour which is generally stable, but different from no-noise behaviour, and $k > 5$ often results in optimisation timing out without finding a minimum. In comparison, a typical human participant from [19] produced motor variability that is comparable to our model at $k = 5$, but ranged between $k = 3$ and $k = 10$ for different participants. As the human participants experience contributions from other variability sources (e.g. planning variability, signal independent noise, measurement noise) in addition to the control-dependent noise, the actual level of control-dependent noise is therefore likely comparable to our selected values. Finally, while this does not strictly mean that $k = 5$ mimics high motor noise, with our implementation we could not consistently test any higher noise values.

2.7 | Specific model implementations

In order to demonstrate the versatility of the mixed-horizon model, we used our model to replicate the experimental behaviour results of previous studies. As different studies have slight differences in implementations, our model had to be adapted for an individual study to comply with the design. While the methods listed above are common across different studies that are modelled in this article, the current section details specific differences for each study.

2.7.1 | Study 1: Česonis & Franklin 2020 [19]

In the original study participants were asked to reach to the target while producing a specific velocity profile: baseline, early peak (movement accelerated early and slowed for the latter portion), or late peak (velocity gradually increased and reached the peak late in the movement) [19]. From the modelling perspective, these three different conditions were implemented in the finite-horizon OFC by making the activation cost ω_r time-dependent:

$$R'(t) = \frac{\exp(p\frac{t+q}{r})}{mean(R')} \tag{7}$$

$$\omega_r = \omega_r R'(t) \tag{8}$$

Here p, q and r are the three parameters governing the temporal shape of the activation cost function. In order to maintain activation cost of an equivalent magnitude between conditions, the R' was normalised to the mean value of 1.

For the finite-horizon part of our model we implemented this scaling to produce the local offset of the peak velocity. However, the same approach can not be applied to the infinite-horizon part of the model. Particularly, the temporal evolution of the ω_r assumes the known movement duration, so that it can then be normalised over the movement duration. If this duration is not known, then this temporal profile becomes undefined. As a result, the infinite-horizon part did not modulate peak velocity locations, but instead we re-fit the control parameters to adapt the movement duration to the experimental condition.

While selecting a new set of controller parameters seems unreasonable for modelling human-like systems, we utilised the model (in)sensitivity to different optimisations to find a parameter space, such that we can switch between conditions within the infinite-horizon part by only changing one of ω_p, ω_v, ω_f and ω_r. Such adaptation is more realistic than changing all parameters – it is reasonable to assume that different required movement kinematics would impart a change in some, but not necessarily all, control parameters.

In order to evaluate the model performance in modelling study 1, we simulated feedback responses equivalent to those in the experiment. In the experiment, as the participants reached towards a target located 25 cm away from the start position, their cursor representation was occasionally perturbed perpendicularly to the movement direction by 2 cm. These perturbations were implemented as cursor jumps either to the left or to the right, and induced as the hand crossed one of five onset locations (P1, P2, P3, P4, P5), evenly spaced along the movement distance (4.2, 8.3, 12.5, 16.7, 20.8 cm). Similarly, our model simulated each of these perturbation conditions – five different perturbation onset locations for each of the three velocity conditions – and generated a total of 15 responses.

2.7.2 | Study 2: Dimitriou et al. 2013 [28]

Experiment 1 in [28] is conceptually very similar to the study described above. As a result, here we focus on simulating the results of experiments 2 and 3. In total, for each experiment we simulated four different perturbed movements. Two of these movements were to a stationary target: cursor-perturbed reaching to a "far" target (25 cm) and cursor-perturbed reaching to a "near" target (17.5 cm), where the cursor was perturbed laterally by 2 cm. In the other two movements, in addition to these cursor perturbations the target was also perturbed in the movement direction: in movements where the starting target was "near", the target was perturbed to "far" and vice-versa. In all conditions of experiment 2 both cursor and target perturbations were induced simultaneously when the participant had moved 15.75 cm from the start position. In experiment 3 the target perturbations were induced at 10.5 cm, instead of 15.75

cm, and the cursor perturbations were still induced at 15.75 cm. In experimental data this resulted in about a 100 ms time difference between the two perturbations [28].

The major difference between study 1 and study 2 is that in study 2, perturbations only happened in channel trials (there were no maintained perturbations as in study 1). As a result, we modified our model implementation to simulate these channel trials by fixing the x-position (position perpendicular to the start-target axis) to 0 when the cursor is not perturbed, and to 2 cm when the cursor is perturbed. We only constrained this position in the simulation of the movement, but not during the planning of the movement (i.e. when the controller is being calculated).

In terms of the baseline movements (no perturbations) study 2 has similar requirements to study 1. As a result, for the finite-horizon part of our model we used the same model parameters as in the study 1 baseline condition. However, we have tuned the parameters of the infinite-horizon part of the model so that durations of simulated movements better fit the durations of the movements in the experimental data. Here for all the 4 types of movement and both experiments we used the same set of parameters.

2.7.3 | Study 3: Franklin et al. 2016 [29]

In this study participants were asked to reach to a target 25 cm away from the start position with some trials being perturbed during the movement. The perturbations could happen either to the target, the cursor, or both, and were always lateral to the movement direction. In terms of magnitude, perturbations could be 0 cm, 1 cm, 2 cm or 3 cm for both target and cursor, and could occur to either left or right. As one of the findings, the authors showed that for isolated perturbations (only cursor or only target perturbations) the response intensity increased with perturbation magnitude in a non-linear, saturating manner (see Figure 1D,E in [29]). We simulated these isolated perturbations in our model in order to first replicate, but also gain further insight into possible mechanisms governing this phenomenon.

In the original study, the lateral cursor or target perturbation always occurred when participants had moved 12.5 cm from the start position. However, in a simulated experiment we induced these perturbations at five equally spaced locations along the movement (one of them being at 12.5 cm). With more simulated perturbations we can analyse the behaviour of the controller not only at the original (experimental) perturbation, but also around it in order to see whether this non-linear modulation arises from the general regulation of the response, or if it can be explained by the time-to-target regulation [19]. In the former case, we would expect to see stronger overall responses with increasing perturbation size, independent of the onset position/time. In the latter case, we would expect to see comparable responses across the conditions as long as the time-to-target matched.

Similar to our model of study 2, we maintained the same finite-horizon parameters as the baseline model of our study 1, as the requirements for movements are similar. However, we again fit the infinite-horizon parameters to better match the movement durations between the model and the data.

2.8 | Code implementation

All code was written in Python 3.7.1 in a Spyder 4.1.3 environment. Array computations and optimisations were executed with `NumPy v1.19.0` and `SciPy v1.5.1` libraries. Where necessary, state-space systems were discretised from continuous time using a zero-order hold method available in `control v0.8.3`.

2.9 | Code availability

The code for all the simulations that are described in the paper, as well as the data obtained via these simulations is freely available online at `https://doi.org/10.6084/m9.figshare.14356346`.

3 | SENSITIVITY ANALYSIS

We performed sensitivity analysis on our model before selecting the model parameters for any given simulation of an experiment. For this analysis we used the mixed-horizon model implementation that we built to simulate our previous results [19]. In order to understand how sensitive the model behaviour is to individual optimisations with the same kinematic cost function (Equation 5), we first performed optimisation sensitivity analysis. This is necessary due to the fact that our optimisation minimum is not a single point within the parameter state-space, but a set of points. Each set of points could potentially result in different model behaviour, as each optimisation run produces a different set of model parameters. However, such a model would not be reliable for our applications, so this has to be tested first. On the other hand, such analysis might allow us to uncover an underlying structure of the optimal parameter space, if multiple parameter sets produce identical optimal behaviours. Second, we performed sensitivity analysis on different kinematic cost functions that we use to optimise the model behaviour. While it is expected that changing the kinematic cost function will affect the model behaviour in general, our purpose with this analysis is to see how sensitive the relative behaviour is between the conditions (e.g. the relative regulation of simulated responses in three different kinematics of our previous work [19]). Specifically, by analysing the cost sensitivity we can test whether particular outcomes of the model simulations result from model fitting, or whether that behaviour arises from the structure within the model. This distinction is important, as the former would indicate that a vast range of behaviours could be fit to a specific model, indicating that the model structure is not meaningful. However, the latter would indicate that the behaviour itself is not an outcome of the parameter choice, but results from the model structure, and thus the behaviour is robust.

3.1 | Optimisation sensitivity

The mixed-horizon OFC consists of two main parts – the infinite-horizon and the finite-horizon OFC working in series. As each part is responsible for different functions within the whole model, namely, determining the movement duration, or, computing the optimal kinematics and dynamics for that duration, the two parts can be optimised separately. In this section we present the optimisation sensitivity analysis for both parts of separable (without control noise) and non-separable (with control noise) models.

3.1.1 | Separable infinite-horizon controller

An infinite-horizon controller in the mixed-horizon framework is used to estimate the total movement duration, which is then passed to the finite-horizon controller to generate movement. In order to fit this controller to produce movement durations, we optimised its state and activation costs $\omega_p, \omega_v, \omega_f$ and ω_r so that the movement durations matched those of human participants in equivalent conditions [19]. Specifically, for each of the three kinematics (baseline, early-peak velocity and late-peak velocity) we simulated a movement where the cursor was perturbed in each of the five perturbation onset locations (once per movement) and evaluated the goodness of fit by calculating the sum of square residuals (SSR) between the generated and actual movement durations. We used this goodness of fit, together with

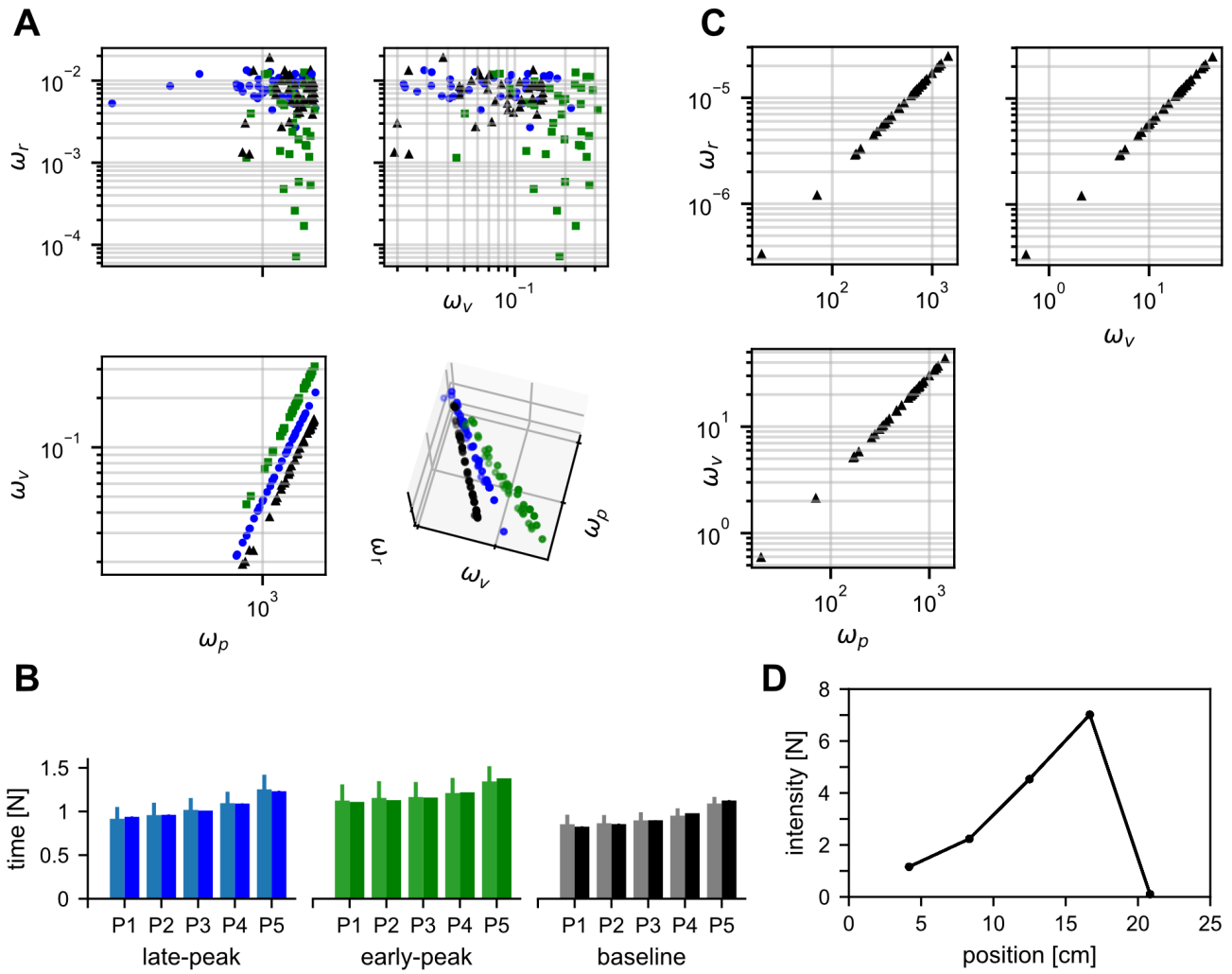


FIGURE 2 Optimisation sensitivity analysis for the separable model. **A.** Distribution of the best LQR infinite-horizon control parameters in the ω_p-ω_v-ω_r parameter space. Each dot represents one set of parameters, while separate plots show the same parameters in a different projection. The fourth control parameter ω_f is not shown, as it is tied to ω_v in a fixed relationship. Black triangles represent the baseline condition, green squares the early-peak velocity condition, blue circles the late-peak velocity condition simulations. Optimal parameters from different conditions do not overlap in $\omega_p - \omega_v$ projection except at the origin, but do in the other two projections. **B.** Movement duration sensitivity to different optimisation instances. Blue (left), green (middle) and black (right) bars represent the movement durations of the late-peak, early-peak and baseline conditions. Lighter shaded bars (left bars for each perturbation) show the mean and 95% confidence intervals (95% CI) of the experimental movement durations for each condition and each of five perturbations (P1-P5). Darker shaded bars (right bars for each perturbation) show the movement durations (and 95% CI) simulated by our separable infinite-horizon model. **C.** Distribution of the best LQR finite-horizon control parameters in the ω_p-ω_v-ω_r parameter space for the baseline condition. Each dot represents one set of parameters, while separate plots show the same parameters in a different projection. The fourth control parameter ω_f is not shown, as it is tied to ω_v in a fixed relationship. Optimal parameters are systematically distributed on a line in the parameter space. Early-peak and late-peak conditions are not shown, as they share the same parameter values (with additional parameters to determine the shape of the velocity profile). **D.** Finite-horizon LQR simulated feedback intensities for the baseline condition. Each set of parameters (one dot) from **C** is used to generate one profile (trace) of simulated feedback intensities by simulating five perturbed movements with different perturbation onset locations. Each of five dots here represents the perturbation onset position, and the corresponding response intensity. All 40 of simulated intensity profiles overlap demonstrating consistency across optimisations.

Nelder-Mead optimisation, to find the ω_p, ω_v, ω_f and ω_r that provide the best fit between model and data durations. We repeated this optimisation 40 times for each condition, generating a set of control parameters per optimisation. These parameters are shown in Figure 2A.

In order to evaluate the sensitivity of the simulated movement durations to different optimisation runs we computed the 95% confidence intervals (95%CI) for these simulated durations and compared them to the equivalent 95%CI for human movements. Despite the wide variations in model parameters for each optimisation, our 95%CI for simulated data is well within the respective intervals for human data, meaning that our simulations are less variable than human movements. Thus, these results show that the model is not sensitive to different instances of the optimisation.

3.1.2 | Separable finite-horizon controller

Similar to the infinite-horizon part of the model, we analysed the sensitivity of the finite-horizon controller to different runs of optimisation. While for the infinite-horizon part we evaluated the goodness of fit during the optimisation via the SSR between movement durations, here we already use the movement duration as the model input. Instead, for the finite-horizon part we minimise the kinematic cost function Γ (Equation 5) in order to produce movements that are kinematically similar to those of the human participants.

The distribution of the optimal state and activation costs over different optimisations is shown in Figure 2C. While every optimisation yielded a different solution numerically, all of these solutions were distributed on the same line in the parameter state-space. In addition, no differences were observed across simulations in either the kinematics or dynamics. Moreover, each simulation produced a consistent variation of the feedback intensities across perturbation locations (Figure 2D). Thus, all together this shows that the separable finite-horizon OFC is extremely robust to different instances of optimisation.

3.1.3 | Non-separable infinite-horizon OFC

Here we analysed the sensitivity of the non-separable infinite-horizon OFC to different optimisation runs. While the non-separable model is similar to the separable model in its implementation, a key difference is the noise in the system which influences model behaviour and leads to a different optimal control solution. Similar to the separable model, the non-separable infinite-horizon OFC is used to compute the required movement duration that would be used by the finite-horizon part. As a result, the sensitivity analysis for the non-separable model is largely similar to the separable model – we used the SSR between the model predicted movement duration and the experimental movement duration as a goodness of fit measure.

In total we simulated the non-separable models with three different levels of noise (Equation 6) $k = 1$, $k = 3$ and $k = 5$. For each of the three noise levels we initially ran the optimisation 120 times per kinematic condition (360 optimisations total per noise level). For each of these optimisations we observed one of three different outcomes:

1. Optimisation converged to a global minimum, with parameters producing consistent and replicable movement durations;
2. Optimisation converged to a local minimum due to random initial conditions, with parameters producing inconsistent movement durations;
3. Optimisation did not converge, or converged with parameters that are out of bounds (e.g. negative costs).

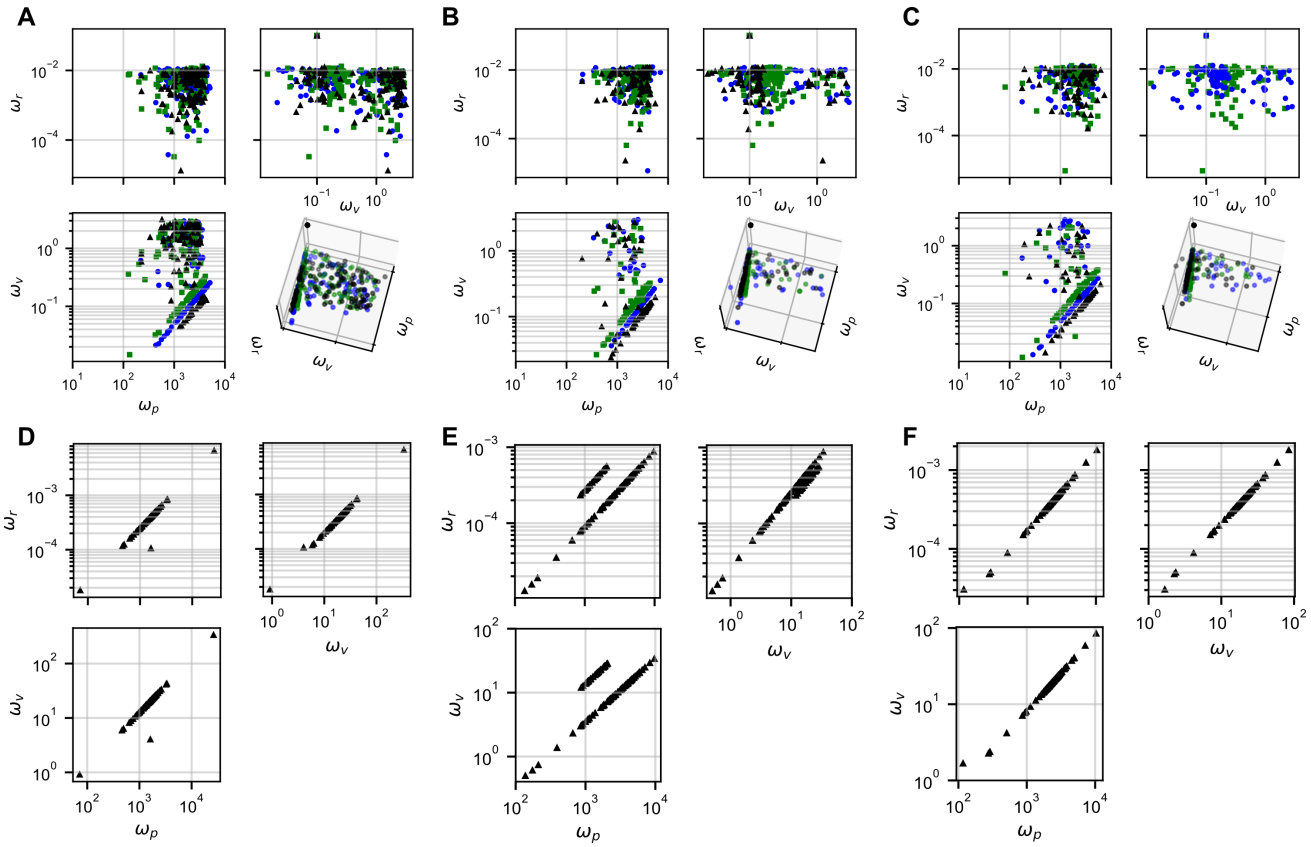


FIGURE 3 Optimisation sensitivity analysis for the non-separable model. **A.** Distribution of the best LQG infinite-horizon control parameters in the ω_p-ω_v-ω_r parameter space for noise level $k = 1$. Each dot represents one set of parameters, while separate plots show the same parameters in a different projection. The fourth control parameter ω_f is not shown, as it is tied to ω_v in a fixed relationship. Black triangles represent the baseline condition, green squares the early-peak velocity condition, blue circles the late-peak velocity condition simulations. Successful optimisations converge onto the parameter plane perpendicular to the $\omega_p - \omega_v$ plane and do not overlap across the conditions. Non-successful optimisation outputs are also shown even though they did not converge, and produce undesirable behaviour. **B, C.** Same as in **A,** but for noise levels $k = 3$ and $k = 5$ respectively. **D.** Distribution of the best LQG finite-horizon control parameters in the ω_p-ω_v-ω_r parameter space for noise level $k = 1$ in the baseline condition. Successful optimisations converged to a line in the parameter space. Unsuccessful optimisations deviate from this line, and are not shown as they are mainly outside of figure boundaries. Early-peak and late-peak conditions are not shown, as they share the same parameter values (with additional parameters to determine the shape of the velocity profile). **E, F.** Same as in **D,** but for noise levels $k = 3$ and $k = 5$ respectively. Interestingly, the optimal parameter space for $k = 3$ results in two lines in the parameter space.

For each noise level we selected only the parameters that converged to a reliable solution (outcome 1). Specifically, with each set of obtained (ω_p, ω_v, ω_f and ω_r) we simulated the movement durations via the infinite-horizon controller three times, and only selected those sets of parameters where the geometric mean of the sums of squared-errors between simulated durations and experimental durations was less than 2.15×10^{-3} s^2. This method helped filter out solutions that belong to option 2, however occasional outliers were still present after such filtering. As a result, we performed an additional outlier removal step using the $1.5IQR$ (interquartile range) rule, applied on mean simulated movement durations. After the outlier removal, there remained 132 parameters sets for $k = 1$, 182 parameter sets for for $k = 3$ and 177 parameter sets for $k = 5$. Finally, in order to keep the number of reliable solutions for $k = 1$ comparable to the other two noise conditions, we ran an additional 60 optimisations per kinematic condition (180 total

optimisations) for this noise condition. This brought the total of remaining parameter sets for $k = 1$ noise condition to 206.

The resulting optimal parameters for all three noise levels are visualised in Figure 3A-C. From all successful simulations we also estimated the mean and 95%CI of movement durations for each noise level (Figure 4A,C,E). Except for $k = 1$ baseline condition, all model simulations show lower or comparable variation than the human data, suggesting that the model is sufficiently insensitive to different optimisations. For the $k = 1$ baseline condition, the model mean is shifted towards high values due to a few outliers. While these outliers generally worsen our results, they can easily be spotted qualitatively and thus removed, not compromising model behaviour when performing simulations.

3.1.4 | Non-separable finite-horizon OFC

As previously, we optimised non-separable finite OFC by evaluating the goodness of fit via a kinematic cost function Γ (Equation 5). For each of the three noise levels $k = 1$, $k = 3$ and $k = 5$ we fit our model to the baseline condition 200 times. As in the infinite-horizon optimisations, some of these optimisations did not converge to a stable solution (e.g. by producing negative parameter values). In total, there were 89, 107 and 114 successful optimisations for $k = 1$, $k = 3$ and $k = 5$ respectively, shown in Figure 3D-F.

We then randomly took 12 sets of our baseline optimisation output parameters, and used them to find best-fit values p, q and r (Equation 7) for generating early-peak and late-peak velocity conditions of the movement. For each set of these 12 baseline parameters we repeated the process 5 times per non-baseline condition to generate a total of 120 sets of (p, q, r) parameters (60 per non-baseline condition). Although these parameters differed numerically, the resulting behaviour was similar (Figure 4B,D,F).

3.1.5 | Mixed-horizon OFC sensitivity to different optimisations

Overall our results of model sensitivity show that the model is reasonably robust to different optimisations. Depending on the noise level in the system the optimisation may time-out without converging to the minimum. However, when the optimisation successfully converged, even though the resultant parameters were numerically different, they produced largely similar behaviour within our model.

3.2 | Cost sensitivity

The optimisation sensitivity analysis presented above is aimed towards testing the reliability of our model. Specifically, it demonstrates that given the actual parameters and cost functions that we use to obtain the final output, our model's behaviour can be reliably replicated, even though the parameter solution is numerically different. However it is also important to test whether the behaviour of our model is a result of model structure and the algorithm itself, or if it is heavily influenced by the optimisation settings (i.e. a cost function). Thus, in addition to optimisation sensitivity we also analyse the model sensitivity to different cost functions. Particularly, as there is no variability in cost function choice for the infinite-horizon part of the model (as the desired output is straightforward), we will focus on the kinematic cost function relative weights $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ (Equation 5), and their effect on the model behaviour.

We performed this sensitivity analysis in the context of the model for our earlier study [19]. Here two alternatives are possible when we optimise the model parameters for different weights $\gamma_1, \gamma_2, \gamma_3, \gamma_4$. On the one hand, modifying the kinematic cost function Γ could impact the kinematics and dynamics of each condition independently, and as a result this would offset the relative regulation of feedback intensities across these conditions. This would imply

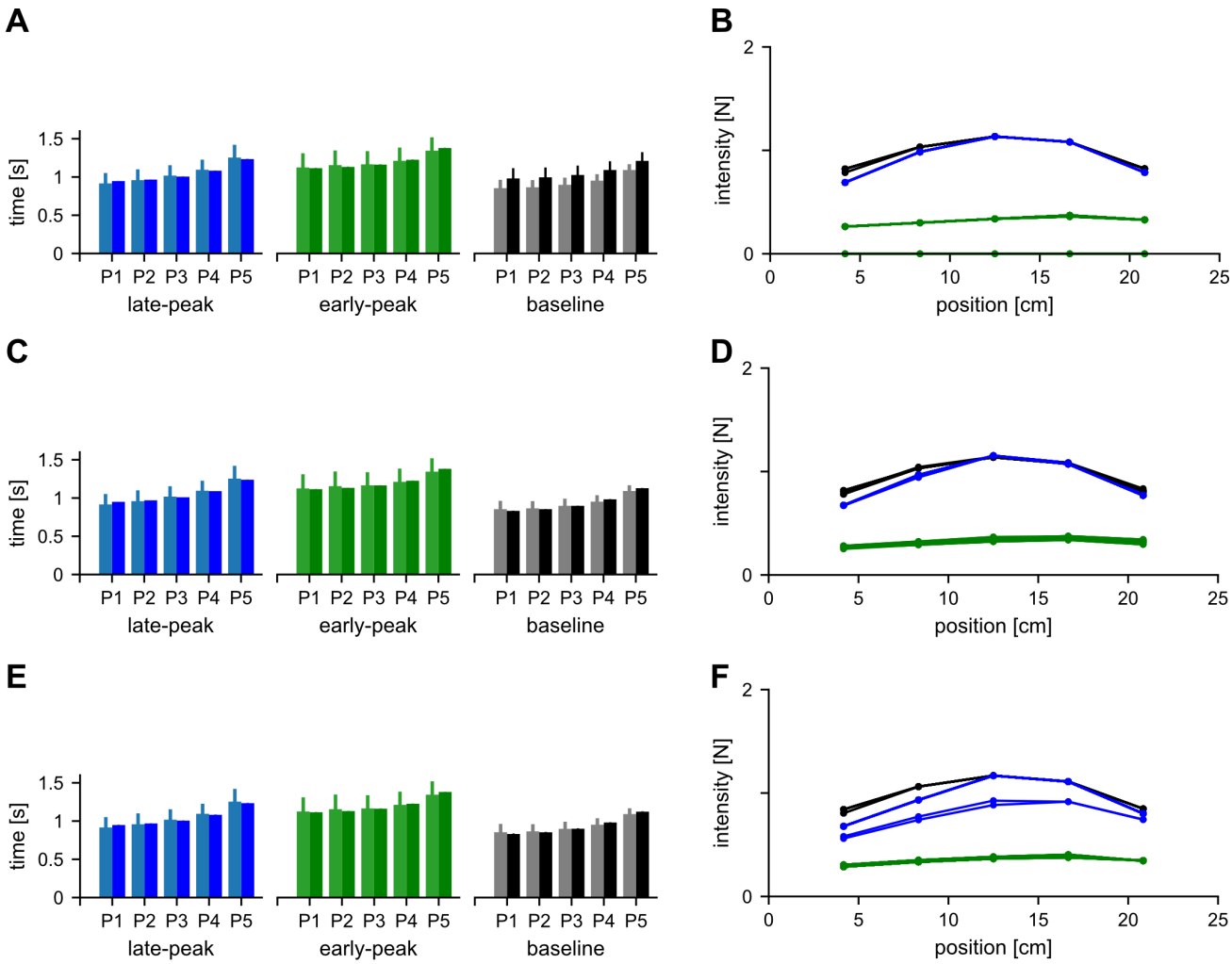


FIGURE 4 Sensitivity of the non-separable mixed-horizon OFC to different optimisation instances. **A.** Blue (left), green (middle) and black (right) bars represent the late-peak, early-peak and baseline movement durations for noise level $k = 1$. Lighter shaded bars (left bars for each perturbation) show the mean and 95% CI of the experimental movement durations for each condition and each of five perturbations (P1-P5). Darker shaded bars (right bars for each perturbation) show the movement durations simulated by our non-separable infinite-horizon model. **B.** Simulated feedback intensity sensitivity to different optimisation instances using the mixed-horizon OFC for noise level $k = 1$. Each line trace represents a feedback intensity profile, generated by simulating five perturbed movements with different perturbation onsets, but with the same control parameters. Twelve simulated profiles are shown for the baseline condition (black lines) along with sixty profiles each for the early peak (green) and the late-peak (blue) conditions. Only convergent optimisation outputs are used. The converging outputs produce consistent behaviour even when the final parameters are different. **C, D.** Same as **A, B.** but noise level $k = 3$. **E, F.** Same as **A, B.** but noise level $k = 5$. In **F.** 58 out of 60 simulation profiles overlap, while two simulations produced lower intensities.

that the observed experimental behaviour replicated by our model is mostly influenced by the cost function used to generate a specific movement. On the other hand, it is also possible that modifying the kinematic cost function will impact the kinematics and dynamics of the model in a way that the relative regulation of feedback intensities remains similar. This result would imply that the observed experimental behaviour is influenced by the underlying model structure, and not by the choice of the control parameters.

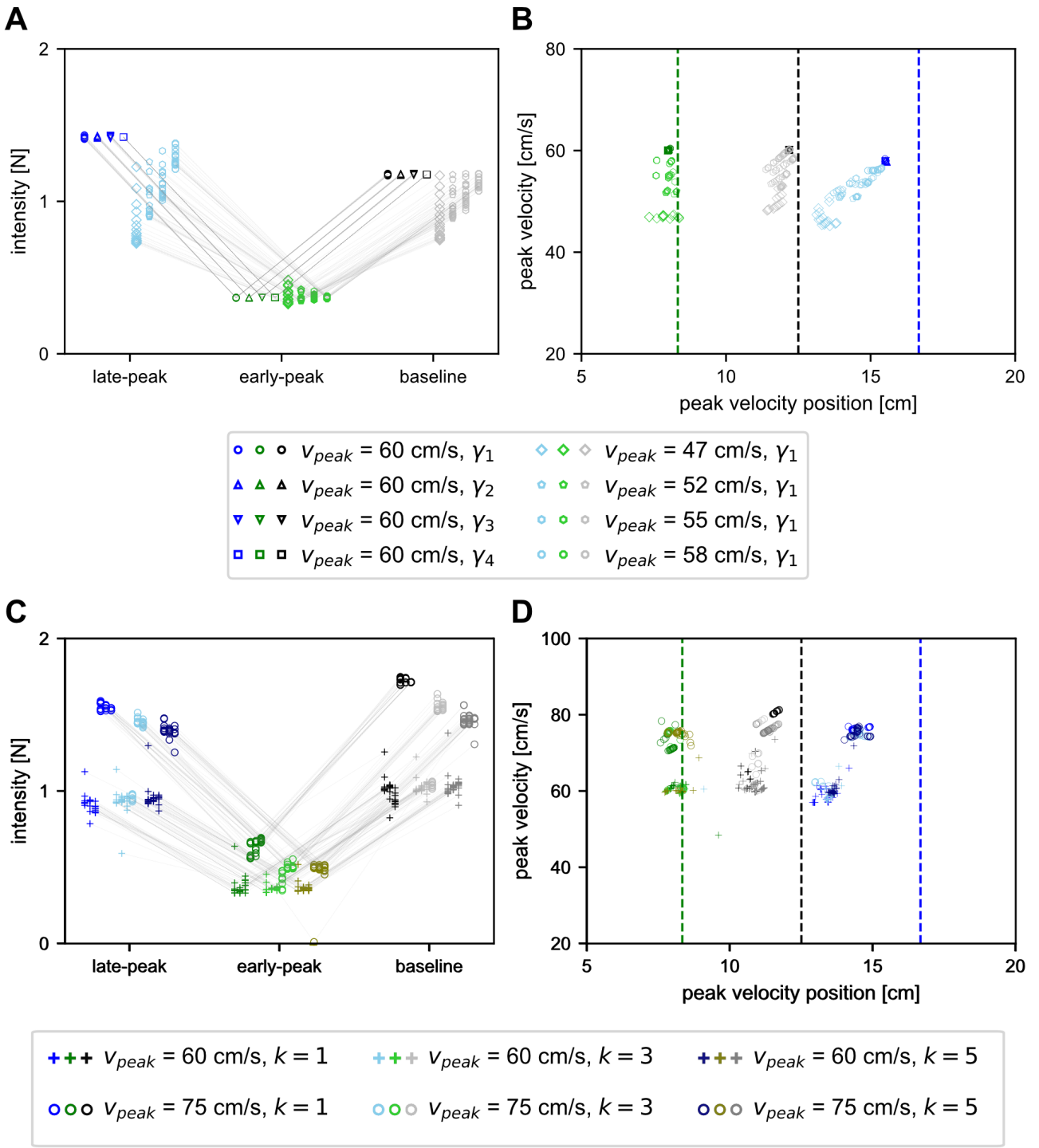


FIGURE 5 Cost sensitivity analysis for the separable and non-separable mixed-horizon OFC models. **A.** Simulated feedback intensity for different kinematic cost functions for the separable model. Every different cost function is represented by three points (blue, green and black), connected via a grey line. Each point represents a mean simulated feedback intensity for the respective simulated condition across five perturbation locations. Different marker styles represent a different family of kinematic cost functions. Overall the relative variation across different experimental conditions is consistent for a variety of cost functions. **B.** Simulated separable model kinematics for different kinematic cost functions represented as a combination of peak velocity location and magnitude. Altering the cost function affects the separation of the kinematic profiles where for some cost functions the baseline and the late-peak condition converge towards similar kinematics. Vertical dashed lines indicate desired location of peak velocity for each condition **C.** Simulated feedback intensities and **D.** simulated kinematics for the non-separable model. Different marker styles and colors represent different families of cost functions and noise levels respectively.

We performed the cost sensitivity analysis by first setting the kinematic cost function Γ to its default value, with $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ = (4, 4, 0.25, 25). In turn, we then varied each of the γs within their boundary range while keeping the other three fixed. We chose the boundary ranges for $\gamma_1 \in [0.125, 16]$, $\gamma_2 \in [0.125, 16]$, $\gamma_3 \in [0.125, 16]$ and $\gamma_4 \in [0.25, 400]$. Furthermore, we also tested different values of $v_{peak,data}$ in the range [47, 75] cm/s.

For the separable model we initially tested 80 different cost functions (20 for each γ_{1-4}, evenly spaced within the boundary range, with $v_{peak,data} = 60$ cm/s). For each simulation we looked at the kinematic features (location and magnitude of peak velocity) and dynamics (qualitative distribution of simulated feedback intensities, as well as mean intensity per condition). The summary of these simulation results is shown in Figure 5A,B. Generally, the results indicate that for separable models, within the range of parameters tested, varying parameters γ_{2-4} has no effect on model behaviour, while changing γ_1 influences the peak velocity magnitudes slightly. However, independent of the cost function, the model maintains the relative regulation of the three kinematic conditions producing consistent kinematics that satisfy the task requirements (Figure 5A).

As we only observed the variance in model behaviour within the modulation of γ_1, we further tested the effect of different target peak velocities in range [47, 60] cm/s and their relative weights (Figure 5AB). By changing the target peak velocity magnitude within the cost function we alter the model behaviour in two ways. First, reducing the peak velocity changes the movement kinematics. While the early-peak velocity condition can be successfully optimised to still produce the target kinematics, both baseline and late-peak conditions deviate further from the target kinematics (Figure 5B). Second, the movement dynamics (feedback intensities) also change. The early-peak condition continues to produce weaker responses than the baseline condition, but the late-peak condition response decreases to the point where it is no longer upregulated in comparison to the baseline. It is important to note, that although we altered the kinematic cost function, we did not alter the movement duration when changing this function – the movement duration is solely produced by the infinite-horizon part of the model. Thus, by changing the target velocity we demonstrate the model behaviour when moving away from its designed optimum point.

We repeated the sensitivity analysis for the non-separable model in the same way as for the separable model, except that we only simulated 10 equally spaced values for each γ_{1-4} due to increased computational complexity (Figure 5CD). In addition, we only analysed two select peak velocities: $v_{peak,data} = 60$ cm/s, and $v_{peak,data} = 75$ cm/s. The 60 cm/s were selected based on our results for the separable model providing the best separation among the kinematic profiles at this velocity (Figure 5B). The 75 cm/s provides even better separation across the kinematics for the non-separable model (Figure 5D). For both peak velocities we performed our simulations at the three different noise levels. Importantly, here again we do not observe any meaningful differences across different noise levels in terms of the relative control gain regulation or kinematic targets (peak velocity position or magnitude, Figure 5C,D). Overall our cost sensitivity analysis shows that, within the range of cost functions that we tested, the model behaviour remains consistent and thus the results are not strongly affected by the choice of the specific kinematic cost function.

4 | RESULTS

Our sensitivity analysis demonstrated that our models can produce kinematics and dynamics with some systematic differences that depend on the parameters and cost functions. The flexibility of our model allows the simulation of many different types of perturbations or movement conditions. To demonstrate and test the generality of our mixed-horizon model, we can simulate specific conditions from previous studies to compare against experimental results. Here we use our model to replicate the experimental behaviour of three previous studies [19, 28, 29]. To do this, we now select a final set of parameters to model our data. Specifically, we select these parameters to fit study 1, and then

apply these same parameters to study 2 and study 3, unless the difference in behaviour requires a specific update (any differences detailed below).

4.1 | Study 1

4.1.1 | Separable mixed-horizon OFC

Final model parameters

In our sensitivity analysis we already demonstrated that for the separable model there exists a set of values (ω_p, ω_v, ω_f, ω_r), with which our mixed-horizon OFC produces movements with durations matching our earlier experiment [19]. However, for each type of kinematics, in order to produce a required movement, three different parameter values are required (ω_p, ω_v and ω_r, as $\omega_f = \omega_v{}^*10$). From the human motor control perspective this becomes challenging – optimising for, and learning a new set of control parameters for each new movement would be computationally expensive. On the other hand, it could be faster to produce a novel movement if a single hyper-parameter could be used to switch between these different kinematics.

Figure 2 demonstrates the relationship of the optimum parameter spaces across all three different kinematics for the infinite-horizon part of the model. As parameters for each condition are distributed in a plane, such that its projection to the ω_p-ω_v plane is a line, we can select parameters ω_p and ω_r, common across all three conditions, and find a value of ω_v, unique for each condition. Indeed, by arbitrarily selecting (ω_p, ω_r) = (1500, 0.006) and performing optimisation on ω_v for each condition we found ω_v = 0.106, ω_v = 0.0713 and ω_v = 0.0501 for the early peak, late peak and baseline kinematics respectively. With these control parameters, our model produced movement durations that fit within the 95% confidence intervals of the experimental durations, showing that infinite-horizon OFC could be used to estimate the duration of the movement for the finite-horizon control (Figure 6A).

In order to now select final model parameters for the finite-horizon part of the model, we first need to select the values for the kinematic cost function Γ (Equation 5). Figure 5 represents the summary of resultant kinematics and relative regulation of dynamics for each Γ tested. In order to best represent the experimental design and requirements for participants, we selected a kinematic function that maximises the separation of the three velocity profiles (i.e. the peak velocity locations are closest to their experimental requirements). As a result, the best parameters of the kinematic function suitable for simulation of the experimental data are ($\gamma_1, \gamma_2, \gamma_3, \gamma_4$) = (4, 4, 0.25, 25), with the $v_{peak,data}$ = 60 cm/s.

Unlike the infinite-horizon part, the finite-horizon part of the model does not require a different set of parameters to produce each of the three different kinematics. Instead, the early-peak and late-peak conditions are generated with the same base parameters ω_p, ω_v, and ω_f. The ω_r is modulated according to Equation 8 to produce the skew in the velocity profile over time, where its mean over the movement duration is identical across the three conditions. In terms of the optimisation to find the parameters for the two conditions, we fixed ω_p, ω_v, ω_f and ω_r to their baseline values and found the p, q and r (Equation 7) that produced closest to desired kinematics. As a result, for the baseline condition we used (ω_p, ω_v, ω_f, ω_r) = (329.5, 9.859, 98.59, 5.568×10^{-6}). For early peak condition, in addition to the baseline parameters we used (p, q, r) = (52.72, 50.37, 741.2), and for late peak condition we used (p, q, r) = (-13.24, 21.08, 347.5).

Model behaviour

Here we propose the mixed-horizon OFC implementation as an extension of the finite-horizon OFC and infinite-horizon OFC when simulating perturbed movements. The goal of this mixed-horizon implementation is to address

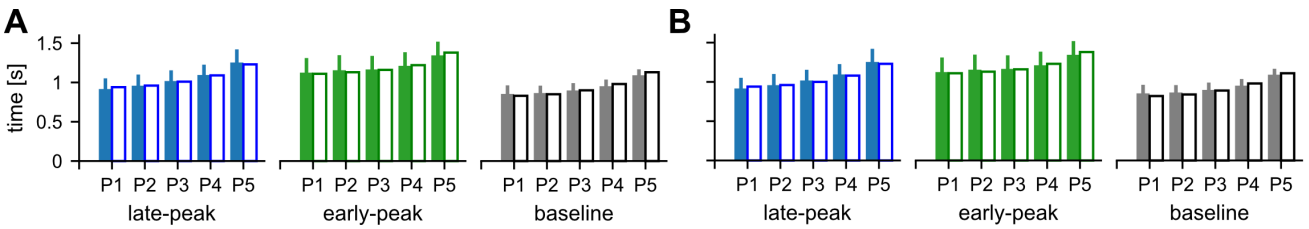


FIGURE 6 Movement duration comparison between the data (solid bars) and the final mixed-horizon model (white bars). Error bars on the data show 95% CI. No error bars are shown for the model results, as the durations result from a single simulation with final model parameters. P1-P5 indicate movements with different perturbation onsets. **A.** Comparison for the separable model. **B.** Comparison for the non-separable model.

the limitations that the fixed horizon OFC encounters when modelling such movements. First, we compared the movement trajectories between the finite-horizon OFC controlled movement and the mixed-horizon OFC controlled movement (Figure 7A). We simulated movements where the target was perturbed laterally at one of five different distances along the movement, requiring a correction to reach the target. For the finite-horizon OFC generated trajectories we observe the stereotypical undershooting for later perturbations, arising from the fact that the perturbation occurs near the end of the planned movement. As a result, it is more optimal for the controller to pay an end-point penalty, than to produce a vigorous and effortful movement in the short remaining time. On the other hand, the mixed-horizon controller produces movements that always converge to the target, consistent with the task requirements and human behaviour. Note that the infinite-horizon controlled movements also always converge to the target due to the nature of the controller and thus are not shown in Figure 7A.

In terms of the model dynamics, the mixed-horizon OFC model qualitatively replicated the feedback intensity profiles of human participants. In the original study, human participants regulated the intensity of visuomotor feedback responses to lateral cursor jumps, both within the same movement, as well as across different movement kinematics. Our mixed-horizon OFC model shows similar behaviour. When compared to the infinite or finite-horizon models, our mixed-horizon model shows a visible improvement (Figure 7B). In addition, while the mixed-horizon model shows similar qualitative behaviour as the time-to-target model (Figure 7C in [19]), it requires fewer inputs and is thus more useful when modelling novel behaviours.

Finally, we compared the separable mixed-horizon model to infinite-horizon and finite-horizon models in their ability to modulate the velocity across three different experimental conditions (Figure 7C). In terms of peak velocity separations, both the finite-horizon and mixed-horizon models produced the required modulations in the shape of the velocity profiles. However, the infinite-horizon OFC could not produce such modulation. Instead, all three conditions in the infinite-horizon OFC manifested in similar velocity profiles with only differences in peak velocity to produce different movement durations. In addition, the velocity profiles were always positively-skewed. However, even with such skewed profiles of the infinite-horizon OFC as a component, the mixed-horizon model is able to capture the appropriate movement durations, condition dependent velocity profiles, and human-like modulation of feedback gains.

4.1.2 | Non-separable mixed-horizon OFC

The non-separable mixed-horizon OFC model differs from the previously presented separable model in its implementation. The main difference between the two models is that the non-separable model contains the multiplicative noise factor, which is equivalent to control dependent noise in humans. Thus, the control policies that emerge from this model may be different than the ones used when no such noise is present. However, at the operational level both

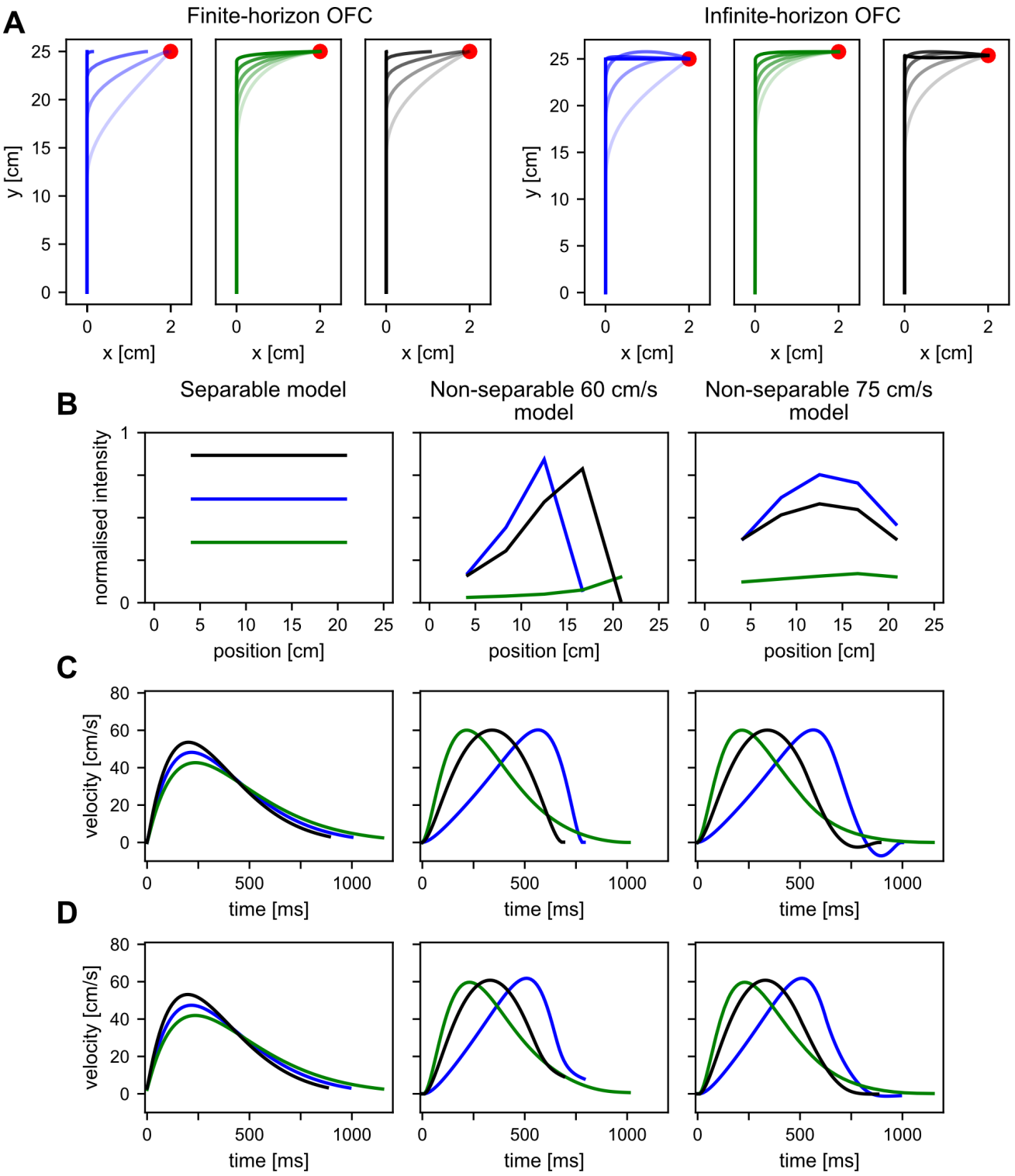


FIGURE 7 Comparison of simulation results across different models. Blue, green and black lines represent late-peak, early-peak and baseline conditions respectively. **A.** Kinematic trajectory comparison between finite-horizon model (left) and an equivalent mixed-horizon model (right). Mixed-horizon model always converges to the target while the finite-horizon model produces undershoots due to the lack of time to complete the movement after the perturbation. Differently shaded lines indicate the five different perturbation onset locations. Three panels in each figure indicate early-peak, late-peak and baseline conditions. **B.** Temporal evolution of feedback intensities for infinite-horizon (left), finite-horizon (middle) and mixed-horizon (right) separable models across different kinematic conditions. Human-like temporal evolution is only observed in the mixed-horizon model. **C.** Velocity profiles in non-perturbed movements across three different kinematic conditions for the three separable models. The separation of different kinematics is not achieved by the infinite-horizon model. **D.** Velocity profiles during non-perturbed movements for non-separable models. Importantly, compared to the separable models, the velocity profiles of non-separable models are skewed towards the beginning of the movement, producing lower separation among different kinematic conditions.

types of models are used in a similar fashion, and hence similar behaviours can be simulated. As a result, here we replicate the same experimentally tested human behaviours as we did with the separable model. Importantly, as we already described the effect of different control-dependent noise levels on the model behaviours, we will only use the control dependent noise with a scaling factor $k = 1$ for further simulations.

Final model parameters

The final parameters for the study 1 model were selected based on our previous sensitivity analysis, using the same criteria as for the separable model. First, based on results shown in Figure 3A we know that for each kinematic condition the optimum parameters are distributed in a plane, similarly to the separable model parameters. Thus, for each kinematic condition we fixed two out of the three controller costs (ω_p, ω_r) = (2000, 0.009) and optimised for the ω_v for each condition so that the resultant movement durations match those in the experimental data. For late-peak, early peak and baseline kinematic conditions we obtained ω_v = 0.0969, ω_v = 0.144 and ω_v = 0.0663 respectively. With the resulting controller parameters our model produced movement durations that fit within the 95% confidence intervals of these durations in experimental data, showing that infinite-horizon OFC could be used to estimate the duration of the movement for the finite-horizon control (Figure 6B).

In order to determine the controller costs for the finite-horizon part of the model we followed the same strategy as for the separable model. Particularly, we first selected a kinematic cost function Γ, and then used it to optimise for the best fit values for controller parameters. As our sensitivity analysis showed that no significant changes in mean behaviour of the model are introduced with the different levels of control-dependent noise, we only analysed the conditions where $k = 1$. However, with the overall presence of control dependent noise we no longer have a clear-cut best Γ to describe our kinematics. Instead, we observe a trade-off between peak velocity location and magnitude for the late-peak velocity condition (Figure 5D). Hence, to fully analyse the model behaviour we will look at the best-fit kinematic cost function for $v_{peak} = 60$ cm/s and $v_{peak} = 75$ cm/s, as the former sufficiently meets the peak velocity criterion while the latter produces the best separation between the peak locations.

First, we chose the values for the kinematic cost function Γ as $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ = (4, 4, 0.25, 25) and the $v_{peak,\ desired} = 60$ cm/s. With these values we then found the baseline parameters $(\omega_p, \omega_v, \omega_f, \omega_r)$ = (1307, 10.7, 107, 2.27×10^{-4}). In addition to the baseline parameters, for the early peak condition we used (p, q, r) = (52.66, 81.05, 772.6), and for the late peak condition we used (p, q, r) = (-13.45, 60.29, 322.3). The best-fit control parameters for $v_{peak,\ desired} = 75$ cm/s were obtained by optimising the kinematic cost function Γ with $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ = (1.44, 4, 0.25, 25), which yielded $(\omega_p, \omega_v, \omega_f, \omega_r)$ = (627.6, 2.243, 22.43, 2.321×10^{-6}) for the baseline condition, and $(p, q, r)_{early}$ = (32.24, 18.61, 520.4) and $(p, q, r)_{late}$ = (-14.83, 37.55, 260.5).

Non-separable model behaviour

In order to evaluate the model behaviour for the non-separable model, we compared the kinematics and feedback intensities of this model with the separable model, and with the experimental feedback intensities [19]. First, the velocity profiles produced by non-separable infinite-horizon OFC still showed similar lack of variation as in the separable case (Figure 7D). In addition, the finite-horizon and mixed-horizon models produced velocity profiles that were positively-skewed compared to the separable models, demonstrating the effect of the control dependent noise. That is, with control-dependent noise and a peak velocity target, the controller chooses to reach this velocity earlier in the movement to be able to have lower velocity and lower noise near the target. While qualitatively this change appears minor (Figure 7CD), quantitatively this results in earlier perturbation onset times (as our model moves faster in the beginning of the movement) and subsequently longer times-to-target for each perturbation. In turn, longer times-to-target for these perturbed movements should result in weaker feedback response intensities, particularly for baseline

and late-peak velocity conditions where the effect of such velocity front-loading is the largest.

Second, we compared just the baseline kinematics of the finite-horizon parts of the separable and non-separable models, as this demonstrates the default, non-perturbed behaviour (Figure 8A). Here the separable model produces the classic bell-shaped velocity profile peaking at the required velocity of 60 cm/s and stopping at the target. In contrast, the 60 cm/s peak velocity requirement for a non-separable model produces a movement that satisfies the required peak velocity, but fails to stop at the target. An attempt to reduce this terminal velocity by increasing γ_3 would only increase the priority of stopping by compromising some of the other requirements – either by shifting the peak velocity (magnitude or location), or by overshooting or undershooting the target. This is illustrated by the kinematics of the 75 cm/s condition, where an increase in the peak velocity allowed the model to stop at the target. The addition of control dependent noise, therefore, changes the default kinematics of the model, particularly affecting the velocity profiles which in turn can affect the dynamics of the system.

Finally, we compared the overall kinematics and feedback intensities across separable, non-separable 60 cm/s and non-separable 75 cm/s models (Figure 8BC). Importantly, all three models were able to reproduce the different kinematic requirements for each condition. In addition, in all three models the response intensities for the early-peak condition were downregulated compared to the other two conditions. However, as expected from the control-dependent noise effects on kinematics, the major difference between the two non-separable models compared with the separable model and experimental data was in the relative regulation of baseline and late-peak response intensities. While the experimental data and separable model showed a clear upregulation for the late-peak intensities, this difference is reduced significantly for the non-separable models. Specifically, for the 60 cm/s model the two profiles are virtually indistinguishable, while for the 75 cm/s model this difference is slight but consistent. Notably, it appears that this absence of regulation in the non-separable 60 cm/s model results from the late-peak condition, as the baseline feedback responses are just slightly lower than for the separable model (Figure 8C).

Overall, the presence of control dependent noise in the mixed-horizon model distorts the experimentally observed regulation of feedback intensities across the three kinematic conditions (Figure 8C, middle), which was successfully simulated without the control dependent noise (Figure 8C, left). The source of this mismatch is due to the regulation of the late-peak velocity condition, which required unusual and unnatural movements and is thus unlikely to be encountered in typical model applications. Finally, for this and the other conditions, our mixed-horizon model produced the variation in temporal evolution closer to that of the human participants than other model candidates (Figures 7B and 8C).

4.2 | Study 2

4.2.1 | Separable mixed-horizon OFC

Final model parameters

Here we modelled the experimental paradigm of [28], experiments 2 and 3, where the reaching target was perturbed in the forward movement direction simultaneously with, or 100 ms before, the cursor perturbations used to measure the feedback intensities. Specifically, we modelled 6 conditions: 2 $\times$ cursor perturbations, 2 $\times$ target perturbations (near to far and far to near), and 2 $\times$ combined target and cursor perturbations. As all conditions appeared in the same experiment in a randomised order, we used the same control parameters across all conditions.

The infinite-horizon part of the model is used initially by the controller in order to produce movements of an appropriate duration. Thus, we optimised the infinite-horizon control parameters for modelling this study by fitting the modelled movement durations to experimentally recorded movement durations for each experiment. For experiment

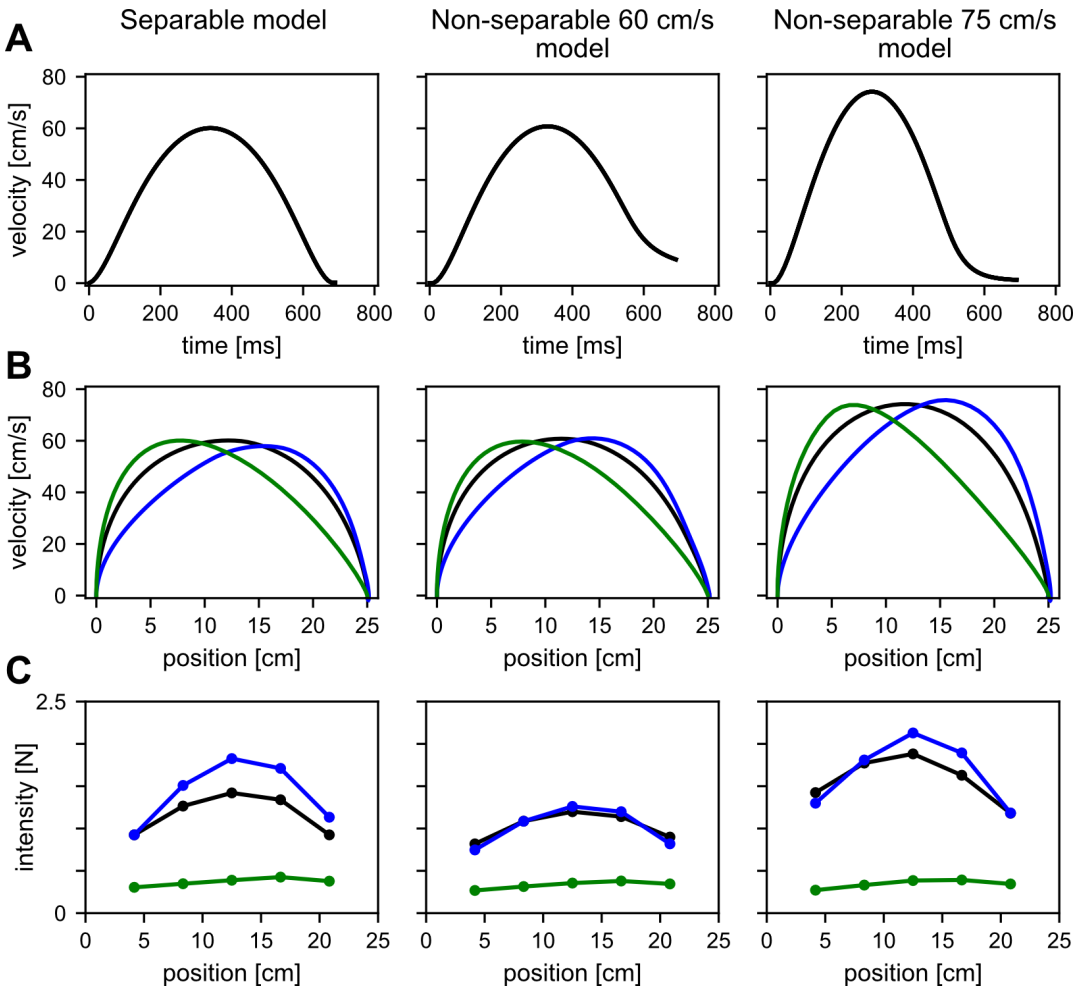


FIGURE 8 Comparison of separable and non-separable model simulations of kinematics and feedback intensities for study 1. Left, middle and right columns demonstrate the results for the separable, non-separable 60 cm/s velocity target, and non-separable 75 cm/s velocity target models. **A.** Forward velocity as a function of time for the finite-horizon part of the mixed-horizon model. The separable model produces a movement that successfully fulfils requirements of both peak velocity and zero final velocity. In the non-separable models, the noise introduces a trade-off between either finishing the movement with non-zero velocity (60 cm/s model) or increasing the peak velocity (75 cm/s model). **B.** Position-velocity dependency. All three models successfully produce the required velocity profiles across the conditions. Note that the 75 cm/s model produces a clearer separation of peak velocities across the conditions compared to the non-separable 60 cm/s model. **C.** Simulated feedback intensities. The separable model successfully replicates experimental results. Non-separable models show limitations in successfully up-regulating the late-peak velocity condition, but capture the down-regulation of the early-peak velocity condition.

2 the obtained values were $(\omega_p, \omega_v, \omega_f, \omega_r)$ = (5448, 0.05285, 0.5285, 0.002460). For experiment 3 the values were $(\omega_p, \omega_v, \omega_f, \omega_r)$ = (5606, 0.04114, 0.4114, 0.0003922). However, for the finite controller we used the identical control parameters to study 1 as the task requirements such as movement distance, speed and perturbation size were comparable.

Model behaviour

We compared the simulated feedback responses of only cursor perturbations to the combined cursor and target perturbations. In experiment 2 of the original study [28], cursor and target perturbations occurred simultaneously.

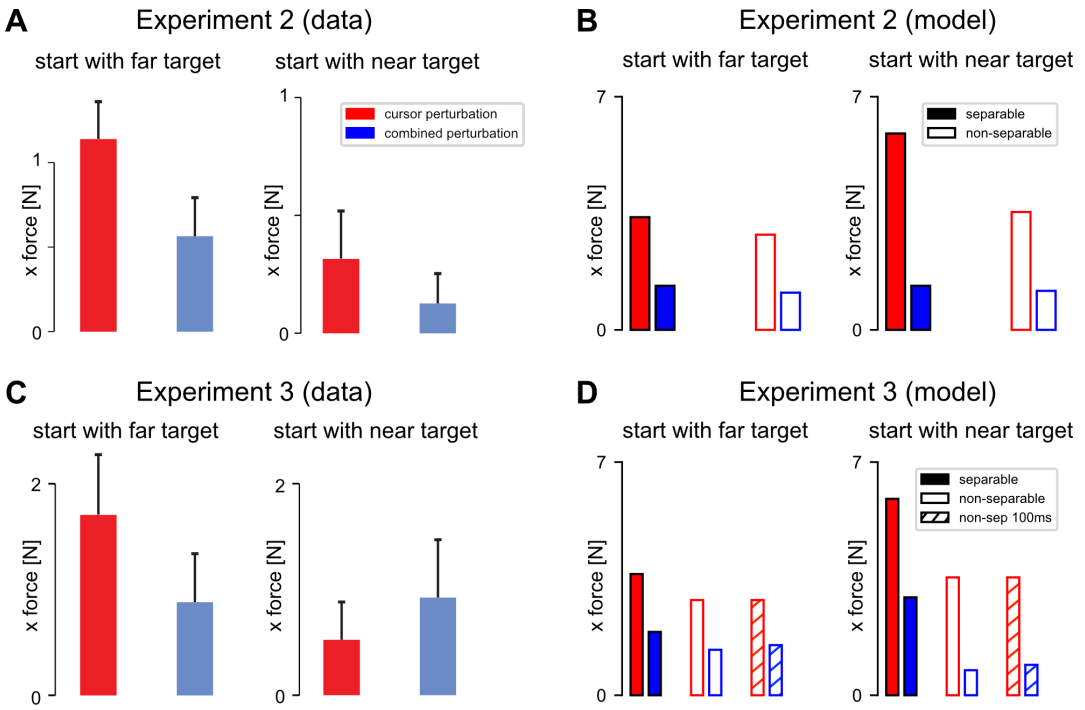


FIGURE 9 Data and model simulation comparisons for study 2. **A,C.** Figures adapted from the original published paper [28] showing the feedback response intensities in experiment 2 (simultaneous perturbations) and 3 (cursor and target perturbations separated by 100 ms) respectively. Red bars show responses to cursor-only perturbations while blue bars show responses to combined perturbations. **B, D.** Respective model simulations with separable and non-separable models. For experiment 3 we also include another non-separable model where the perturbation onset times are adjusted to time-match the perturbations in the data.

Participants produced stronger corrective responses to perturbations of only the cursor, than when both the target and the cursor were perturbed (Figure 9A, Figure 3FG in [28]). Our mixed-horizon model similarly produced weaker corrective responses for simultaneous cursor and target perturbations, compared to just cursor perturbations (Figure 9B). However, while experimental data showed overall stronger responses for conditions where the starting target was "far" compared to when the starting target was "near", our model simulations showed the opposite.

In experiment 3 of the original study, target jumps (when present) were induced earlier in the movement than the cursor jumps. For the conditions with the "far" starting target, the results of this experiment were similar to the results of the equivalent conditions from experiment 2 (Figure 9C, Figure 4FG in [28]). Our model also produced such responses, with corrections to cursor perturbations being stronger than the combined cursor and target perturbations (Figure 9D). However, participants produced stronger responses to the combined perturbations than only to the cursor perturbations when the starting target was "near", which was not consistent with the results of our model.

4.2.2 | Non-separable mixed-horizon OFC

Final model parameters

As previously, in the separable model, we fit the infinite-horizon part of the model to reproduce the movement durations recorded by experimental participants. For both experiment 2 and experiment 3 we set $(\omega_p, \omega_v, \omega_f, \omega_r) = (3585.9, 0.04073, 0.4073, 0.009247)$ for the infinite models. For the finite-horizon part of the model we maintained the parameters we previously found for study 1 60 cm/s condition. The 60 cm/s parameters were chosen over the

75 cm/s parameters because for both settings the baseline condition was producing kinematics consistent with the task requirements – only the late-peak condition was causing discrepancies between models and data. As for both study 2 and study 3 we only use the natural, baseline kinematics, here we opted for the 60 cm/s as it is closer to the kinematics in the data.

Model behaviour

We implemented the non-separable model to simulate experiment 2 and experiment 3 of study 2, similar to the separable model simulations. Qualitatively the non-separable model produced similar results as the separable model (Figure 9BD). Specifically, conditions with combined target and cursor perturbations produced weaker responses to the cursor perturbations, and conditions that started with the near target also produced stronger responses to isolated cursor perturbations. In addition, the non-separable model, similar to the separable model, also failed to replicate the response modulation when the cursor and target were perturbed at different times (starting with near target, experiment 3). Instead, similar to the separable model, the combined perturbation produced weaker responses than the isolated cursor perturbation.

The original study describes combined perturbations in experiment 3 as "separated by 100 ms", however both of these perturbations are induced via the hand position crossing an onset location. Thus, such time dependency is dependent strongly on the shape of the velocity profile. Indeed, in our simulations these perturbations were only separated by 60-70 ms which could have influenced the results. We performed another simulation where instead of perturbing the target once the cursor crosses 10.5 cm distance from the start, we induced this perturbation at 8.5 cm resulting in 100 ms delay between the target and cursor perturbations. Still this modification did not change our results significantly, with combined perturbation still producing lower gains than isolated perturbation for the near target condition (Figure 9D).

4.3 | Study 3

4.3.1 | Separable mixed-horizon OFC

Final model parameters

In order to simulate the control behaviour when subjected to different perturbation sizes we modelled a part of the experimental paradigm of [29]. Particularly, we simulated the conditions where either a cursor, or a target, was perturbed by 1, 2 or 3 cm, at the mid-point in the forward movement. In order for our modelled movement durations to match the experimentally recorded movement durations we again optimised the infinite-horizon part of the model, and we used the same set of infinite-horizon parameters across all conditions. The obtained values were (ω_p, ω_v, ω_f, ω_r) = (2292, 0.05868, 0.5868, 0.006573). However, in terms of control requirements, movements in this study were similar to the ones in study 1 baseline condition, thus we used the previous parameters for the finite-horizon part of the model.

Model behaviour

In the original study, 1, 2, and 3 cm lateral perturbations in the middle of the forward movement resulted in visuomotor feedback intensities (early responses 170-230 ms) that did not scale linearly, but saturated or even reduced for larger perturbation sizes (Figure 10A, Figure 1DE in [29]). We simulated these conditions with our mixed-horizon model. First, our feedback response intensities showed similar regulation – responses to increasing perturbation sizes initially increased and then slightly reduced. In the experiment, the late responses (370-430ms) corresponded roughly to the

peak forces within the correction. Therefore, we used the simulated peak response intensities as a proxy for these late responses. While these responses were stronger in intensity compared to the early simulated responses, they also saturated with increasing perturbation size, similar to the early human responses (Fig. 10B).

It is worth noting that, due to the linearity of our controller, one could reasonably expect the feedback intensities to scale linearly with perturbation size, particularly as the perturbations in the original experiment happen at the same time, speed and forward distance. Therefore, it is important to understand how our linear, mixed-horizon OFC can produce non-linearly increasing responses with linearly increasing stimuli. To illustrate the source of this behaviour, we can use the strength of the mixed-horizon implementation and simulate additional conditions (i.e. perturbations at different onsets) that provide broader context to the available data, but were not experimentally tested. Thus, for each of three perturbation sizes (1, 2 and 3 cm) we additionally simulated four other perturbed movements: 1/6, 1/3, 2/3 and 5/6 of the distance along the movement, similar to study 1. In total, at each of the five perturbation onset locations we simulated 3 perturbations, one of each magnitude (1 cm, 2 cm and 3 cm). As the velocity requirements were identical for non-perturbed movements independent of the upcoming perturbation size, the perturbations at the same onset location therefore happened at the same position, same movement time, and same movement velocity.

Even though the movement kinematics and dynamics, as well as perturbation onset completely matched before the perturbations, introduction of the perturbation resulted in the remaining movement time (time-to-target) being differently adjusted by the controller after the perturbation. Importantly, the resultant times-to-target were always longer for larger perturbations, which is an intuitive result, given that larger corrections were necessary to reach the target (Figure 10C). Also notably, perturbations of a specific magnitude occurring close to the target could extend the movement duration beyond the same perturbation occurring earlier in the movement. For example, in some cases (e.g. perturbations at 2/3 and 5/6 of the movement) a perturbation that occurred closer to the target produced a longer time-to-target than an earlier perturbation with the same magnitude. While this result is less intuitive, it is consistent with our previous experimental results ([19]).

The time-to-target is an important parameter in the LQG and LQR implementation, as the control gains are computed in the backwards-pass calculation, where at every future time point, starting from the end of the movement, the movement cost-to-go and thus the appropriate optimal control gains are computed [12]. Thus, we looked at our simulated feedback intensities for the newly induced perturbations with respect to the time-to-target (Figure 10D). Two important results were observed: first, there is a systematic relation between the time-to-target and the feedback intensities, with higher intensities produced for shorter times-to-target within these responses. Second, due to the extension in movement time post-perturbation, the perturbations happen at different times-to-target, even when they occur at the same location, or same movement time. Due to this difference we also observe a non-monotonic response regulation with perturbation onset location or onset time. Thus, even though, due to the linearity of the controller, the responses that happen at the same time-to-target scale linearly with perturbation size as expected (Figure 10E), the overall regulation of feedback intensities induced at the same location or onset time is non-linear (Figure 10AB).

4.3.2 | Non-separable mixed-horizon OFC

Final model parameters

For the non-separable model of study 3 we again re-fit the infinite-horizon parameters so that the movement durations generated by our model match those in the data. This resulted in the following infinite-horizon parameters: (ω_p, ω_v, ω_f, ω_r) = (2321, 0.05999, 0.5999, 0.001706). For the finite-horizon part of the model we maintained the parameters we previously found for study 1 60 cm/s condition and already used for study 2.

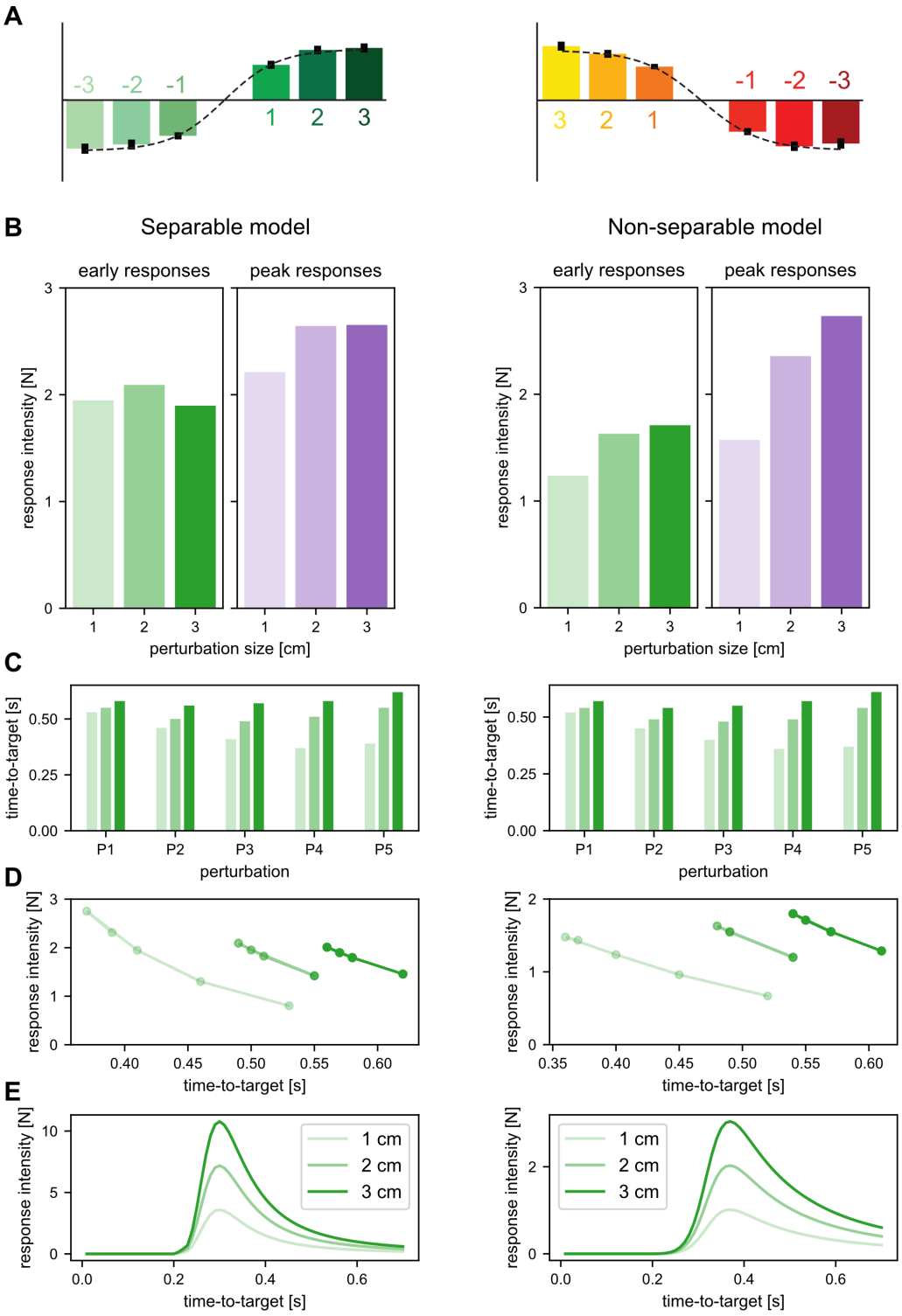


FIGURE 10 Data and model simulations for study 3. **A.** Experimental results. Early feedback intensity to a perturbation at 12.5 cm along the movement in relation to perturbation size in cm. Left plot shows the modulation to target perturbations while the right plot shows for the equivalent cursor perturbations. A non-linear saturation in response intensity is observed. Figure adapted from the original study [29]. **B-E.** Model simulations. Left half shows results for the separable model, right half for the non-separable model. **B.** Simulated intensities for different perturbation sizes. Green bars show early responses, purple bars show peak responses. Similar behaviour to experimental results is achieved by the model simulations. **C.** Simulated time-to-target for perturbations of sizes 1 cm, 2 cm and 3 cm, and with onset at one of the five locations P1-P5. Perturbation onset location and magnitude has a non-monotonic effect on the time-to-target. **D.** Same perturbation responses visualised against time-to-target. Perturbations with matching onset locations produce different times-to-target. **E.** Time-to-target effect on response intensity generalised to all perturbations simulated by the mixed-horizon model. Modelling results show a linear increase in response intensity to perturbations happening at the same time-to-target.

Model behaviour

We did not observe any significant qualitative or quantitative changes in our simulations for study 3 when comparing the non-separable model results with previously simulated separable model results (Figure 10B-E). Thus, introducing control-dependent noise to this particular model did not have a behavioural effect, demonstrating consistency in the simulated behaviour.

5 | DISCUSSION

Here we built and analysed a novel implementation of the optimal feedback controller (OFC) for reaching movements – a mixed-horizon OFC. In order to do this, we not only evaluated the behaviour of the controller when applied to motor control problems, but also analysed the parameters that generate the optimal behaviour. Specifically, our mixed-horizon OFC could successfully replicate kinematics as well as many dynamic features shown by human participants [19, 28, 29], that were not previously possible with either a finite-horizon or infinite-horizon OFC. In addition, our model allowed us to test new hypotheses and alternative explanations to previously published results that could further strengthen original conclusions (study 2) or find new explanations to previously observed interesting features (study 3). Furthermore, by testing the model sensitivity we demonstrated the distinction between fundamental behavioural differences across experimental conditions of study 1, and showed that these differences are not the effect of deliberate model fitting, but maintained across different cost functions.

Prior implementations of either a finite-horizon OFC or infinite-horizon OFC had strong limitations in predicting human behaviour. First, a conventional finite-horizon OFC requires a movement duration as an input. Real world movements on the other hand are almost always free from a specified movement duration. For example, no-one reaches for their coffee cup in the morning with a specified duration of 550 ms. Even if we consider laboratory experiments, most motor control study paradigms are paced not by movement duration, but rather by movement speed. For such paradigms, movement duration has to be non-trivially estimated, or recorded from the data before modelling is even possible. Furthermore, for both visual and physical perturbation paradigms the movement duration is further modulated by movement speed or perturbation timing [19, 35–37]. Our mixed-horizon controller allows us to bypass the requirement of movement duration all together, making this controller more human-like. Second, the requirement of a movement time can be reliably bypassed via infinite-horizon control [20, 25] or receding-horizon control [21] to generate controller movements with kinematics similar to experimental results. However, such simulated movements do not contain any variations in the feedback gains throughout a movement (Figure 7B and [19]), while the presence of such feedback variations is foundational to numerous motor control studies [19, 26–28, 34, 36, 38–40]. Instead, the mixed-horizon OFC is able both select an appropriate movement duration and predict the temporal pattern of feedback responses.

We constructed the mixed-horizon OFC by employing the infinite-horizon controller and finite-horizon controller in series. One of our key motivations for such a combination was that the finite-horizon controller needs movement durations as an input argument. In particular cases, where only data-descriptive modelling is of interest, this is not a major technical issue – simply extracting movement durations from the data for different types of movements, and using it with the finite-horizon OFC would produce identical behaviour to the behaviour of the mixed-horizon model. However, the mixed-horizon control allows us to also generalise the model behaviour beyond only the available data-points to either make predictions about unseen conditions, for novel studies, or to provide some context to the available data. The latter was particularly critical for our simulations of study 3 [29], where we could simulate new experimental predictions even in the absence of experimental data, and therefore movement durations. On the

other hand, adjoining a finite-horizon controller to the infinite-horizon controller produces appropriate modulation of feedback responses throughout the movement that better resemble those of human participants. While this approach may not be required if the only goal is to reproduce kinematics recorded in the data, it nevertheless produces a more human-like response.

Our mixed-horizon controller aims to combine the benefits of finite-horizon and infinite-horizon controllers, here focusing on perturbed, goal-directed movements. The receding-horizon control [21] is another OFC implementation that combines the strengths of the infinite and finite-horizon control, although there are important differences between the two implementations. The receding-horizon OFC uses a similar architecture as the finite-horizon OFC, however it neither needs, nor has any information about the movement duration prior, or even during the movement. Instead, the movement is executed as a series of control signals with the horizon a fixed time away from the current state, until the end of the movement simply "happens" as the hand arrives at the goal. Such implementation is very powerful when modelling long, slow movements, as these movements can be subdivided into movements with via-points, and consequently can accurately replicate the human behaviour in these situations. However, for short, goal-directed movements without via-points the receding-horizon OFC could not replicate some features like visuomotor feedback responses in humans [19]. On the other hand, our mixed-horizon OFC combines the strengths of infinite and finite implementations into a different architecture, which continuously monitors the remaining time horizon for the movement and produces a control signal that matches that of humans. While for long or via-point movements the assumption that the whole time-horizon is predefined in advance may not be realistic, this is critically important to simulate the human-like behaviour in the fast, goal-directed reaching movements.

The combination of two separate control stages in series has previously been discussed in the neuroscience literature [22, 41, 42]. In such a context the process of the infinite-horizon part of the control would be considered motor selection, whereas the finite-horizon part would be considered motor planning. If the finite-horizon part is switched off, the controller is still capable of producing movements to the target, albeit showing less task-dependent modulation. Similarly, human participants are also capable of reaching towards a target (with reduced accuracy) even if no motor planning is allowed [43]. A more recent study has also associated the movement planning stage with trajectory calculation in obstacle avoidance tasks [44]. In that study, the authors demonstrated that a significant portion of reaction time after a go cue is spent calculating the movement trajectory if a specific trajectory was required, while these reaction times were reduced when no trajectory planning was necessary. Here the finite-horizon controller is again consistent with the motor planning stage, as it adds the complexity to the movement features (in our case to movement dynamics). While we have interpreted our two controllers as motor selection and motor planning, the literature also suggests other possible interpretations, e.g. motor planning and execution [23, 42].

While motion planning and execution are generally considered as two processes in series, with planning occurring before execution, it is important to note that some kind of re-planning must occur in order to complete the task if the movement goal suddenly changes [30]. In our mixed-horizon implementation, the perturbation triggers the recalculation of the remaining movement duration (time-to-target) via the infinite-horizon OFC whenever it occurs. This new duration is then used by the finite-horizon OFC to update the remaining control policy. While these processes appear to be in series at any particular time point, both processes continually operate throughout the entire movement: the infinite-horizon is continuously waiting for any possible changes to trigger the re-planning and sending an updated time-to-target (or "urgency" signal [19, 38]) to the finite-horizon controller. Importantly, both controllers still function with the same control weights, thus the only necessary stimulus from the environment to successfully complete the control is the error signal of the perturbation. A similar idea has been previously described by [45], where fast visuomotor responses were explained via the pre-computation of feedback gains for possible perturbations. Here, instead of pre-computing the responses to particular perturbations, the controller maintains fixed time-variable finite-horizon

gains. In turn, the infinite-horizon controller updates the time-to-target in case of any perturbations, which modulates the response of the finite-horizon, and thus overall the mixed-horizon controller. An alternative explanation is that both the infinite and finite controllers are continually re-calculated throughout the entire movement (rather than waiting for a discrete error signal), producing similar updates of the time-to-target or urgency after any perturbation. While this would not change the model predictions in this paper, it is an open question whether the biological sensorimotor control system implements such continuous computation or re-computation only in the case of errors signals.

In this article we were interested not only in building the model and evaluating its dynamics, but also in how the model parameters and costs affect the outcome of the simulation. Particularly, we analysed our mixed-horizon OFC sensitivity to individual optimisations for the same kinematic cost function, as well as sensitivity to different kinematic cost functions 5. Across multiple optimisations to the same cost function (optimisation sensitivity analysis) we observed the structure within controller costs ω_p, ω_v and ω_r (ω_f was fixed to ω_v, consistent with [13]). For the infinite-horizon part, outputs ω_p, ω_v and ω_r of repetitive identical optimisations were distributed on the same plane perpendicular to the plane ω_p-ω_v . As such, instead of optimising for all three parameters simultaneously to obtain best fit controller costs, the problem can be reduced to a one-dimensional optimisation on ω_v for an arbitrary selected pair of (ω_p,ω_r). Similarly, the finite-horizon part of controller outputs to the same repeated optimisation were generally distributed along a single line in the parameter space, which can also be reduced to a one-dimensional optimisation. This is conceptually similar to the idea of structural learning [46, 47], where one meta-parameter could represent a movement along a task structure which is otherwise non-trivial if the entire parameter space needs to be searched. Computationally it can be used to significantly reduce optimisation durations for similar problems by reducing the dimensionality of the problem. That is, new movements may not require optimisation across a huge parameter space, but instead only optimisation along a single dimension which might easily be implemented during learning.

In addition to examining multiple optimisations to the same cost function, we also looked at the outcomes of optimisations with different kinematic cost functions. We analysed this cost sensitivity in the context of our model for study 1 [19] as it involves three different kinematics conditions and their corresponding dynamics. Our mixed-horizon model provided mixed results when simulating the relative regulation of the response intensities between the conditions in our previous work [19]. Specifically, while the separable model could reliably simulate the upregulation and downregulation of the two conditions with changed kinematics, the non-separable model struggled to convincingly simulate the upregulation of the late-peak condition responses (Figure 8). We suggest that this limitation arises, at least partially, from specific task requirements and their interactions with features of the OFC. First, evidence from simulations suggests that contrary to humans, OFC simulates movements that are not exactly bell-shaped [13, 19, 39, 48] but instead have positively skewed asymmetric velocity profiles (peak velocities early in the movement). This feature is particularly emphasised in the infinite-horizon problems [19, 20, 25], as the movement costs are constant throughout the movement. Furthermore, as we previously showed [19], we can not easily modulate the shape of this velocity profile for the infinite-horizon control, which means that our infinite-horizon controller is always projecting a movement with a velocity profile that peaks within the first half of the movement. In turn, even if we fit the infinite-horizon controller to produce appropriate movement durations after perturbations, the difference in profile still introduces differences in perturbation timings, which then influences the responses. This is most strongly emphasised for the late-peak condition, as the separation between the finite-horizon and infinite-horizon kinematics is largest. As a result, we posit that while this is indeed a model limitation, it minimally affects the movements that could be considered natural.

The cost sensitivity analysis of our models allows us to evaluate the general features of model behaviour. That is, we can separate model predictions that are outcomes of the parameter fitting from the predictions that are fundamental to the model (i.e. independent of the cost function or parameters). Our results for study 1 [19] and for sensitivity

analysis (Figure 5) demonstrate consistent relative regulation of feedback responses across the three experimental conditions, with early-peak condition down-regulated and late-peak condition up-regulated from the baseline. Importantly, these results are consistent throughout our vast set of explored kinematic cost functions that result in an even broader set of model parameters. Furthermore, our model has no further assumptions beyond the LQG-governed optimal feedback controller, which is enough to replicate the behaviour of human participants. This suggests that the behaviour observed in the data could be entirely governed by the response of the optimal control policy to the movement requirements, and does not require any additional control components.

Our results for study 2 replicated the main behaviour of experiment 2 from Dimitriou and colleagues [28], showing that combined cursor and target perturbations produce weaker responses than those to an isolated perturbation (Figure 9). We obtained this result for both separable and non-separable models. While the original authors discussed this down-regulation as arising due to the uncertainty of the limb, our models included no assumptions about the certainty and thus such behaviour is a direct outcome of the OFC. In contrast, we were not able to replicate the behavioural results of experiment 3 [28], where the feedback gains were conditionally modified (either up- or down-regulated) when the target perturbation preceded the cursor perturbation. Instead, our model continued to produce similar down-regulation for combined perturbations as in experiment 2. The original authors suggested that the experiment 3 results may have occurred from a recalculation of the control gains during the 100 ms between the target and cursor perturbations. However, our model implementation immediately recalculates optimal control gains after any perturbations by default (including in the model for experiment 2). Here, based on the model-data discrepancies, we presume that instead of just recalculating the control policy after the target is perturbed along the movement direction, human participants could have also updated the kinematic cost function in a non-trivial way that generated such behaviour.

Our simulations of study 3 [29] demonstrate the applicability of the mixed-horizon OFC in situations where data is not currently available. Specifically, in the original study only a single perturbation location was used to measure the visuomotor feedback responses across a range of conditions. One particular observation was a non-linear increase in feedback responses with perturbation size (Figure 10A), which was not explained. Here we examined whether our mixed-horizon optimal controller could explain this non-linear increase in feedback intensities by simulating more data points than experimentally collected. In turn, we demonstrated that such scaling is a combination of a linear gain regulation with perturbation size further modulated by the extension in movement duration, which also depends on perturbation size. It is important to note that we would not have been able to generate such results with only finite, infinite or receding horizon models, either due to the lack of input data (movement durations for not recorded perturbation locations in finite-horizon simulations), or due to model's inability to simulate the temporal evolution of feedback gains (infinite and receding horizon).

Our proposed mixed-horizon model is able to both predict movement durations for untrained conditions, and generate human-like modulation of feedback intensities throughout movements, by combining features of both the infinite and finite-horizon controllers. However, one limitation in combining these two separate models, as currently implemented, is that each model has their individual control parameter sets. While in the ideal case the two parts would share parameters, this is not possible due to the mathematical implementations of infinite and finite-horizon OFC. Finite-horizon OFC models utilise cost functions Q and R that are (or at least can be) variable over time, while infinite-horizon models use stationary costs. As both of these models serve different purposes within the whole mixed-horizon control, such architecture implies that the infinite and finite-horizon models are two separate systems within the overall control, rather than one holistic system. Despite these limitations, the mixed-horizon OFC is able to model a large range of behaviours and paradigms that were not previously possible, as well as make predictions for new studies.

In this paper we have demonstrated a new mixed-horizon approach to optimal feedback control modelling of human behaviour. Specifically, we have contrasted the results of the finite and infinite-horizon simulations. While each implementation is valuable on its own, it also has limitations in modelling specific features of human motor control, such as the temporal evolution of feedback gains or variable movement duration of perturbed movements. By combining the infinite-horizon and finite-horizon controllers into a mixed-horizon controller we were able to better model the results of previously published studies and provide alternative explanations of their results [28], further reinforce the results [19], or model the results that could not previously be simulated [29]. All together our results demonstrate a novel and powerful approach to optimal feedback control models that can more accurately represent behaviours observed in humans.

acknowledgements

We thank Clara Günter and Raz Leib for their comments on this manuscript.

references

[1] Hogan, N. (1984) An organizing principle for a class of voluntary movements. *Journal of Neuroscience*, **4**, 2745–2754.

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

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

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

[5] Scott, S. H. (2004) Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews Neuroscience*, **5**, 532–546.

[6] Franklin, D. W., Burdet, E., Peng Tee, K., Osu, R., Chew, C.-M., Milner, T. E. and Kawato, M. (2008) CNS Learns Stable, Accurate, and Efficient Movements Using a Simple Algorithm. *Journal of Neuroscience*, **28**, 11165–11173.

[7] Crevecoeur, F. and Scott, S. H. (2013) Priors Engaged in Long-Latency Responses to Mechanical Perturbations Suggest a Rapid Update in State Estimation. *PLoS Computational Biology*, **9**.

[8] Ueyama, Y. (2014) Mini-max feedback control as a computational theory of sensorimotor control in the presence of structural uncertainty. *Frontiers in Computational Neuroscience*, **8**, 119.

[9] Crevecoeur, F., Scott, S. H. and Cluff, T. (2019) Robust Control in Human Reaching Movements: A Model-Free Strategy to Compensate for Unpredictable Disturbances. *Journal of Neuroscience*, **39**, 8135.

[10] Bian, T., Wolpert, D. M. and Jiang, Z.-P. (2020) Model-Free Robust Optimal Feedback Mechanisms of Biological Motor Control. *Neural Computation*, **32**, 562–595.

[11] Todorov, E. and Jordan, M. I. (2002) Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, **5**, 1226–1235.

[12] Todorov, E. (2005) Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. *Neural Computation*, **17**, 1084–1108.

[13] Liu, D. and Todorov, E. (2007) Evidence for the Flexible Sensorimotor Strategies Predicted by Optimal Feedback Control. *Journal of Neuroscience*, **27**, 9354–9368.

[14] Shadmehr, R., Huang, H. J. and Ahmed, A. A. (2016) A Representation of Effort in Decision-Making and Motor Control. *Current Biology*, **26**, 1929–1934.

[15] Izawa, J., Rane, T., Donchin, O. and Shadmehr, R. (2008) Motor Adaptation as a Process of Reoptimization. *Journal of Neuroscience*, **28**, 2883–2891.

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

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

[18] Yeo, S.-H., Franklin, D. W. and Wolpert, D. M. (2016) When Optimal Feedback Control Is Not Enough: Feedforward Strategies Are Required for Optimal Control with Active Sensing. *PLOS Computational Biology*, **12**, e1005190.

[19] Česonis, J. and Franklin, D. W. (2020) Time-to-target simplifies optimal control of visuomotor feedback responses. *eNeuro*, **7**, 1–17.

[20] Jiang, Y., Jiang, Z. P. and Qian, N. (2011) Optimal control mechanisms in human arm reaching movements. *Proceedings of the 30th Chinese Control Conference, CCC 2011*, 1377–1382.

[21] Guigon, E., Chafik, O., Jarrassé, N. and Roby-Brami, A. (2019) Experimental and theoretical study of velocity fluctuations during slow movements in humans. *Journal of Neurophysiology*, **121**, 715–727.

[22] Cisek, P. and Kalaska, J. F. (2010) Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience*, **33**, 269–298.

[23] Kodl, J., Ganesh, G. and Burdet, E. (2011) The CNS stochastically selects motor plan utilizing extrinsic and intrinsic representations. *PLoS ONE*, **6**, 1–10.

[24] Todorov, E. and Li, W. (2005) A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems. *Proceedings of the 2005, American Control Conference, 2005.*, 300–306.

[25] Qian, N., Jiang, Y., Jiang, Z.-P., Mazzoni, P. and Forster, D. (2013) Movement duration, Fitts's law, and an infinite-horizon optimal feedback control model for biological motor systems. *Neural Computation*, **25**, 697–724.

[26] Franklin, D. W. and Wolpert, D. M. (2008) Specificity of Reflex Adaptation for Task-Relevant Variability. *Journal of Neuroscience*, **28**, 14165–14175.

[27] Oostwoud Wijdenes, L., Brenner, E. and Smeets, J. B. J. (2011) Fast and fine-tuned corrections when the target of a hand movement is displaced. *Experimental Brain Research*, **214**, 453–462.

[28] Dimitriou, M., Wolpert, D. M. and Franklin, D. W. (2013) The Temporal Evolution of Feedback Gains Rapidly Update to Task Demands. *Journal of Neuroscience*, **33**, 10898–10909.

[29] Franklin, D. W., Reichenbach, A., Franklin, S. and Diedrichsen, J. (2016) Temporal Evolution of Spatial Computations for Visuomotor Control. *Journal of Neuroscience*, **36**, 2329–2341.

[30] Georgopoulos, A. P., Kalaska, J. F. and Massey, J. T. (1981) Spatial trajectories and reaction times of aimed movements: effects of practice, uncertainty, and change in target location. *Journal of Neurophysiology*, **46**, 725–743. PMID: 7288461.

[31] Sarlegna, F., Blouin, J., Bresciani, J. P., Bourdin, C., Vercher, J. L. and Gauthier, G. M. (2003) Target and hand position information in the online control of goal-directed arm movements. *Experimental Brain Research*, **151**, 524–535.

[32] Saunders, J. A. and Knill, D. C. (2003) Humans use continuous visual feedback from the hand to control fast reaching movements. *Experimental Brain Research*, **152**, 341–352.

[33] Brenner, E. and Smeets, J. B. (2003) Fast corrections of movements with a computer mouse. *Spatial Vision*, **16**, 365–376.

[34] Reichenbach, A., Franklin, D. W., Zatka-Haas, P. and Diedrichsen, J. (2014) A dedicated binding mechanism for the visual control of movement. *Current Biology*, **24**, 780–785.

[35] Shadmehr, R. and Mussa-Ivaldi, F. A. (1994) Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience*, **14**, 3208–24.

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

[37] Cross, K. P., Cluff, T., Takei, T. and Scott, S. H. (2019) Visual Feedback Processing of the Limb Involves Two Distinct Phases. *Journal of Neuroscience*, **39**, 6751–6765.

[38] Crevecoeur, F., Kurtzer, I., Bourke, T. and Scott, S. H. (2013) Feedback responses rapidly scale with the urgency to correct for external perturbations. *Journal of Neurophysiology*, **110**, 1323–1332.

[39] Nashed, J. Y., Crevecoeur, F. and Scott, S. H. (2014) Rapid Online Selection between Multiple Motor Plans. *Journal of Neuroscience*, **34**, 1769–1780.

[40] Zhang, Y., Brenner, E., Duysens, J., Verschueren, S. and Smeets, J. B. (2018) Postural responses to target jumps and background motion in a fast pointing task. *Experimental Brain Research*, **236**, 1573–1581.

[41] Wong, A. L., Haith, A. M. and Krakauer, J. W. (2015) Motor planning. *Neuroscientist*, **21**, 385–398.

[42] Gallivan, J. P., Logan, L., Wolpert, D. M. and Flanagan, J. R. (2016) Parallel specification of competing sensorimotor control policies for alternative action options. *Nature Neuroscience*, **19**, 320–326.

[43] Ghez, C., Favilla, M., Ghilardi, M. F., Gordon, J., Bermejo, R. and Pullman, S. (1997) Discrete and continuous planning of hand movements and isometric force trajectories. *Experimental Brain Research*, **115**, 217–233.

[44] Wong, A. L., Goldsmith, J. and Krakauer, J. W. (2016) A motor planning stage represents the shape of upcoming movement trajectories. *Journal of Neurophysiology*, **116**, 296–305.

[45] Franklin, D. W. (2016) Rapid Feedback Responses Arise From Precomputed Gains. *Motor Control*, **20**, 171–176.

[46] Braun, D. A., Aertsen, A., Wolpert, D. M. and Mehring, C. (2009) Motor Task Variation Induces Structural Learning. *Current Biology*, **19**, 352–357.

[47] Braun, D. A., Waldert, S., Aertsen, A., Wolpert, D. M. and Mehring, C. (2010) Structure learning in a sensorimotor association task. *PLoS ONE*, **5**, 3–10.

[48] Berniker, M. and Penny, S. (2019) A normative approach to neuromotor control. *Biological Cybernetics*, **113**, 83–92.

3.6 Study III

Contextual cues are not unique for motor learning: Task-dependant switching of feedback controllers

This study, authored by Justinas Česonis and David W. Franklin has been accepted for publication at PLOS Computational Biology. At the time of submission of this thesis the article is in print ("forthcoming"), and therefore a plain-formatting version of the article is submitted with this thesis. In this study we definitively test the time-to-target based control hypothesis by first simulating the expected behaviour under a selectively tweaked movement requirements compared to our earlier work. In addition, we then experimentally test this control under two separate movement goals, validating the model predictions. Finally, we compare human behaviour under different experimental schedules – blocked (steady state) and mixed, showing that task specification acts as a contextual cue, allowing humans to rapidly switch between multiple feedback controllers.

Contributions

Justinas Česonis was the primary contributor and lead author in this research. In addition, Justinas Česonis designed the hypothesis, implemented the experimental design, collected and analysed the data, built the computational models and wrote the first draft. Both authors designed the experimental paradigm, selected the analysis methods and finalised the manuscript together.

Abstract

The separation of distinct motor memories by contextual cues is a well known and well studied phenomenon of feedforward human motor control. However, there is no clear evidence of such context-induced separation in feedback control. Here we test both experimentally and computationally if context-dependent switching of feedback controllers is possible in the human motor system. Specifically, we probe visuomotor feedback responses of our human participants in two different tasks – stop and hit – and under two different schedules. The first, blocked schedule, is used to measure the behaviour of stop and hit controllers in isolation, showing that it can only be described by two independent controllers with two different sets of control gains. The second, mixed schedule, is then used to compare how such behaviour evolves when participants regularly switch from one task to the other. Our results support our hypothesis that there is contextual switching of feedback controllers, further extending the accumulating evidence of shared features between feedforward and feedback control.

Contextual cues are not unique for motor learning: Task-dependant switching of feedback controllers

Justinas Česonis[1], David W. Franklin[1,2,3,*]

1 Neuromuscular Diagnostics, Department of Sport and Health Sciences, Technical University of Munich, Germany
2 Munich Institute of Robotics and Machine Intelligence (MIRMI), Technical University of Munich, Germany
3 Munich Data Science Institute (MDSI), Technical University of Munich, Munich, Germany

* david.franklin@tum.de

Abstract

The separation of distinct motor memories by contextual cues is a well known and well studied phenomenon of feedforward human motor control. However, there is no clear evidence of such context-induced separation in feedback control. Here we test both experimentally and computationally if context-dependent switching of feedback controllers is possible in the human motor system. Specifically, we probe visuomotor feedback responses of our human participants in two different tasks – stop and hit – and under two different schedules. The first, blocked schedule, is used to measure the behaviour of stop and hit controllers in isolation, showing that it can only be described by two independent controllers with two different sets of control gains. The second, mixed schedule, is then used to compare how such behaviour evolves when participants regularly switch from one task to the other. Our results support our hypothesis that there is contextual switching of feedback controllers, further extending the accumulating evidence of shared features between feedforward and feedback control.

Author summary

Extensive evidence has demonstrated that humans can learn distinct motor memories (i.e. independent feedforward controllers) using contextual cues. However, there is little evidence that such contextual cues produce similar separation of feedback controllers. As accumulating evidence highlights the connection between feedforward and feedback control, we propose that context may be used to separate feedback controllers as well. It has not been trivial to test experimentally whether a change in context also modulates the feedback control, as the controller output is affected by other non-contextual factors such as movement kinematics, time-to-target or the properties of the perturbation used to probe the control. Here we present a computational approach based on normative modelling where we separate the effects of the context from other non-contextual effects on the visuomotor feedback system. We then show experimentally that task context independently modulates the feedback control in a particular manner that can be reliably predicted using optimal feedback control.

Introduction

Whether it is touching a hot surface, returning a tennis serve or simply lifting an object, the human body utilises a variety of sensory inputs to produce movements of any complexity. Indeed, different feedback modalities of human motor control, such as stretch reflex [1–3], vestibulo-ocular reflex [4, 5], visuomotor [6–14], or even auditory feedback [15, 16] have extensively been studied in prior literature. However, most studies have investigated feedback control in paradigms of either a single task [17–21], or multiple tasks presented in their own dedicated blocks [22–26]. While such designs provide key insights into the behaviour of the feedback controller in isolation, they are not entirely reflective of human behaviour in real-life situations. For example, a realistic sequence of events could require a volleyball player to first pick up the ball from the ground by reaching for it with their hand and stopping on contact, only then to hit the same ball with the same hand a few moments later while serving. While studying both components independently has received focus in the field of motor control, any interactions between the feedback controllers in the context of rapid switching have not been broadly studied.

While feedback control in human movement is critical in correcting for random errors within movements, feedforward control corrects for movement errors that are predictable. In order to systematically predict and compensate for specific errors upcoming in a given movement, the mechanism of contextual switching via contextual cues is broadly accepted. It is now well understood that performing two opposing tasks in an alternating manner will lead to interference [27–29], resulting in behaviour that is averaged between the two tasks, failing to deal with either task. However, if the two tasks are performed in sufficiently different contexts, such as separate physical or visual workspaces [30–32], or different lead-in [33, 34] or follow-through movements [35, 36], this interference can be reduced, allowing the formation of two separate motor memories. Hence, it is reasonable to expect, that a similar contextual regulation could be present in feedback controllers. Therefore, in this study we test whether the feedback control policies exhibit such modulation when humans are presented with different tasks in an alternating manner.

One difference between studying contextual switching in feedforward and feedback control is that it is difficult to evaluate whether the feedback control policy has changed after the intervention. Specifically, it has been shown computationally that the optimal feedback controller (OFC) with fixed parameters can produce variable responses when correcting for perturbations within the movement, for example, when the comparable perturbations are induced in different parts (e.g. early or late) of otherwise identical movements [26, 37–39]. Furthermore, such behaviour was also observed in experimental studies [7, 20, 26, 39–41]. Hence, merely observing a difference in the feedback response is not enough to conclude a change in the control policy. However, recently we demonstrated that as long as two perturbations of the same magnitude are induced at the same time-to-target (which is defined as a difference between the perturbation onset and movement end), the same feedback control policy produces the same magnitude response, independent of whether the two perturbations occurred at the same location, time from the beginning of the movement, or the movement velocity [26]. Thus, we can utilise this relationship between the magnitude (or intensity) of the feedback response to a perturbation at the same time-to-target to quantify whether the difference in the response is due to the change in the control policy or not.

There are several studies that have already looked into contextual regulation in feedback control tasks. Most of such studies, to our knowledge, approached this question by modulating the structure of a target (wide vs. narrow, long vs short, etc.) [42–45], or by including obstacles along the reaching path [42, 45]. Results of these studies are consistent with optimal control-like behaviour with separate controllers for

different tasks, even when the target structure is changed on random trials [42, 44, 45], however one study suggests that control when the target is unpredictable may be sub-optimal [43]. In this study we test whether such rapid switching also holds true for tasks, where target structure remains unchanged, but the tasks themselves require a change of movement properties. Specifically, we test how the feedback control policies are affected when our participants are presented with a "multitasking" scenario where they have to switch between performing two distinct tasks, i.e. reaching to and stopping at the target, or hitting through the target and stopping behind it. While the two tasks are fundamentally different, and in isolation should require different feedback control policies, here we also test whether the same relationship holds true in the mixed schedule (as it would for contextual switching in feedforward control), or if the interference between two control policies results in a single policy, averaged or weighted between the two independent controllers.

Results

In this study we tested the behaviour of the human feedback controller when switching between two different tasks. Specifically, we presented our human participants with two tasks requiring different control policies – a stop task, where participants had to reach and stop at the target, and a hitting task, where participants had to punch through the target and stop behind it. In our previous work we demonstrated computationally that these two different types of movements trigger feedback responses of different magnitudes, even if the perturbations occur at the same position, time, or time-to-target [26]. However, if the two movements share the same goal (for example the goal of stopping at the same target), then these feedback responses match in magnitude if the time-to-target matches in both movements, irrespective of other movement parameters like peak velocity, movement distance, distance to the target or current velocity. Therefore, such a relation between time-to-target and feedback response intensity could be used to characterise the feedback control policy.

We use the relationship between the time-to-target and the feedback response intensity (which serves as a proxy for feedback controller gain) as a means to analyse the controller behaviour when the task changes. Specifically, we propose two alternatives for the architecture of such control: a single universal feedback controller that exhibits adaptation to a given task (Fig 1A), or multiple task-specific controllers, gated by task context (Fig 1B). When presented with a single task in a blocked schedule (e.g. blocked stop or blocked hit), both the universal controller and task specific controllers are expected to behave similarly, as the universal controller should easily adapt its gains appropriately for the required task. However, if multiple tasks are presented in a mixed schedule (i.e. task can randomly switch from trial to trial), the different control architectures predict different responses. Particularly, a single universal controller would aim to adapt to each presented task, thus on average producing responses somewhere in between the two given tasks within the mixed schedule (Fig 1C). In contrast, a set of task-specific controllers would produce similar responses in the mixed schedule as they would in a blocked schedule, as for every trial an appropriate controller is selected from a set of controllers, rather than being adapted for the task (Fig 1D).

In order to probe the control policies of human participants within these different tasks, we occasionally perturbed participants during the movement by visually shifting the target perpendicular to movement direction and inducing a reactive visuomotor feedback response (Fig 2A). Recently it has become common practice to maintain these perturbations until the end of the movement, such that an active correction is required to successfully complete the trial [19, 26, 41, 46–51]. However, we have noticed in our previous work that such maintained perturbations significantly impact the overall

time-to-target, which in turn affects the visuomotor feedback gains [26]. Thus, to keep the measurements of visuomotor feedback responses consistent within time-to-target, in this study we only perturbed our participants laterally in channel trials [7, 40, 52] and maintained these perturbations for 250 ms before switching them off, making any corrections redundant. As a result, even when producing the feedback response, participants' hands are constrained along the path of forward movement, resulting in matching movement durations independent of different perturbation onsets.

Participants produced involuntary feedback responses to the target jumps. These responses, observed as a lateral force exerted by the participants on the handle of the robotic manipulandum, were modulated by the different perturbation onsets (Fig 2B and Fig 2C). From these force responses we computed feedback intensities, by averaging individual responses over a time window 180 ms - 230 ms relative to the perturbation onset on each individual trial. This time window has now been used in numerous studies to quantify such responses and is associated with the involuntary, early

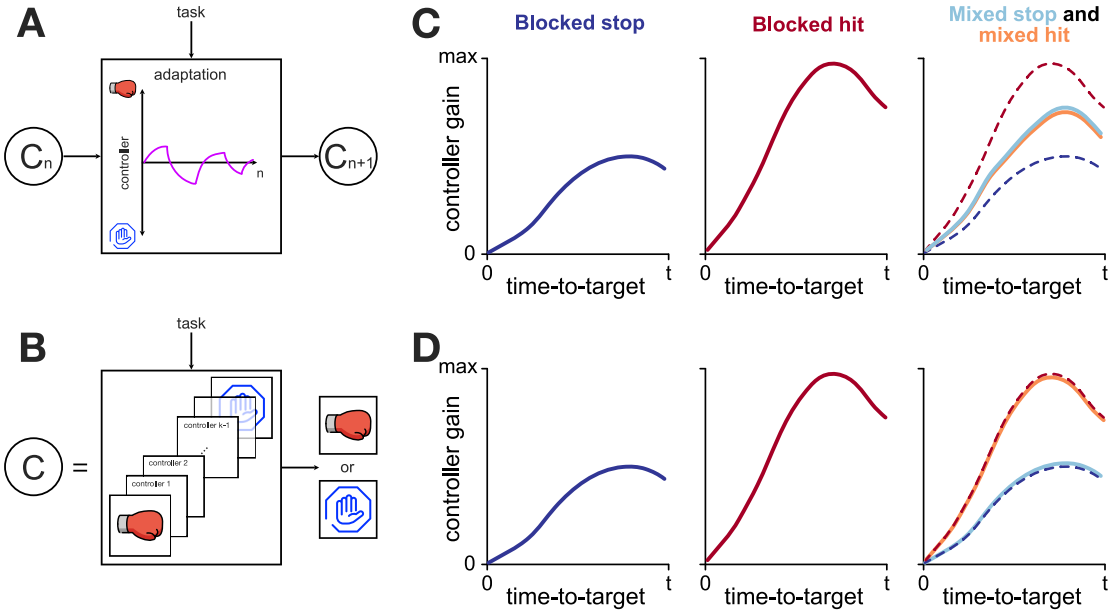


Fig 1. Theoretical predictions of two different architectures for feedback regulation. **A.** Universal feedback controller. A single feedback controller is used to produce both stop and hit movements, and is adapted to the given task over multiple trials. Such adaptive behaviour is reminiscent of the behaviour of the feedforward controller when learning two opposing force-fields without separable context. C_n indicates a feedback controller at trial n **B.** A feedback controller as a set of task-specific controllers. A task-specific controller (stop or hit) is selected based on the task-related context and is used during the given movement. Such contextual switching behaviour is reminiscent of the behaviour of the feedforward controller when learning two opposing force fields with separable context. **C.** Expected regulation of feedback responses by the universal feedback controller. When exposed to a single task for a long time (blocked schedule) the controller adapts to the given task, producing optimal responses for both stop and hit conditions. However, due to interference within the mixed schedule, such a controller would settle to the average (or weighted average) gains between the two blocked conditions. **D.** Expected regulation of feedback responses by a set of task specific controllers. Within the blocked schedule, similar regulation is expected between hit and stop as in the case of the universal controller (**C**). However, in the mixed schedule, due to the task-related context, an appropriate controller is recalled on a trial-to-trial basis, producing similar regulation as within the blocked schedule.

visuomotor responses [7, 25, 40, 50, 51].

OFC model predicts differences between hit and stop conditions

We utilised the mixed-horizon OFC [38] model, presented in our earlier work, to generate predictions of feedback control policies in our current study. Due to the experimental design of this study not requiring an extension in movement times after perturbations, the predictions of the mixed-horizon model also matched the predictions of our earlier time-to-target OFC model [26]. In order to compare differences in control throughout hit and stop movements, we first simulated two movement conditions: a 25 cm long movement with 60 cm/s peak velocity and velocity at the target distance <1 cm/s (stop condition), and a similar movement, but with velocity at the target >20 cm/s (hit condition) (Fig 3A). Both models were implemented using a linear quadratic regulator (LQR), and were identical, apart from the difference in state-dependent costs of terminal velocity and terminal force. Here we reduced these cost parameters for the hit model by a factor of 50 in order to reduce the incentive to stop at the target, and thus successfully simulate hit-like movements. In addition, we also simulated a third condition, that we term the long-stop condition, where we used the same position, velocity, force and mean activation costs as in the stop model, but applied for reaching movements of 28 cm. The concept of the long-stop model is to compare the actual hit behaviour, executed through a different controller, with "cheating" behaviour where the same stop movement is performed to an imaginary target, located beyond the actual target, resulting in non-zero velocity at the actual target, and thus appearing as a hit movement. For all three conditions we then induced virtual target perturbations by shifting a target laterally by 2 cm at every time step from movement onset to movement end. With these simulations we obtained one continuous feedback response profile per condition, showing a dependency of feedback response intensity on time-to-target (Fig 3B and Fig 3C). This feedback response profile is characteristic of the particular

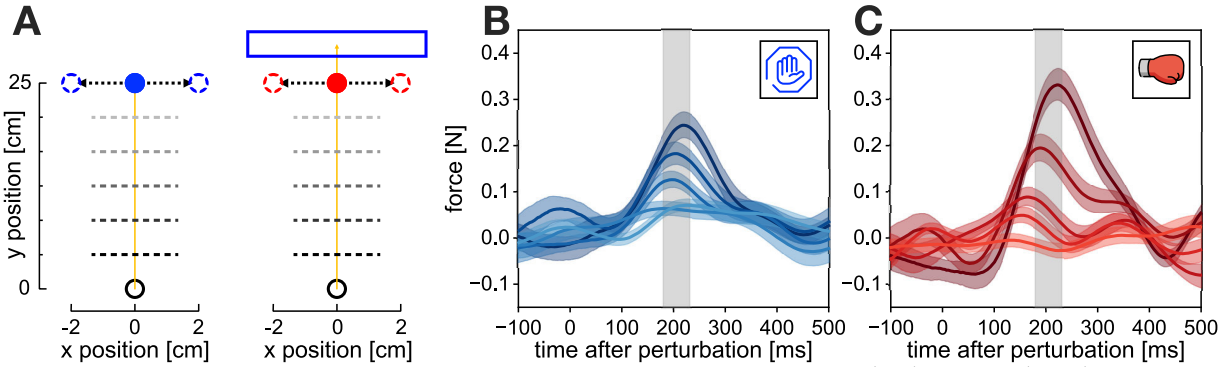


Fig 2. Experimental perturbations and responses. **A.** Perturbations in stop (left) and hit (right) conditions. Participants performed a forward reaching movement towards a target, positioned 25.0 cm in front of the start position. When the hand crossed one of five evenly spaced locations (dashed lines), a perturbation could be induced by shifting the target by 2 cm laterally for 250 ms and then returned back to the original position. Participants were instructed to either stop at the blue target (stop condition), or hit the red target and stop within the blue rectangle (hit condition). **B.** Net unscaled feedback responses to the target perturbations in the stop condition, measured via the force channel. Participants produced corrective responses to the target perturbations that varied by different perturbation onsets. Different traces represent different perturbation onsets, with darker colours indicating earlier perturbations. Shaded areas represent one standard error of the mean (SEM). The grey rectangle represents the time window of 180 – 230 ms, where the visuomotor feedback intensities are measured. **C.** Net unscaled feedback responses to the target perturbations in the hit condition.

movement control policy associated with the movement goal, as it is maintained even if the kinematics of the movement change (Fig 8 in [26]).

Even with similar simulated kinematics, that deviate from each other only in the last portion of the movement, the OFC model predicts striking differences in the control policies for stop and hit conditions (Fig 3B and Fig 3C, blue and red traces) or hit and long-stop conditions (Fig 3B and Fig 3C, red and green traces) when expressed against time-to-target. On the other hand, when expressed against position, even different controllers (hit vs stop) show no differences in feedback intensities, while identical control (stop vs long-stop) exhibit clear differences (Fig 3D). Among other things, these results point out limitations of position as a dependent variable in determining the changes of control policies, and provide yet additional support for time-to-target.

Our models make a few predictions for the behaviour of human participants. First, it challenges the classic assumption that visuomotor feedback response profiles are always bell shaped, if probed at evenly spaced locations or movement times. Instead, we propose that the bell-shaped feedback response profiles are consequential to the specific kinematic values imposed by the experiments, and other, for example monotonically decreasing intensity profiles, are also possible with faster movements (Fig 3D). Second, our simulations also make predictions on relative differences between the feedback intensity profiles in stop and hit conditions. Particularly, we expect the hit condition to produce stronger responses than the stop condition for short times-to-target, with this relationship inverting for long times-to-target if the two types of movements require different feedback controllers (Fig 3B and Fig 3C). Note, that while in previous studies it is typical to compare such response profiles in terms of perturbation onset location, here no difference between hit and stop is predicted in position-dependent profiles (Fig 3D).

Human control policies match model predictions in hit and stop conditions

In order to compare the behaviour of our participants to the model predictions, we first analysed our results from the blocked schedule of the experiment. Here every participant has completed a block of 416 trials of hit condition and another block of 416 trials of stop condition, with the order counterbalanced across all participants. Our experimental results qualitatively match the predictions of our model. First, participants successfully differentiated between the kinematics of the hit and the stop condition, with both types of movements resulting in matching early and peak velocity ($v_{peak,stop} = 58.9$ cm/s, $v_{peak,hit} = 58.1$ cm/s), but with differences towards the end of the movement such that the velocity at the target is higher for the hit condition (Fig 3E). Specifically, in the hit condition participants produced movements with average velocity at the target of 38.5 cm/s, while successfully stopping at the target in the stop condition. In addition, similar to the model simulations, movements in the hit condition were of slightly shorter duration (630 ms vs 700 ms).

Qualitatively, the experimental feedback responses also match the model predictions (Fig 3F and Fig 3G). First, due to relatively fast reaching velocities in our experiment, as well as the lack of maintained perturbations, all perturbations were induced at short times-to-target (under 550 ms). For comparison, in our previous study [26] perturbations were induced at times-to-target that ranged between 300 ms and 1000 ms, with peak feedback intensities recorded for perturbations with time-to-target at 400 ms. Second, both our data and the model produce feedback intensities at short times-to-target that are higher for the hit condition than for the stop condition, even in movement segments where the kinematics are otherwise similar. In addition, we also observe no learning effects within this regulation, as the relative behaviour across conditions is present in the first few blocks of the study, and remains throughout the

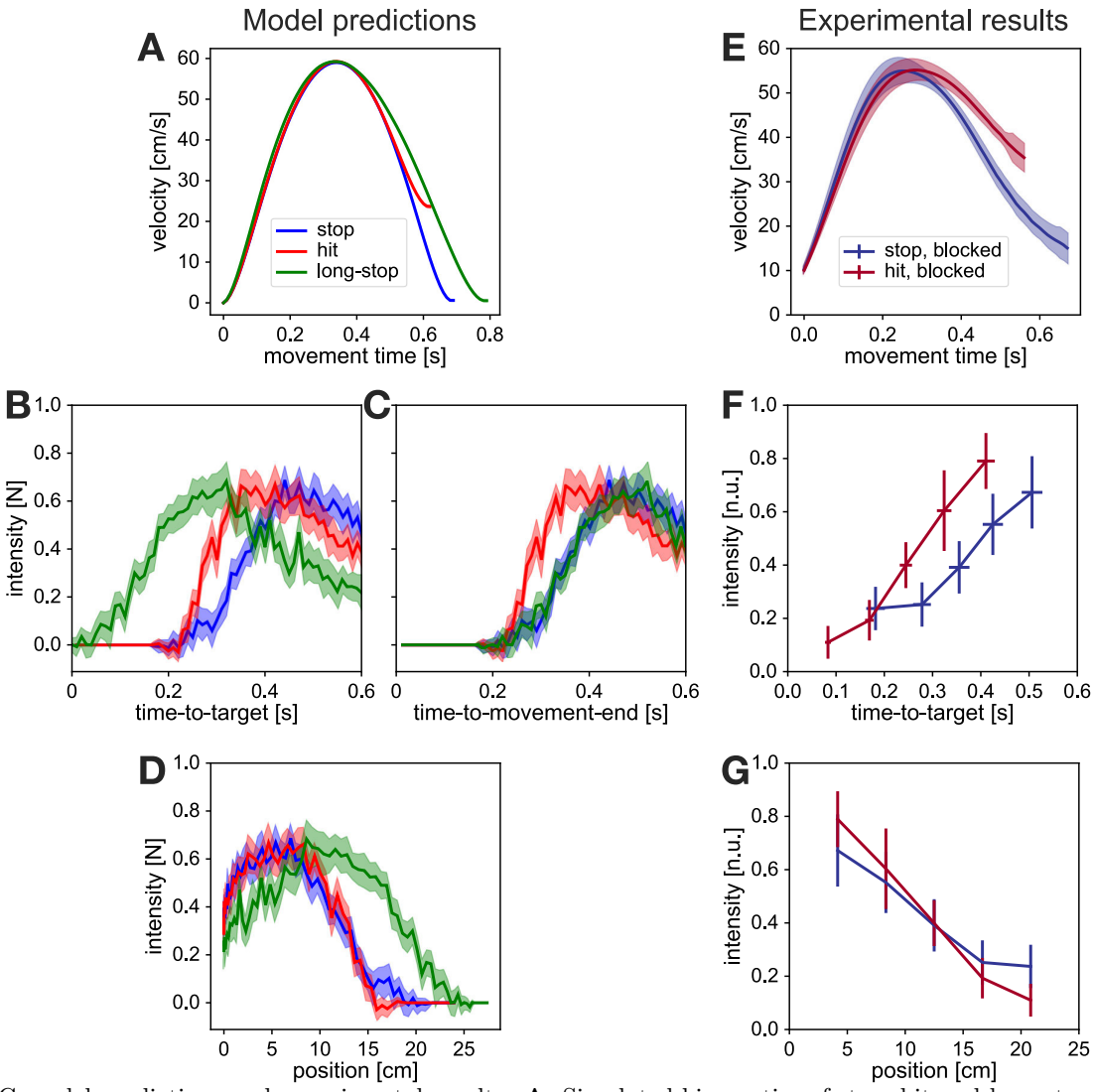


Fig 3. Comparison of OFC model predictions and experimental results. **A.** Simulated kinematics of stop, hit and long-stop conditions. Stop and hit conditions produce matching kinematics that only deviate shortly before movement end. The long-stop condition is a control simulation, that matched the kinematics of the hit condition for the duration of the hit movement, but was achieved with the same stop controller. **B.** Simulated feedback intensities as a function of time-to-target and **C.** time-to-movement-end. Simulations predict a faster increase of response intensities for hit condition than for stop condition. As the long-stop condition is simulated via a longer (28 cm) movement, the time-to-target represents a time until the simulated movement crosses a point of 25 cm distance (the target distance). For hit and stop conditions, time-to-target and time-to-movement-end are identical. When expressed against time-to-movement-end, long-stop produces matching responses to the stop condition, as the feedback controller used for these movements is identical. With respect to the time-to-target, long-stop responses are time-shifted from the stop responses. **D.** Simulated feedback intensities as a function of the position. Stop and hit simulations with these particular kinematics produce matching feedback intensity profiles when expressed against position, even if the feedback controllers are different. In contrast, the long-stop simulation with a feedback controller matching that of the stop condition still produces different intensity profile against position. Shaded areas in simulated traces represent 95% confidence intervals for simulated results. **E.** Velocity profiles of participants in blocked stop and blocked hit conditions. The profiles match the task requirements. **F.** Normalised feedback intensity profiles of participants in blocked stop and blocked hit conditions, expressed against time-to-target. Participants produce stronger responses at matching time-to-target in the hit condition, consistent with simulation results for hit and stop. **G.** Normalised feedback intensity profiles of participants in blocked stop and blocked hit conditions, expressed against position. Participants produce matching responses within hit and stop conditions, supporting model simulations for stop and hit conditions, and not stop and long-stop. Error bars in experimental results represent 95% confidence intervals.

entire experiment (S1 Text). Importantly, we do not fit the model to match the data, but instead use it to qualitatively describe the relative regulation of stop and hit conditions. As such, matching features between the intensity profiles of the model (Fig 3B, Fig 3C and Fig 3D) and the data (Fig 3F and Fig 3G) imply that similar computational mechanisms may be in action. Finally, our results also indicate that participants utilise different feedback controllers for the hit and stop conditions, as the experimental results for the blocked hit condition match the model simulations of the hit, rather than the long-stop condition.

Human participants utilise contextual switching of feedback controllers

In the previous sections we established the differences between the baseline control policies of hit and stop conditions. Here, we test how these policies change when the exposure to these conditions is no longer blocked. For example, it is natural in our daily activities to continuously switch between tasks, rather than doing a single task for many repetitions before switching to a new task. However, the question remains, how switching between different tasks affects the underlying feedback control policies. To test this, in the second half of the experiment we presented our participants with the same two types of movements (hit and stop), but now with the conditions randomly mixed across trials, instead of being presented in two separate blocks. As such, we could test for one of two possible outcomes:

1. Control policies for stop and hit movements in the mixed schedule match respectively the control policies in the stop and hit movements in the blocked schedule (Fig 1D). Such an outcome would indicate that participants are able to easily switch between different control policies (at least within consecutive trials).

2. Control policies for stop and hit movements in the mixed schedule do not match with the respective baseline policies, indicating interference when switching among multiple conditions (Fig 1C).

While both outcomes have previously been discussed from the sensorimotor adaptation perspective, to our knowledge they have not yet been demonstrated for feedback control.

Our participants successfully produced the movements required in the experiment (Fig 4A). Particularly, we observed clear distinctions in the terminal velocity between the hit and stop conditions, independent of the experimental schedule (blocked or mixed). A two-way repeated-measures ANOVA showed a significant main effect on condition (hit or stop, $F_{1,13} = 544.2$, $p \ll 0.001$), but no significant main effect on experiment schedule (blocked or mixed, $F_{1,13} = 0.710$, $p = 0.42$) or schedule/condition interactions ($F_{1,13} = 0.681$, $p = 0.42$). In addition, a complementary Bayesian repeated-measures ANOVA analysis showed similar results, with a very strong effect [53] of condition (hit or stop, $BF_{incl} = 1.6 \times 10^{25}$), and with a tendency towards no effect of schedule (blocked or mixed, $BF_{incl} = 0.379$), or condition/schedule interaction ($BF_{incl} = 0.409$). A similar analysis for peak velocities showed a significant main effect of condition (hit or stop, $F_{1,13} = 5.94$, $p = 0.03$; although $BF_{incl} = 1.12$ indicates not enough evidence to either reject or accept the null hypothesis) and condition/schedule interaction ($F_{1,13} = 19.3$, $p \ll 0.001$; $BF_{incl} = 32.6$), but not on schedule (blocked or mixed, $F_{1,13} = 1.52$, $p = 0.24$; $BF_{incl} = 0.56$ shows a weak tendency towards accepting null hypothesis). The Holm-Bonferroni corrected post-hoc analysis for the interaction term revealed that participants produced slightly faster movements in the mixed-hit condition, with the peak velocities matching otherwise.

We examined the evolution of the experimental visuomotor responses as a function of perturbation onset position or onset time across the four different conditions (Fig 4C

and Fig 4D). When expressed against either position or time, the visuomotor intensity profiles do not show the classical bell-shaped profile where strongest responses occur in the middle of the movement and are reduced towards the beginning and end. Instead, our participants produced the strongest responses for the earliest perturbations, induced at 1/6 of the total forward movement, with further responses decaying in intensity as perturbations occurred closer to the target. Moreover, we observed no significant differences in visuomotor responses across the different conditions and schedules. Three-way repeated-measures ANOVA with condition (stop or hit), schedule (blocked or mixed) and perturbation location (5 levels) as main factors showed no effect of condition ($F_{1,13} = 0.486$, $p = 0.50$; $BF_{incl} = 0.238$ shows substantial evidence towards no effect), schedule ($F_{1,13} = 0.096$, $p = 0.76$; $BF_{incl} = 0.142$ shows substantial evidence towards no effect) or condition/schedule interaction ($F_{1,13} = 0.657$, $p = 0.43$; $BF_{incl} = 0.305$ shows substantial evidence towards no effect). While we observed a significant main effect of

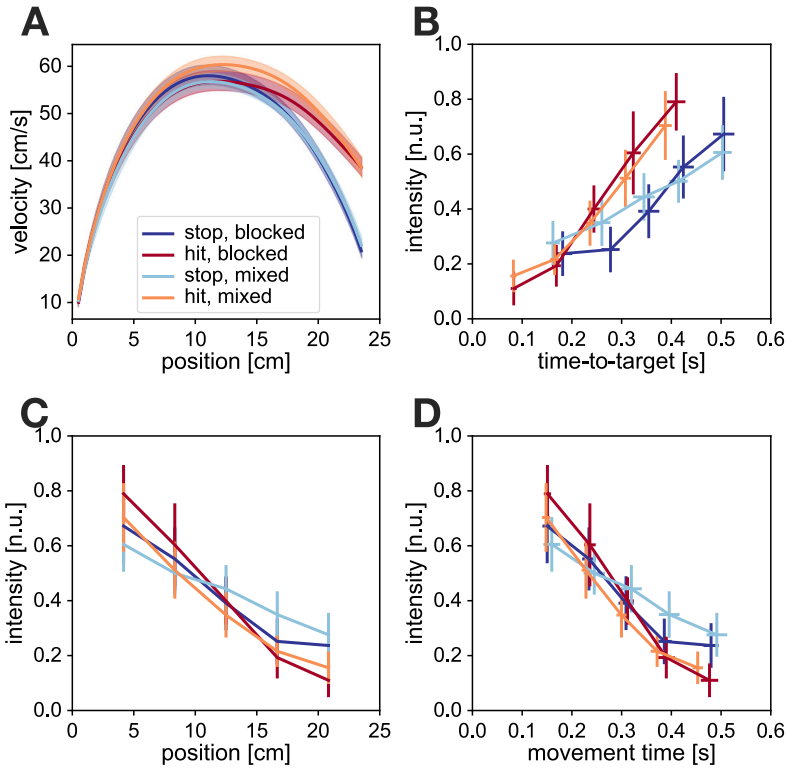


Fig 4. Experimental results of stop and hit conditions in both blocked and mixed schedules. **A.** Velocity profiles against position. Both stop conditions and both hit conditions produce respectively similar velocity profiles, showing that participants successfully performed the task in the mixed schedule. **B.** Normalised feedback response intensities represented as a function of time-to-target. Hit and stop movements in the mixed schedule demonstrate differences when expressed against time-to-target, that match the differences between hit and stop conditions in the blocked schedule. This supports the hypothesis of contextual controller switching between multiple task-specific controllers. **C.** Normalised feedback intensities in all four conditions show no differences when expressed against position or **D.** movement time at perturbation onset, as predicted by the OFC simulations. This questions the appropriateness of position or movement time as the reference frames in which to compare multiple feedback controllers. Error bars and shaded areas indicate 95% confidence intervals of the mean.

the perturbation location ($F_{2.9,37.7} = 61.2$, $p \ll 0.001$ after Greenhouse-Geisser sphericity correction; $BF_{incl} = 9.3 \times 10^{36}$), such an effect was expected due to the temporal evolution of feedback responses. In addition, we observed a significant interaction between perturbation onset location and the condition ($F_{2.1,27.0} = 6.26$, $p = 0.005$ after the sphericity correction; $BF_{incl} = 6.86$), however a Holm-Bonferroni corrected post-hoc analysis on the interaction term did not indicate any meaningful interaction effects, with none of the significant interactions appearing at the same perturbation onset location. Finally, the remaining interactions of schedule/perturbation ($F_{2.6,33.9} = 2.67$, $p = 0.07$ after Greenhouse-Geisser sphericity correction; $BF_{incl} = 0.289$) and condition/schedule/perturbation ($F_{2.8,36.8} = 0.233$, $p = 0.86$ after Greenhouse-Geisser sphericity correction; $BF_{incl} = 0.075$) showed no significant effects. Thus, as a whole our analysis indicates that the feedback controllers could not be differentiated when expressed as a function of the position within the movement.

When expressed against time-to-target, the visuomotor feedback responses show decreasing feedback intensities with decreasing time-to-target, with responses virtually vanishing when the time-to-target approaches zero (Fig 4B). This behaviour is consistent with our previous models describing the time-gain relationship [26]. In addition, we observe stronger increases in visuomotor feedback intensity with increasing time-to-target for the hit condition compared to the stop condition, in both blocked and mixed schedules. Such regulation was previously predicted by our time-to-target OFC model (see Fig 9C in [26]) for short times-to-target. Finally, we also observe a qualitative match between the two stop conditions (mixed and blocked) as well as between the two hit conditions (mixed and blocked), suggesting first evidence of rapid feedback controller switching in the mixed schedule. This finding holds equally well in trials immediately after a condition switch, as well as after the trials of the same movement condition (S2 Text). Qualitatively the increase of visuomotor response intensities with time-to-target for our specific results could be well described by a line function for each of the four combinations of condition and schedule. In order to get a quantitative estimate of the differences between the conditions we performed a Two-way ANCOVA analysis of visuomotor response intensity, with schedule and condition as the two factors, and time-to-target as the covariate. The results showed a significant main effect of condition (hit or stop, $F_{1,275} = 24.8$, $p \ll 0.001$; $BF_{incl} = 9.46 \times 10^3$), and time-to-target ($F_{1,275} = 222.8$, $p \ll 0.001$; $BF_{incl} = 1.04 \times 10^{33}$), but no effect of the experimental schedule (blocked or mixed, $F_{1,275} = 0.098$, $p = 0.75$; $BF_{incl} = 0.138$) or of schedule/condition interaction ($F_{1,275} = 1.06$, $p = 0.30$; $BF_{incl} = 0.304$ shows tendency towards no effect). Such results indicate that we can successfully separate the two different controllers when expressing their feedback response intensities (or their gains) against the time-to-target. Furthermore, we also show that such differences are only present when comparing the controllers for different tasks, and are not dependent on the presentation schedule of these tasks. Thus, we demonstrate that our participants successfully selected an appropriate controller for a hit or a stop task, even in a schedule where the task could change on consecutive trials.

Discussion

In this study we have demonstrated that humans are capable of rapid switching between appropriate feedback controllers in the presence of different contextual cues. Specifically, our participants show systematic differences in feedback responses when performing hitting movements, compared to reach-and-stop movements. Moreover, the same systematic differences are present, both when the two tasks are performed in isolation (blocked schedule), or when rapidly switching from one task to the other (mixed schedule), showing that these differences are evoked within a single trial, and not

gradually adapted. Finally, these feedback responses are also well matched with the optimal feedback control predictions for these responses in hit and stop tasks, further reinforcing accumulating evidence of optimality principals in the feedback control of human movements.

In order to gain insight into computational mechanisms that are employed when humans switch between hit and stop conditions, in this study we formulate our hypothesis through normative modelling [37, 54–58]. Such an approach compares the behavioural experimental data to the results simulated computationally through a known bottom-up design. In turn, any mismatch between the data and the model rules out the mechanism, while matching behaviour provides support for the likelihood of such a mechanism. Specifically, here we simulate three different types of control movements: stop movement, where a point mass is stopped at a target 25 cm away from the start position; hit movement, where the point mass is instead brought to the same target with nonzero terminal velocity; and a long-stop movement, with similar kinematics to the hit movement within the 25 cm segment, generated by a stop movement to a secondary virtual target at 28 cm distance. The hit and stop simulations differed in the implementation of the feedback controller, with the state dependent costs for the terminal velocity and terminal acceleration reduced by a factor of 50 for the hit condition. As a result, the two models inherently simulate the behaviour that is achieved via different controllers. On the other hand, the long-stop condition was simulated by using the same controller as the stop condition, but to a target at 28 cm instead of 25 cm. Consequently, such a movement still maintained a non-zero velocity at 25 cm, virtually simulating a hit-like movement. Notably, in order to better match the kinematics of a long-stop movement to the kinematics of the hit and stop movements, we temporally modulated the activation cost R of the long-stop controller, which we have previously shown does not affect the overall feedback responses in terms of time-to-target [26]. In general, while kinematics of hit and long-stop models matched well, the two simulations predicted very different feedback response profiles when expressed both against time-to-target and against position. Finally, the responses of our participants in the hit condition matched better with the model simulation of the hit condition, rather than the simulation of the long-stop, providing evidence that humans use different feedback controllers for different tasks.

Principles of contextual switching have been extensively studied in the context of feedforward adaptation [30, 31, 34, 59–62]. While these cues vary in effectiveness [30, 59] and are typically considered as relative weightings of multiple feedforward models [63], strong dynamic cues such as differences in follow-through [35, 36], lead-in [33, 61], or visual workspace [31, 32] can effectively separate the feedforward models. As multiple recent papers have demonstrated that voluntary (feedforward) and feedback control likely share neural circuits [24, 39, 64–67], it is reasonable to believe that similar contextual regulation would also be present in feedback control. However, studies that have shown this parallel changes in the feedback responses to the learning of the feedforward dynamics, either examined before and after adaptation to novel dynamics [24, 64, 68, 69], or during the process of adaptation [19, 70–72], meaning that the they could not distinguish between the slow adaptation of the feedback controller to each condition or the rapid switching between two controllers. Moreover, other studies have suggested that feedforward and feedback controllers are learned separately [73, 74] and may even compete with one another [75], suggesting that these share different neural circuits and may have different properties. In this study we showed that in the mixed schedule, where the task goal unpredictably switched between hit or stop tasks on consecutive trials, participants evoked different control policies for each task. Furthermore, these policies, evoked within mixed schedule, well matched with the respective policies in the blocked schedule, suggesting that they were not only different

from one another, but also appropriate for each task, showing the strong separation of the two contexts. While this is not unexpected, as the two hit and stop tasks are significantly different in their dynamics and thus should act as a strong contextual cue, one important result is that we demonstrated that the context regulates the feedback, and not only feedforward control. Finally, our results are also consistent with the accumulating evidence of the shared relationship between feedforward and feedback control in human reaching.

One reason why contextual effects on feedback control have not been broadly studied, is that it is difficult to quantify what really constitutes a change in feedback control policy. For example, we can trigger responses of different magnitudes by changing the size of the perturbation [41, 42], inducing perturbations at different positions [26, 40] or at different times [20, 37, 76]. However, computationally such differences in response intensity can be achieved within the same optimal feedback controller without ever changing control parameters [38]. On the other hand, experimental tasks, presented in some of these studies, e.g. reaching towards narrow, wide or long targets, inherently require different feedback controllers. Specifically, assuming a similar controller to the one we present in this work, a wide target implies a reduced $\omega_{p,t}$ weight in x-axis compared to the narrow target, thus leading to a different optimum of control matrix L (Eq. 2 and 3). Indeed, human responses in tasks where the target structure changed (either by shape or by the presence of obstacles [42, 44, 45]) were consistent with the OFC predictions of two independent controllers [42]. In this article we present two tasks that also require different feedback controllers, but achieve that while maintaining the target shape. Instead, we invoke different controllers by modulating the task requirement of either stopping at the target, or hitting through it. In addition, by combining an OFC model predictions with our previous work, showing that the time-to-target is a strong predictor of the feedback intensity in optimal control tasks [26], we not only show that the human behaviour is consistent with two independent controllers, but also that it cannot be explained by one controller. Specifically, we simulate the behaviour either by recomputing the controller L (Fig 3B, Fig 3C and Fig 3D, hit and stop), or by updating the state estimate $\hat{x}$ and using the same controller L (Fig 3B, Fig 3C and Fig 3D, stop and long-stop), to compare with the experimental results (Fig 3F and Fig 3G). These results show that humans indeed change their control policies when the task goal (e.g. hit or stop) changes. Thus, by combining behavioural results with normative control models we can clearly identify that it is specifically the change in control, and not other mechanisms, that is responsible for the regulation observed in the experimental data.

Previous studies have demonstrated that visuomotor feedback intensity profiles are roughly bell-shaped along the movement – low at the beginning and the end, and peaking in the middle [26, 40] – leading to assumptions that these gains might parallel the velocity [39, 67]. Our simulations and experimental results (Fig 3D and Fig 3G) demonstrate that this bell-shape profile is not fixed, and that other profiles are possible. In our previous work, we established a robust relationship between the visuomotor feedback intensities and time-to-target, demonstrating that time-to-target is the fundamental variable that modulates the responses, given that the task goal (and thus the feedback controller) remains the same [26]. This means that the bell-shaped profile is simply a by-product of a specific timing of perturbations, and is not regulated by their onset location. As a consequence, the shape of these feedback intensity profiles can be modulated away from the bell-shaped profile by changing movement speed, target distance or acceleration profile. Such results illustrate possible caveats in the experimental paradigms of motor control: historically, some of the task requirements have been largely consistent, particularly in terms of reaching distance, reaching speed or duration. This may result in some measured behavioural outcomes being specific to

these kinematics or conditions rather than representing the general features of the motor control system. Thus, while we do not advocate for routinely altering the standard experimental and analytical methods, it is worth considering the specific biases that such methods may contribute to a given study.

One popular way of looking at the visuomotor responses in humans is how they vary with position in a movement. Indeed, numerous studies either analyse the evolution of responses against position [26, 39, 40], or induce perturbations based on a fixed position [7, 9, 25, 41, 49–51], with the expectation that these perturbations induce similar responses unless the control changes. For example, [9] demonstrated different feedback responses, induced at a matched position in movements towards different targets. While we believe that these different target properties indeed suggest different feedback controllers, such a distinction cannot be reliably tested with only one perturbation, matched by position. Our results clearly demonstrate the limitations of position as the main variable to probe such control. On one hand, even with similar kinematics for the majority of the movement, simulations of stop and long-stop movements predict radically different responses at matching positions (Fig 3D), despite the fact that these are generated with identical controllers. On the other hand, different controllers for hit and stop conditions still produced roughly matching feedback responses at the same position, consistent with the experimental data (Fig 3D and Fig 3G). In contrast to position as the main variable, OFC simulations in both this study and our previous work [26] show that the same controller, when expressed against time-to-target, produces matching response profiles, independent of other kinematic factors such as movement velocity or position of the perturbation onset (Fig 3B and Fig 3C). Furthermore, different controllers, such as hit and stop, produce feedback responses with systematic differences when expressed against time-to-target, exactly as demonstrated by our participants. Thus, we propose that time-to-target is the better reference frame for comparing feedback responses.

In this study we have raised two alternative hypotheses about the regulation of feedback controllers within the mixed schedule. The first possibility is that the feedback control gradually adapts to a given task over a few consecutive trials, similar to the feedforward control during learning of a force field or visuomotor rotation. If such adaptation was true, we expect different feedback intensities between the hit and stop conditions in the blocked schedule as the controller has enough trials to reach steady-state behaviour. However, in the mixed schedule the controller would drift between the equilibrium of hit and stop conditions, producing similar responses for mixed hit and mixed stop conditions. Note that even in such a case where only a single controller is performing both hit and stop trials, we would not necessarily expect any effects on the kinematics or the participant's ability to complete the task. Instead, due to the feedback nature of the control, a sub-optimal controller would still complete the movement, but produce sub-optimal (e.g. more costly) responses in the presence of external disturbances. The second possibility is that an appropriate controller is selected before each movement based on the provided context, allowing immediate switching between tasks. In this case, the feedback intensity profiles would match for the same task, regardless of the schedule of their presentation. That is, we expect to see similarities between both hit conditions, as well as between both stop conditions, but differences between any two hit and stop conditions. Our experimental results strongly support the latter option, as we not only observe differences between mixed hit and mixed stop conditions, but also observe their respective match with the blocked conditions. While our results do not rule out the adaptation of feedback controllers in general, we do demonstrate that different optimal controllers can be rapidly selected and switched between for familiar tasks.

One important aspect of the relationship between feedforward and feedback control

is that modulating one of them should affect the behaviour of the other. Indeed, previous work has demonstrated that human participants changed their feedback gains after adapting their feedforward models to novel dynamics [11, 19, 24, 64, 68–72]. However, an adapted movement in the force field typically produces kinematics that are similar to those in baseline movements, suggesting that such change of gains is achieved at matching times-to-target, and with the same task goal. Thus, our proposed framework that the relation between feedback intensities and time-to-target is unique for a unique controller would predict that the feedback gains would remain unchanged. As a result, we can not directly explain this change of control gains, unless the feedback controller somehow changes during adaptation. One factor driving such a change is that adapted movements in the force field are more effortful than baseline movements, due to additional muscle activity required to compensate for the force. An increased effort in the context of OFC simulations would thus increase the model activation cost R, resulting in a change of optimal feedback gains and intensities at matching times-to-target. In addition, the presence of a force field likely influences the biomechanics of the movement (particularly the muscle viscosity b), changes the state transition due to the external dynamics (via state transition matrix A), and updates the state uncertainty [77], resulting in the same controller being applied to a different control plant, and thus producing different responses. Moreover, if the controller is optimised to to this new control plant, adaptation will inevitably require a new feedback controller. Therefore, such changes in feedback control are expected, even though conventionally it appears that the task goal remains the same after adaptation to the novel dynamics.

Even though many recent studies use force channel trials [52] to accurately measure the visuomotor feedback responses [7, 11], often these brief perturbation trials are complemented with maintained perturbation trials [19, 26, 41, 46–48, 50, 51]. This is because brief perturbations within a channel trial are task-irrelevant, and can be ignored without compromising the task, whereas maintained perturbations strengthen these responses as they require an active correction for the participant to reach the target. However, we have recently shown that these maintained perturbations also force a non-trivial extension of the movement duration compared to the non-perturbed movement, and thus complicate the relationship between the perturbation onset location and the time-to-target. Hence, in order to consistently evaluate the control behaviour and its relation to the time-to-target, here we deliberately chose to only induce perturbations within the force channels and not to include the maintained perturbations. Although this generally decreases overall feedback intensities, our participants produced clear responses that exhibited the temporal evolution as predicted by the OFC model simulations.

Another possible limitation of using channel trials to probe the feedback control is the potential interference of the stretch reflex. Specifically, small forces produced on the hand by the channel could set on the feedback corrections [78] that superimposed onto the measured visuomotor responses. However, as the onset of the force channel occurs prior to any movement of the participant, long before the time of the visual perturbation, the channel onset will not produce a stretch reflex response timed to the visual stimuli. More importantly, such effects, as well as any other corrections to the channel onset would be present in all channel trials (including zero-perturbation trials), and thus would cancel-out in the net feedback responses, as they would not depend on the direction of the visual perturbation. As a result, it is unlikely that the force channels introduced systematic effects into our recorded visuomotor feedback signal. Similarly, we also saw no behaviour differences between simulations of free movements, that we presented in our results, and analogous simulations of matching duration movements in channel trials (S1 Fig).

Most studies of motor learning study contextual switching in conjunction with dual adaptation by introducing participants to novel force fields or visuomotor rotations, for which they do not have pre-existing feedforward controllers. In turn, we typically see slow, simultaneous adaptation to applied perturbations, as well as context-induced switching of the memories after the transient learning phase is over. Importantly, for familiar tasks this switching is evoked immediately, without the need to re-learn the dynamics again on re-exposure. This is clearly seen on the second or later days after adapting to dual force fields [35, 79]. In this study, our main goal was to demonstrate that such contextual switching is also possible for feedback controllers, rather than to demonstrate gradual adaptation. Therefore, in our experimental design we consciously selected two tasks (stop or hit) that were not novel to our participants. While there remains a possibility that due to the dynamics of the vBot environment both tasks were different to stop or hit movements outside of the lab, and thus novel to participants, we always started our studies with the blocked schedule, and only then followed with the mixed schedule to make sure that the different baseline controllers are already available to our participants. An interesting control would be to first test the participants in the mixed schedule, followed by the blocked schedule. However, we believe such control would mainly test whether the two task choices were novel to participants or not, which is not the focal point of our study.

In summary, here we again demonstrate that time-to-target (which could be considered as one form of urgency) [20, 26, 39, 76, 80], and not position or velocity, act as a primary predictor for the feedback response intensity when the task goal is fixed. Moreover, when comparing multiple tasks, the time-to-target reference frame consistently separates the feedback control policies for these tasks – an outcome that fails when comparing two different controller gains within the position reference frame. While position within the movement, and velocity at the time of a perturbation, definitely influence the controller responses, our results clearly demonstrate that the effect of these variables on overall control may be somewhat exaggerated in the previous literature. For example, our participants produced temporal evolution of the responses to visual perturbations that neither paralleled the velocity, nor showed the typical variation with position (with peak responses achieved mid-movement), but could be explained by the time-to-target dependency that was predicted by OFC. In addition, participants were able to switch their feedback controller from one trial to another, demonstrating the principle of contextual switching for feedback control. Such switching, well known in feedforward control, further reinforces accumulating evidence of the shared connections between feedforward and feedback control. Most importantly, our results demonstrate that the visuomotor feedback control in humans not only follows the principles of optimal control for a singular task, but also selects an appropriate controller for that task upon presenting the relevant context.

Methods

Ethics statement

The study was approved by the Ethics Committee of the Medical Faculty of the Technical University of Munich. All participants have provided a written informed consent before participating in the study.

Experimental setup

Fourteen right-handed [81] human participants (age 21-29 years, 5 females) with no known neurological diseases and naïve to the purpose of the study took part in the

experiment. Participants performed forward reaching movements either to a target (stop condition) or through the target (hit condition) while grasping the handle of a robotic manipulandum (vBOT, [82]) with their right hand, with their right arm supported on an air sled. Participants were seated in an adjustable chair and restrained using a four-point harness in order to limit the movement of the shoulder. A six-axis force transducer (ATI Nano 25; ATI Industrial Automation) measured the end-point forces applied by the participant on the handle. Position and force data were sampled at 1 kHz, while velocity information was obtained by differentiating the position over time. Visual feedback was provided via a computer monitor and mirror system, such that this system prevented direct vision of the hand and arm, and the virtual workspace appeared in the horizontal plane of the hand (Fig 5A). The exact timing of any visual stimulus presented to the participant was determined from the graphics card refresh signal.

Participants controlled a yellow cursor (circle of 1.0 cm diameter) by moving the robotic handle. The centre position of this cursor in the virtual workspace always matched the physical position of the handle. Every experimental trial was initiated when the cursor was brought into the start position (grey circle of 1.6 cm diameter), which was located 20 cm in front of participants' chest and centred with the body. When the cursor was within this start position, the circle changed from grey to white and the type of experimental trial was indicated by the presentation of a target. After a random delay, sampled from an exponential distribution with $\lambda = 0.7$ and truncated outside 1.0 s - 2.0 s interval, a tone was played to indicate the start of the movement. If participants failed to leave the start position within 1000 ms after this tone, the procedure of the current trial was aborted and restarted.

Over the course of the experiment the participants were tasked to complete two types of movements: stop movements, where they were required to stop within the target (a circle of 1.2 cm diameter, located 25.0 cm in front of the start position) (Fig 5B), or hit movements, where they had to intercept the target without stopping, and instead stop in a designated stopping area (a blue rectangle, [width, height] = [15 cm, 4 cm], centred 5 cm beyond the target) (Fig 5C). The reaching movement was considered complete once the centre of the cursor was maintained for 600 ms either within the area of the target in stop trials, or within the stopping area in the hit trials. In addition, if the movement duration was longer than 4.0 s, the trial was timed-out and had to be repeated. After each trial, the participant's hand was passively moved back to the start position by the vBOT, while the feedback of the previous trial was provided on screen (Fig 5D). All movements were self-paced, with short breaks provided every 208 trials, and a longer break (5-10 minutes) provided at the half-way point of the experiment.

Experimental paradigm

Participants performed reaching movements in four conditions – blocked stop, blocked hit, mixed stop and mixed hit – that were part of a single experiment. Across these conditions, participants were required to either reach to the target and stop there (the stop conditions), or to reach through the target and stop in the designated stopping area (hit conditions). In order to easily cue the distinction between the hit and stop conditions, the two types of trials had small visual differences. For the hit condition participants were presented with a red target (a red circle of 1.2 cm diameter) and a rectangular blue stopping area of dimensions 15 cm by 4 cm, centred 5 cm beyond the target (Fig 5C). For the stop condition participants were presented with a target that was otherwise identical to the target in hit condition, but was blue in colour, and with a horizontal, 15 cm wide red line, that was placed 3 cm before the target (Fig 5B). While this line had no functional interaction with the experiment, it allowed us to consistently instruct the participants to always perform reaching movements so that they intercept the red element in the workspace, and stop within the blue element.

In order to probe the visuomotor feedback responses of participants, during some reaching movements we briefly perturbed the target by shifting it 2.0 cm laterally for 250 ms before returning back to the original position (Fig 2A). These perturbed trials were always performed within the virtual mechanical channel, where participants were free to move along the line between the start position and the target, but were laterally

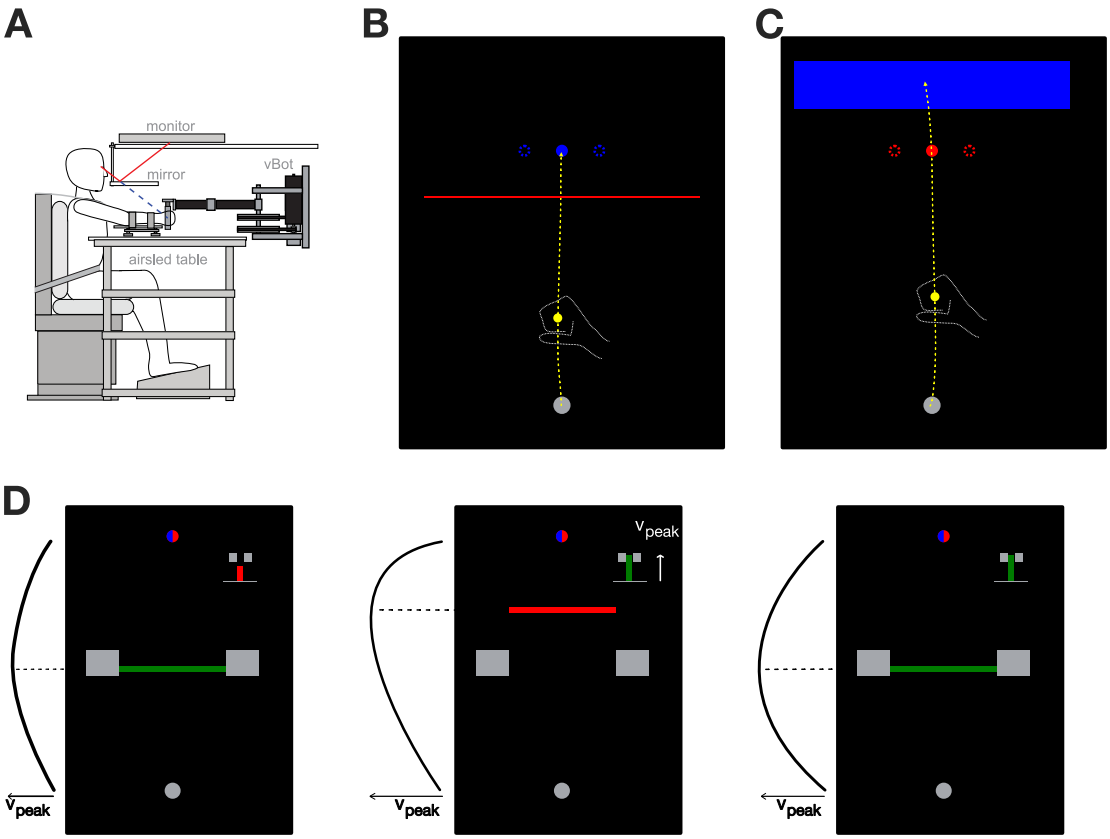


Fig 5. Experimental setup. **A.** Participants controlled a yellow cursor by moving a robotic handle. The cursor was projected via a screen-mirror system directly into the plane of the participant's hand. Figure copyright 2008 Society for Neuroscience. **B.** Stop condition. Participants were instructed to reach with the cursor through a red line and stop within the blue target. Target perturbations were occasionally induced via target jumps of 2 cm laterally. **C.** Hit condition. Participants were instructed to reach through the red target and stop within the blue area. Target perturbations (2 cm laterally) were again induced on random trials. **D.** Visual feedback was presented after each trial. Participants were shown the workspace with the start position and the target still present. In addition, two indicators were displayed. A bar chart at the top-right part of the workspace scaled proportionally with the absolute peak velocity, and was green if the velocity was within the required range as indicated by two grey brackets. A horizontal bar indicating the actual forward location where this peak velocity was achieved was displayed between the start and target positions. This bar was green if the peak location matched the experimental requirements, indicated by two large rectangular blocks. If both location and peak amplitude criteria were successfully fulfilled, participants were rewarded with one point. If at least one of the two criteria was not fulfilled, the respective indicator turned red instead of green, and no point was provided. In both hit and stop experiments participants were instructed to move through the red workspace element and stop at the blue, and were rewarded with one point if they both intercepted the target and fulfilled both velocity requirements.

constrained by a virtual viscoelastic wall with stiffness of 2 N/m and damping 4000 Ns/m [7, 40, 52]. As the perturbations were always task-irrelevant, this channel therefore did not obstruct participants to complete the trial. However, as participants still produced involuntary feedback responses due to the target shift, the virtual channel allowed us to record the forces that participants produced due to the perturbations and measure the intensities of the visuomotor feedback responses.

For each type of movement (i.e. hit or stop) there was a total of 11 different perturbations. Ten of these perturbations were cued during the reaching movement as participants crossed one of the five perturbation onset locations, equally spaced between the start position and the target position (4.2, 8.3, 12.5, 16.7, and 20.8 cm from the centre of the start position). At all of these five locations the target could either shift to the left or to the right. In addition, one zero-magnitude perturbation was also included, where the movement was simply performed within the channel without any target shift in order to probe the force profile of the natural movement. Finally, in addition to the perturbation trials we also included non-perturbed trials where participants simply reached towards the target without any target perturbation and without the virtual channel constraining the hand.

In order to present the different perturbations in a balanced manner, we combined different types of trials in blocks of 16 trials. One block of 16 trials contained 11 perturbed trials (5 perturbation onset locations x 2 directions, plus one neutral movement in the force channel), and 5 non-perturbed movements without the force channel. Each of the four experimental conditions consisted of 26 such blocks, with the order of trials fully randomised within each block, resulting in 416 trials per condition and 1664 trials overall.

In the first half of the experiment, participants were always presented with the two blocked-design conditions (blocked hit and blocked stop), with the order of the conditions balanced across the population of participants. That is, each participant started with 416 trials of stop trials, followed by 416 hit trials or vice-versa. In the second half of the experiment, the two final conditions – mixed hit and mixed stop – were presented in a pseudo-random order within the same blocks. While individual trials within mixed conditions were identical to the individual trials within the respective blocked conditions, they were now presented in a pseudo-randomised order. Specifically, the remaining 832 trials were divided into 26 blocks of 32 trials, with each block consisting of 16 hit and 16 stop trials fully randomised within this block. Such randomisation resulted in a percentage split where 52% of trials were presented after a condition switch, 26% of trials were presented after exactly one trial of the same condition, 12% – after exactly two trials of the same condition, and larger clusters with diminishing frequency.

Feedback regarding movement kinematics

In theory, the movements in hit condition could be interpreted as the movements, where the goal is to go through the via-point (the red target) and stop at the blue stopping area. As a result, such movements could simply be treated by participants as the stop movements with longer movement distance and a less restrictive target. Typically for such reaching movements, humans would produce a velocity profile that is bell-shaped, with peak velocity near the middle of the movement, and therefore further along the movement than in the stop condition. In order to avoid such differences and keep the velocity profiles comparable between the two conditions, we provided the task-relevant feedback on the velocity profiles, specifically the peak velocity and peak velocity location, to our participants (Fig 5D).

Independent of the experimental condition, participants were required to produce the movements with the peak velocity of 60 cm/s $\pm$ 8 cm/s, and the peak velocity location

within 11.25 cm - 13.75 cm movement distance (or 45%-55% of the distance between the start location and the target). The peak velocity was indicated as the small bar chart at the top-right of the screen, with the required velocity range indicated by two grey brackets. If the velocity target was matched, the bar chart turned green, otherwise it was red. Similarly, the peak velocity location was shown as a horizontal bar, centred around the movement distance where the peak velocity was reached. If this location was within the target range (also indicated by grey brackets), it was displayed as green, otherwise it was red. Participants were rewarded one point if both velocity requirements were successfully met, and the cursor intercepted the target during the movement.

Data Analysis and Code Availability

All data was pre-processed for the analysis in MATLAB 2017b: force and kinematic time series were low-pass filtered with a tenth-order zero-phase-lag Butterworth filter with 15 Hz cutoff and resampled at 1 kHz to account for an occasional missed sample during the signal recording. All subsequent analysis was performed in Python 3.9.4 and JASP v0.14.1 [83]. First, raw visuomotor feedback intensities were calculated from the force responses, recorded after the induction of a target perturbation. Specifically, for every perturbation trial we averaged the lateral force response over a time window of 180 ms - 230 ms after the onset of the perturbation, and subtracted a neutral force profile over the matching time window. This method and the particular time window has now been used in numerous studies to calculate the intensity of the early involuntary visuomotor feedback response [7, 25, 26, 40, 41, 50, 51]. As the direction of the response differed based on the perturbation direction, we reversed the direction of the intensities of responses to the leftward perturbations, so that positive intensities always indicate movements in correct direction, and grouped all intensities by the perturbation onset location. Second, we normalised mean feedback responses between 0 and 1 for each participant in order to avoid the group effect being biased towards participants with stronger responses. Finally, in our analysis the start of all movements was defined as the last time sample where the cursor is still within the area of the start circle, and the end of the movement was defined as the last time sample before the cursor enters the target circle. Time-to-target values were extracted from the data for every perturbation trial by subtracting the perturbation onset time from the movement end time.

In this article we provide two types of statistical analysis: the conventional frequentist statistics, as well as complementary Bayesian analysis that is presented as Bayesian factors [53], which instead of a simple hypothesis testing provides evidence for or against the null hypothesis. As a result, among other things, Bayesian analysis allows us to distinguish between accumulating evidence for the null hypothesis, and simply lacking evidence in either direction due to low power or small sample size.

All the Jupyter notebooks for the data analysis, pre-processed experimental data and statistical analysis conducted in this article are available at `https://doi.org/10.6084/m9.figshare.17113904.v1`.

Computational modelling

In this study we formulated our initial hypothesis about the feedback control mechanisms in humans by first simulating the behaviour of the optimal feedback controller (OFC). Specifically, we used a finite-horizon linear-quadratic regulator framework – a relatively simple OFC that assumes perfect sensory input, as well as no control-dependent noise, while still being able to capture a significant part of the variance of human reaching movements [38, 84]. In order to model the feedback behaviour of our human participants, we first simulated virtual movements of a point mass with $m = 1$ kg, and an intrinsic muscle damping $b = 0.1$ Ns/m. This point mass

was controlled in two dimensions by two orthogonal force actuators that simulated muscles, and regulated by a control signal u_t via a first-order low-pass filter with a time constant $\tau = 0.06$ s. At time t within the movement, such system could be described by the state transition model:

$$x_{t+1} = Ax_t + B(u_t + \xi_t), \tag{1}$$

where A is a state transition matrix, B is a control matrix, and ξ_t is additive control noise. For one spacial dimension A and B are defined in discrete time as:

$$A = \begin{bmatrix} 1 & \delta t & 0 \\ 0 & 1 - b\delta t/m & \delta t/m \\ 0 & 0 & 1 - \delta t/\tau \end{bmatrix}$$

,

$$B = \begin{bmatrix} 0 \\ \delta t/\tau \\ 0 \end{bmatrix}$$

Finally, to simulate our model in discrete time we used the sampling rate $\delta t = 0.01$ s

State x_t exists in the Cartesian plane and consists of position $\vec{p}$, velocity $\vec{v}$ and force $\vec{f}$ (two dimensions each). The control signal u_t is produced via the feedback control law:

$$u_t = -Lx_t \tag{2}$$

where L is a matrix of optimal feedback control gains, obtained by optimising the performance index (also known as the cost function):

$$J = \sum_{t=0}^{N} x_t^T Q_t x_t + u_t^T R_t u_t = \sum_{t=0}^{N} \omega_{p,t}(\vec{p_t} - \vec{p^*})^2 + \omega_{v,t}||\vec{v}_t||^2 + \omega_{f,t}||\vec{f}_t||^2 + \omega_{r,t}||u_t||^2. \tag{3}$$

Here $x_t^T Q x_t$ and $u_t^T R u_t$ are two components of the total cost, known as state-cost and a control-cost respectively. In addition, ω_p, ω_v and ω_f are position, velocity and force state cost parameters, $\vec{p^*}$ is a target position, ω_r is the activation cost parameter and N is the duration of the movement, here required as a model input. Within the finite-horizon formulation, the cost parameters can be non-stationary and thus be different for every time-point. However, in our simulations we set $Q = 0$ for $t \neq N$, consistent with [37, 85].

In this study we simulate three different controllers that we call stop, hit and long-stop. While the stop and long-stop controllers are derived from the identical set of costs state-costs Q, they are used for slightly different movements (25 cm and 700 ms for stop, 28 cm and 800 ms for long-stop). We used $\omega_p = 1.5$, $\omega_v = 1$ and $\omega_f = 0.1$ as the values for the state cost parameters in this model, and the activation cost $R = 3 \times 10^{-6}$. Furthermore, in order to better match the forward velocity profiles, we also introduced a non-stationarity in the activation cost R of the long stop movement, where the total integral of the activation cost over the movement is not changed, but this cost develops over time during the movement. Specifically, at a time t in the trial, the activation cost for the long-stop movement was computed by:

$$R_{long-stop}(t) = RC(t), \tag{4}$$

where

$$C(t) \propto \exp\left(p\frac{t+q}{r}\right), \tag{5}$$

and the mean of $C(t)$ equals 1 for the duration of the trial, so that $R_{long-stop}$ produces the same amount of activation as R over the duration of the trial. Here $p = 1$, $q =$ -1000 and $r = 65$ are constants, fit via trial and error in order to produce the forward velocity profile of long-stop condition that matches the velocity of stop and hit conditions. We have previously shown that such modulation only affects the kinematics of the movement, but does not change the feedback responses when expressed against the time-to-target [38]. On the other hand, in order to incentivise the hit controller to produce faster movements at the target, we reduced the cost parameters for terminal velocity and terminal force by a factor of 50. As a result, such controller produced hit-like movements that were aimed directly at a target, positioned at 25 cm distance, over 620 ms, which matched the kinematics of the long-stop controller over this movement segment.

Finally, for each controller we simulated feedback response intensity profiles along the movement, which we then used to compare the control policies predicted by each controller. To do so, we induced lateral target perturbations of 2 cm magnitude during the simulated movement to the target and recorded the corrective force, produced by each controller as a result of these perturbations. While in the experimental study we only induced such perturbations at five different onsets due to practical reasons, in our simulations we could perturb the movements at every point in time and fully map the response intensity profiles over the movement. Thus, for each model we simulated different movements with perturbations at each movement time-step (i.e. every 10 ms), with one perturbation only happening once per movement. In addition, to simulate the visuomotor delay that is present in humans, we delayed the onset of each perturbation by 150 ms, so that for the perturbation triggered at time t, the target is shifted at time $t + 150$ ms. We then averaged the force, produced by our model over a time interval 10 ms – 60 ms after the target was shifted (160 ms – 210 ms after the perturbation was triggered), representing the visuomotor response window of 180 ms – 230 ms in human subjects. Note that we used an earlier window for the model simulations than for the human subjects as the responses in the simulations ramp up fast due to muscles simplified to a single low pass filter.

Acknowledgements

We thank Hanna Hoogen, Isabelle Hoxha and Oliver Gerke for contributions to preliminary projects related to this manuscript. We thank Clara Günter, Jing Zhang, Sae Franklin, and Marion Forano for their feedback on this manuscript.

References

1. Houk JC. An Assessment of Stretch Reflex Function. Progress in Brain Research. 1976;44(C):303–314. doi:10.1016/S0079-6123(08)60741-4.

2. Pruszynski JA, Scott SH. Optimal feedback control and the long-latency stretch response. Experimental Brain Research. 2012;218(3):341–359. doi:10.1007/s00221-012-3041-8.

3. Dimitriou M. Enhanced Muscle Afferent Signals during Motor Learning in Humans. Current Biology. 2016;26(8):1062–1068. doi:10.1016/j.cub.2016.02.030.

4. Barr CC, Schultheis LW, Robinson DA. Voluntary, non-visual control of the human vestibulo-ocular reflex. Acta Oto-Laryngologica. 1976;81(5-6):365–375. doi:10.3109/00016487609107490.

5. Tabak S, Collewijn H. Human vestibulo-ocular responses to rapid, helmet-driven head movements. Experimental Brain Research. 1994;102(2):367–378. doi:10.1007/BF00227523.

6. Prablanc C, Martin O. Automatic control during hand reaching at undetected two-dimensional target displacements. Journal of Neurophysiology. 1992;67(2):455–469. doi:10.1152/jn.1992.67.2.455.

7. Franklin DW, Wolpert DM. Specificity of Reflex Adaptation for Task-Relevant Variability. Journal of Neuroscience. 2008;28(52):14165–14175. doi:10.1523/JNEUROSCI.4406-08.2008.

8. Izawa J, Shadmehr R. On-Line Processing of Uncertain Information in Visuomotor Control. Journal of Neuroscience. 2008;28(44):11360–11368. doi:10.1523/JNEUROSCI.3063-08.2008.

9. Knill DC, Bondada A, Chhabra M. Flexible, Task-Dependent Use of Sensory Feedback to Control Hand Movements. Journal of Neuroscience. 2011;31(4):1219–1237. doi:10.1523/JNEUROSCI.3522-09.2011.

10. Reynolds RF, Day BL. Direct visuomotor mapping for fast visually-evoked arm movements. Neuropsychologia. 2012;50(14):3169–3173. doi:10.1016/j.neuropsychologia.2012.10.006.

11. Franklin S, Wolpert DM, Franklin DW. Visuomotor feedback gains upregulate during the learning of novel dynamics. Journal of Neurophysiology. 2012;108(2):467–478. doi:10.1152/jn.01123.2011.

12. Pruszynski JA, Gribble PL, Corneil BD. A rapid visuomotor response on the human upper limb is selectively influenced by implicit motor learning. Journal of Neurophysiology. 2018;121(1):85–95. doi:10.1152/jn.00720.2018.

13. Zhang Y, Brenner E, Duysens J, Verschueren S, Smeets JBJ. Postural responses to target jumps and background motion in a fast pointing task. Experimental Brain Research. 2018;236(6):1573–1581. doi:10.1007/s00221-018-5222-6.

14. Saijo N, Murakami I, Nishida S, Gomi H. Large-Field Visual Motion Directly Induces an Involuntary Rapid Manual Following Response. Journal of Neuroscience. 2005;25(20):4941–4951. doi:10.1523/JNEUROSCI.4143-04.2005.

15. Baram Y, Miller A. Auditory feedback control for improvement of gait in patients with Multiple Sclerosis. Journal of the Neurological Sciences. 2007;254(1-2):90–94. doi:10.1016/j.jns.2007.01.003.

16. Oscari F, Secoli R, Avanzini F, Rosati G, Reinkensmeyer DJ. Substituting auditory for visual feedback to adapt to altered dynamic and kinematic environments during reaching. Experimental Brain Research. 2012;221(1):33–41. doi:10.1007/s00221-012-3144-2.

17. Saunders JA, Knill DC. Humans use continuous visual feedback from the hand to control fast reaching movements. Experimental Brain Research. 2003;152(3):341–352. doi:10.1007/s00221-003-1525-2.

18. Saunders JA, Knill DC. Humans use continuous visual feedback from the hand to control both the direction and distance of pointing movements. Experimental Brain Research. 2005;162(4):458–473. doi:10.1007/s00221-004-2064-1.

19. Franklin S, Wolpert DM, Franklin DW. Rapid visuomotor feedback gains are tuned to the task dynamics. Journal of Neurophysiology. 2017; p. jn.00748.2016. doi:10.1152/jn.00748.2016.

20. Oostwoud Wijdenes L, Brenner E, Smeets JBJ. Fast and fine-tuned corrections when the target of a hand movement is displaced. Experimental Brain Research. 2011;214(3):453–462. doi:10.1007/s00221-011-2843-4.

21. Oostwoud Wijdenes L, Brenner E, Smeets JBJ. Comparing online adjustments to distance and direction in fast pointing movements. Journal of Motor Behavior. 2013;45(5):395–404. doi:10.1080/00222895.2013.815150.

22. Day BL, Lyon IN. Voluntary modification of automatic arm movements evoked by motion of a visual target. Experimental Brain Research. 2000;130(2):159–168. doi:10.1007/s002219900218.

23. Diedrichsen J. Optimal Task-Dependent Changes of Bimanual Feedback Control and Adaptation. Current Biology. 2007;17(19):1675–1679. doi:10.1016/j.cub.2007.08.051.

24. Maeda RS, Cluff T, Gribble PL, Pruszynski JA. Feedforward and feedback control share an internal model of the arm's dynamics. Journal of Neuroscience. 2018;38(49):10505–10514. doi:10.1523/JNEUROSCI.1709-18.2018.

25. Cross KP, Cluff T, Takei T, Scott SH. Visual Feedback Processing of the Limb Involves Two Distinct Phases. Journal of Neuroscience. 2019;39(34):6751–6765. doi:10.1523/jneurosci.3112-18.2019.

26. Česonis J, Franklin DW. Time-to-target simplifies optimal control of visuomotor feedback responses. eNeuro. 2020;7(2):1–17. doi:10.1523/ENEURO.0514-19.2020.

27. Shadmehr R, Brashers-Krug T, Mussa-Ivaldi FA. Interference in Learning Internal Models of Inverse Dynamics in Humans. Advances in Neural Information Processing Systems 7. 1995;7:1117–1124.

28. Bock O, Schneider S, Bloomberg J. Conditions for interference versus facilitation during sequential sensorimotor adaptation. Experimental Brain Research. 2001;138(3):359–365. doi:10.1007/s002210100704.

29. Sing GC, Smith MA. Reduction in learning rates associated with anterograde interference results from interactions between different timescales in motor adaptation. PLoS Computational Biology. 2010;6(8). doi:10.1371/journal.pcbi.1000893.

30. Howard IS, Wolpert DM, Franklin DW. The effect of contextual cues on the encoding of motor memories. Journal of Neurophysiology. 2013;109(10):2632–2644. doi:10.1152/jn.00773.2012.

31. Forano M, Franklin DW. Timescales of motor memory formation in dual-adaptation. PLoS Computational Biology. 2020;16(10):1–33. doi:10.1371/journal.pcbi.1008373.

32. Hirashima M, Nozaki D. Distinct motor plans form and retrieve distinct motor memories for physically identical movements. Current Biology. 2012;22(5):432–436. doi:10.1016/j.cub.2012.01.042.

33. Howard IS, Ingram JN, Franklin DW, Wolpert DM. Gone in 0.6 seconds: The encoding of motor memories depends on recent sensorimotor states. Journal of Neuroscience. 2012;32(37):12756–12768. doi:10.1523/JNEUROSCI.5909-11.2012.

34. Sarwary AME, Stegeman DF, Selen LPJ, Medendorp WP. Generalization and transfer of contextual cues in motor learning. Journal of Neurophysiology. 2015;114(3):1565–1576. doi:10.1152/jn.00217.2015.

35. Howard IS, Wolpert DM, Franklin DW. The value of the follow-through derives from motor learning depending on future actions. Current Biology. 2015;25(3):397–401. doi:10.1016/j.cub.2014.12.037.

36. Sheahan HR, Franklin DW, Wolpert DM. Motor Planning, Not Execution, Separates Motor Memories. Neuron. 2016;92(4):773–779. doi:10.1016/j.neuron.2016.10.017.

37. Liu D, Todorov E. Evidence for the Flexible Sensorimotor Strategies Predicted by Optimal Feedback Control. Journal of Neuroscience. 2007;27(35):9354–9368. doi:10.1523/JNEUROSCI.1110-06.2007.

38. Česonis J, Franklin DW. Mixed-horizon optimal feedback control as a model of human movement. Neurons, Behavior, Data analysis, and Theory. 2021; p. 1–36. doi:10.51628/001c.29674.

39. Poscente SV, Peters RM, Cashaback JGA, Cluff T. Rapid Feedback Responses Parallel the Urgency of Voluntary Reaching Movements. Neuroscience. 2021;doi:10.1016/j.neuroscience.2021.07.014.

40. Dimitriou M, Wolpert DM, Franklin DW. The Temporal Evolution of Feedback Gains Rapidly Update to Task Demands. Journal of Neuroscience. 2013;33(26):10898–10909. doi:10.1523/JNEUROSCI.5669-12.2013.

41. Franklin DW, Reichenbach A, Franklin S, Diedrichsen J. Temporal Evolution of Spatial Computations for Visuomotor Control. Journal of Neuroscience. 2016;36(8):2329–2341. doi:10.1523/JNEUROSCI.0052-15.2016.

42. Nashed JY, Crevecoeur F, Scott SH. Influence of the behavioral goal and environmental obstacles on rapid feedback responses. Journal of Neurophysiology. 2012;108(4):999–1009. doi:10.1152/jn.01089.2011.

43. Orban de Xivry JJ. Trial-to-Trial Reoptimization of Motor Behavior Due to Changes in Task Demands Is Limited. PLoS ONE. 2013;8(6). doi:10.1371/journal.pone.0066013.

44. Lowrey CR, Nashed JY, Scott SH. Rapid and flexible whole body postural responses are evoked from perturbations to the upper limb during goal-directed reaching. Journal of Neurophysiology. 2017;117(3):1070–1083. doi:10.1152/jn.01004.2015.

45. de Comite A, Crevecoeur F, Lefèvre P. Online modification of goal-directed control in human reaching movements. Journal of neurophysiology. 2021;125(5):1883–1898. doi:10.1152/jn.00536.2020.

46. Reichenbach A, Costello A, Zatka-Haas P, Diedrichsen J. Mechanisms of responsibility assignment during redundant reaching movements. Journal of Neurophysiology. 2013;109(8):2021–2028. doi:10.1152/jn.01052.2012.

47. Reichenbach A, Franklin DW, Zatka-Haas P, Diedrichsen J. A dedicated binding mechanism for the visual control of movement. Current Biology. 2014;24(7):780–785. doi:10.1016/j.cub.2014.02.030.

48. Reichenbach A, Diedrichsen J. Processing reafferent and exafferent visual information for action and perception. Journal of Vision. 2015;15(8):1–12. doi:10.1167/15.8.11.

49. Gallivan JP, Logan L, Wolpert DM, Flanagan JR. Parallel specification of competing sensorimotor control policies for alternative action options. Nature Neuroscience. 2016;19(2):320–326. doi:10.1038/nn.4214.

50. de Brouwer AJ, Jarvis T, Gallivan JP, Flanagan JR. Parallel Specification of Visuomotor Feedback Gains during Bimanual Reaching to Independent Goals. Eneuro. 2017;4(2):ENEURO.0026–17.2017. doi:10.1523/ENEURO.0026-17.2017.

51. de Brouwer AJ, Gallivan JP, Flanagan JR. Visuomotor feedback gains are modulated by gaze position. Journal of Neurophysiology. 2018;120(5):2522–2531. doi:10.1152/jn.00182.2018.

52. Scheidt RA, Reinkensmeyer DJ, Conditt MA, Rymer WZ, Mussa-Ivaldi FA. Persistence of motor adaptation during constrained, multi-joint, arm movements. Journal of neurophysiology. 2000;84(2):853–862. doi:0022-3077/00.

53. Raftery AE, Kass RE. Bayes Factors. Journal of the American Statistical Association. 1995;90(430):773–795.

54. Harris CM, Wolpert DM. Signal-dependent noise determines motor planning. Nature. 1998;394:780.

55. Todorov E, Jordan MI. Optimal feedback control as a theory of motor coordination. Nature Neuroscience. 2002;5(11):1226–1235. doi:10.1038/nn963.

56. Verstynen T, Sabes PN. How Each Movement Changes the Next: An Experimental and Theoretical Study of Fast Adaptive Priors in Reaching. Journal of Neuroscience. 2011;31(27):10050–10059. doi:10.1523/JNEUROSCI.6525-10.2011.

57. Rigoux L, Guigon E. A Model of Reward- and Effort-Based Optimal Decision Making and Motor Control. PLoS Computational Biology. 2012;8(10). doi:10.1371/journal.pcbi.1002716.

58. Berniker M, Penny S. A normative approach to neuromotor control. Biological Cybernetics. 2019;113(1-2):83–92. doi:10.1007/s00422-018-0777-7.

59. Wada Y, Kawabata Y, Kotosaka S, Yamamoto K, Kitazawa S, Kawato M. Acquisition and contextual switching of multiple internal models for different viscous force fields. Neuroscience Research. 2003;46(3):319–331. doi:10.1016/S0168-0102(03)00094-4.

60. Lee JY, Schweighofer N. Dual Adaptation Supports a Parallel Architecture of Motor Memory. Journal of Neuroscience. 2009;29(33):10396–10404. doi:10.1523/JNEUROSCI.1294-09.2009.

61. Howard IS, Ford C, Cangelosi A, Franklin DW. Active lead-in variability affects motor memory formation and slows motor learning. Scientific Reports. 2017;7(1):1–12. doi:10.1038/s41598-017-05697-z.

62. Forano M, Schween R, Taylor JA, Hegele M, Franklin DW. Direct and indirect cues can enable dual adaptation, but through different learning processes. Journal of Neurophysiology. 2021;126(5):1490–1506. doi:10.1152/jn.00166.2021.

63. Wolpert DM, Kawato M. Multiple paired forward and inverse models for motor control. Neural Networks. 1998;11(7-8):1317–1329. doi:10.1016/S0893-6080(98)00066-5.

64. Ahmadi-Pajouh MA, Towhidkhah F, Shadmehr R. Preparing to reach: Selecting an adaptive long-latency feedback controller. Journal of Neuroscience. 2012;32(28):9537–9545. doi:10.1523/JNEUROSCI.4275-11.2012.

65. Hayashi T, Yokoi A, Hirashima M, Nozaki D. Visuomotor Map Determines How Visually Guided Reaching Movements are Corrected Within and Across Trials. eNeuro. 2016;3(3):1–13. doi:10.1523/ENEURO.0032-16.2016.

66. Maeda RS, Gribble PL, Pruszynski JA. Learning New Feedforward Motor Commands Based on Feedback Responses. Current Biology. 2020;30(10):1941–1948.e3. doi:10.1016/j.cub.2020.03.005.

67. Voudouris D, Fiehler K. Dynamic temporal modulation of somatosensory processing during reaching. Scientific Reports. 2021;11(1):1–12. doi:10.1038/s41598-021-81156-0.

68. Wagner MJ, Smith MA. Shared internal models for feedforward and feedback control. Journal of Neuroscience. 2008;28(42):10663–10673. doi:10.1523/JNEUROSCI.5479-07.2008.

69. Maeda RS, Kersten R, Pruszynski JA. Shared internal models for feedforward and feedback control of arm dynamics in non-human primates. European Journal of Neuroscience. 2021;53(5):1605–1620. doi:https://doi.org/10.1111/ejn.15056.

70. Cluff T, Scott SH. Rapid Feedback Responses Correlate with Reach Adaptation and Properties of Novel Upper Limb Loads. Journal of Neuroscience. 2013;33(40):15903–15914. doi:10.1523/JNEUROSCI.0263-13.2013.

71. Franklin S, Franklin DW. Feedback Gains modulate with Motor Memory Uncertainty. Neurons, Behavior, Data analysis, and Theory. 2021;5(2):1–28. doi:10.51628/001c.22336.

72. Coltman SK, Gribble PL. Time course of changes in the long-latency feedback response parallels the fast process of short-term motor adaptation. Journal of Neurophysiology. 2020;124(2):388–399. doi:10.1152/jn.00286.2020.

73. Gritsenko V, Kalaska JF. Rapid online correction is selectively suppressed during movement with a visuomotor transformation. Journal of Neurophysiology. 2010;104(6):3084–3104. doi:10.1152/jn.00909.2009.

74. Yousif N, Diedrichsen J. Structural learning in feedforward and feedback control. Journal of Neurophysiology. 2012;108(9):2373–2382. doi:10.1152/jn.00315.2012.

75. Kasuga S, Telgen S, Ushiba J, Nozaki D, Diedrichsen J. Learning feedback and feedforward control in a mirror-reversed visual environment. Journal of Neurophysiology. 2015;114(4):2187–2193. doi:10.1152/jn.00096.2015.

76. Oostwoud Wijdenes L, Van Beers RJ, Medendorp WP. Vestibular modulation of visuomotor feedback gains in reaching. Journal of Neurophysiology. 2019;122(3):947–957. doi:10.1152/jn.00616.2018.

77. Izawa J, Rane T, Donchin O, Shadmehr R. Motor Adaptation as a Process of Reoptimization. Journal of Neuroscience. 2008;28(11):2883–2891. doi:10.1523/JNEUROSCI.5359-07.2008.

78. Crevecoeur F, Kurtzer I, Scott SH. Fast corrective responses are evoked by perturbations approaching the natural variability of posture and movement tasks. Journal of Neurophysiology. 2012;107(10):2821–2832. doi:10.1152/jn.00849.2011.

79. Howard IS, Franklin S, Franklin DW. Asymmetry in kinematic generalization between visual and passive lead-in movements are consistent with a forward model in the sensorimotor system. PLoS ONE. 2020;15(1):1–21. doi:10.1371/journal.pone.0228083.

80. Crevecoeur F, Kurtzer I, Bourke T, Scott SH. Feedback responses rapidly scale with the urgency to correct for external perturbations. Journal of Neurophysiology. 2013;110(6):1323–1332. doi:10.1152/jn.00216.2013.

81. Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia. 1971;9(1):97–113. doi:10.1016/0028-3932(71)90067-4.

82. Howard IS, Ingram JN, Wolpert DM. A modular planar robotic manipulandum with end-point torque control. Journal of Neuroscience Methods. 2009;181(2):199–211. doi:10.1016/j.jneumeth.2009.05.005.

83. JASP Team. JASP (Version 0.14.1)[Computer software]; 2021. Available from: `https://jasp-stats.org/`.

84. Kuo AD. An optimal control model for analyzing human postural balance. IEEE Transactions on Biomedical Engineering. 1995;42(1):87–101. doi:10.1109/10.362914.

85. Todorov E. Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. Neural Computation. 2005;17(5):1084–1108. doi:10.1162/0899766053491887.

S1 Text Initial learning of feedback controllers. Here we verify whether our participants developed different feedback controllers for hit and stop tasks over the blocked schedule, or if these controllers were innate. To do so, we analyse the visuomotor responses in the first few blocks of the study, showing that these responses can be considered innate.

S2 Text Effect of condition clustering in mixed schedule. Visuomotor responses were analysed in trials immediately following the condition switch (hit to stop or stop to hit) in mixed schedule. Analysis shows same regulation as in the entirety of the mixed schedule, implying rapid switching.

S1 Fig OFC model simulations in channel trials. Model simulations performed in cbannel trials, instead of free movement responses. **A.** Velocity profiles for stop (blue), hit (red) and long-stop (green) conditions. **B.** Model simulations of feedback intensities as a function of time-to-target and **C.** position for the three conditions. Simulations in the channel trials qualitatively predict the same regulation as do the simulations of free movements.

S1 Text. Initial learning of feedback controllers

One of our main goals in this study was to test whether participants could or could not rapidly switch between two feedback controllers. As a result, we selected two tasks (stop and hit) that we expected participants would have innate controllers for. However, it is possible that these controllers were not in fact innate, and participants only developed them over the extended exposure to different tasks during blocked schedule. We therefore analysed the feedback responses of our participants over the first three blocks of each condition qualitatively (Fig S1). Even within the early portion of the experiment, we observe a clear distinction between feedback responses in hit and stop conditions, suggesting that controllers were indeed innate to our participants, or at least developed rapidly enough to be considered as innate for the purposes of our study.

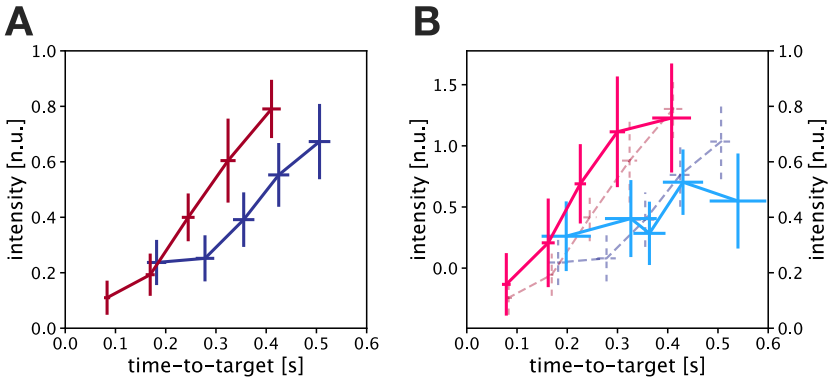


Fig S1. Visuomotor feedback intensities during the first three blocks of the experiment. **A.** Normalised feedback intensity profiles of participants in blocked stop and blocked hit conditions, expressed against time-to-target. Participants produce stronger responses at matching time-to-target in the hit condition, consistent with simulation results for hit and stop (same as Fig 3E). **B.** Normalised feedback intensity profiles of participants in blocked stop and blocked hit conditions during the first three blocks of each condition, expressed against time-to-target (solid lines, primary axis). Dashed lines represent the responses over the entirety of the blocked schedule (secondary axis). Error bars represent 95% CI

S2 Text. Effect of condition clustering in mixed schedule

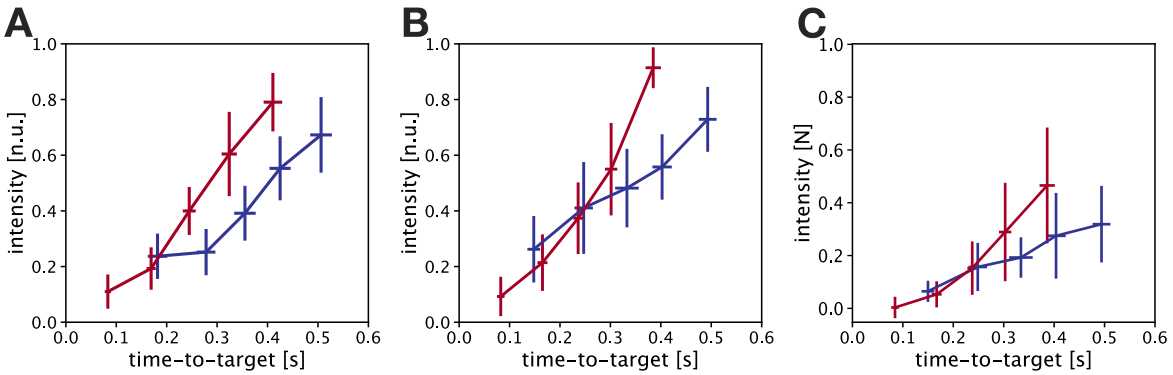
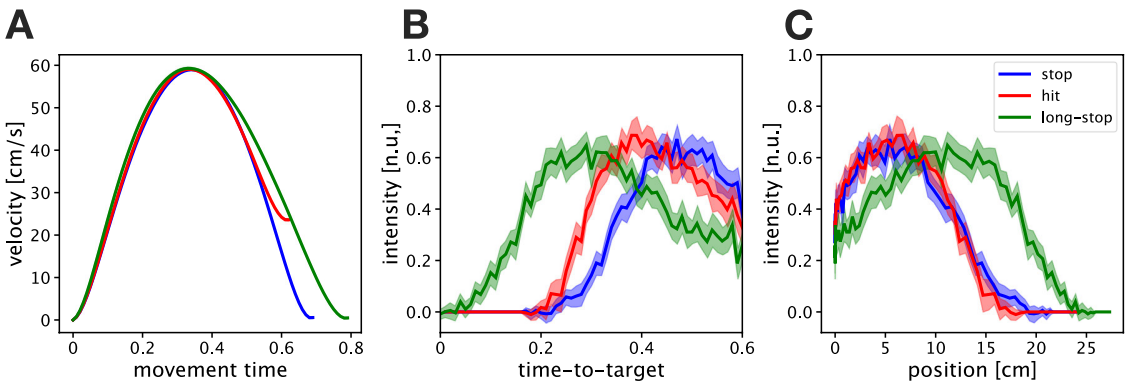


Fig S2. Visuomotor feedback intensities in the trials after the switch of condition. **A.** Normalised feedback intensity profiles of participants in blocked stop and blocked hit conditions, expressed against time-to-target. Participants produce stronger responses at matching time-to-target in the hit condition, consistent with simulation results for hit and stop (same as Fig 3E). **B.** Normalised feedback intensity profiles of participants in mixed stop and mixed hit conditions, only in trials immediately after condition switch, expressed against time-to-target. Similar regulation is observed as in the blocked schedule. **C.** Raw feedback intensity profiles of participants in mixed stop and mixed hit conditions, only in trials immediately after condition switch, expressed against time-to-target. Error bars represent 95% CI

In our study design we took deliberate care in shuffling the conditions in mixed schedule. Particularly, if we simply shuffled the 832 trials in the mixed schedule together, we are likely to observe large clusters of the same condition without switching, which may allow for adaptation of feedback controllers and thus would provide false support for the rapid switching hypothesis. In order to avoid such effects we used a pseudo-random design, where we divided the 832 trials in the mixed schedule into 26 consecutive blocks of 32 trials each. Every block contained 16 trials of hit condition, and 16 trials of stop condition, where the 16 trials contained 11 perturbation trials (5 perturbation onset locations $\times$ 2 directions, plus one zero-perturbation trial) and 5 free (null-field) trials. Within each block all trials were randomly shuffled, but one block had to be completed entirely before the next block started. Similar shuffling was also performed in the blocked schedule, only that each block contained 16 trials of a single condition, and all 26 blocks of one condition had to be completed before the opposite condition was first introduced.

Even under the aforementioned constraints of the randomisation there still remains a chance of condition clustering. In fact, on average in the mixed-schedule each participant experienced 52% of trials after condition switch, with remaining trials following the same condition as on the previous trial (26% after exactly one trial of the same condition, 12% after exactly two trials, and the larger clusters in diminishing quantity). As a result, we performed a control analysis where we compared the hit and stop conditions only in trials immediately after the condition switch (Figure S2). A one-way ANCOVA analysis of normalised visuomotor response intensity in the mixed schedule with condition as a fixed factor and time-to-target as the covariate still showed a significant main effect of condition (hit or stop, $F_{1,127} = 5.51$, $p = 0.02$), despite reduced statistical power. However, a similar Bayesian ANCOVA analysis only showed a weak tendency towards main effect for condition ($BF_{incl} = 1.96$). Both analyses showed a significant effect of time-to-target ($F_{1,127} = 94.3$, $p \ll 0.001$; $BF_{incl} = 2.86 \times 10^{1}3$). Thus, qualitatively and quantitatively the control behaviour of our participants during the first trials after condition switching matched that demonstrated for the whole

duration of the experiment, showing no effect of learning due to consecutive trials of the same condition. As this analysis uses only a subset of our data, these results are sensitive to noise, and thus the normalisation of the data, shown in Figure S2B could be affected by an occasional outlier. As a result, for completeness, and a clearer picture, present both normalised data (S2B), and raw data (S2C).


A
B
C
velocity [cm/s]
movement time
intensity [n.u.]
time-to-target
intensity [n.u.]
position [cm]
stop
hit
long-stop

3.7 Study IV

A Simulated Inverted Pendulum to Investigate Human Sensorimotor Control. © 2018 IEEE. Reprinted, with permission, from [80]

This study, authored by Justinas Česonis, Sae Franklin and David W. Franklin was published in the proceedings of *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. In this study we presented an inverted pendulum balancing task, that was completely virtual, and therefore allowed for flexible modifications of the design when studying the human control behaviour. In addition, we also modelled this control behaviour computationally by using a set of PD controllers, showing that our results can be best described by a PD control with a time delay of 150 ms.

Contributions

Justinas Česonis was the lead author in this research. In addition, Justinas Česonis designed and built the virtual inverted pendulum within the vBot system, implemented the experimental design, analysed the data and built the control models, as well as wrote the first draft of the manuscript. Sae Franklin collected the experimental data. Justinas Česonis, Sae Franklin and David W. Franklin designed the study, selected the analysis methods, interpreted the results and wrote the final version of the paper together.

Abstract

Sensorimotor control regulates balance and stability as well as adaptation to the external environment. We introduce the use of a simulated inverted pendulum to study human sensorimotor control, demonstrating that this system introduces similar control challenges to human subjects as a physical inverted pendulum. Participants exhibited longer stabilization of the system as the pendulum length between the hand and the center of mass increased while the required control input varied in a non-monotonic, yet predictable manner. Finally, we show that the experimental results can be modelled as a PD controller with a time delay of τ = 140 ms, matching the human visuomotor delay. Our results provide evidence of the importance of vision in a control of unstable systems and serve as a proof of concept of a simulated inverted pendulum.

A Simulated Inverted Pendulum to Investigate Human Sensorimotor Control

Justinas Česonis, Sae Franklin, and David W. Franklin, *Member, IEEE*

***Abstract*—Sensorimotor control regulates balance and stability as well as adaptation to the external environment. We introduce the use of a simulated inverted pendulum to study human sensorimotor control, demonstrating that this system introduces similar control challenges to human subjects as a physical inverted pendulum. Participants exhibited longer stabilization of the system as the pendulum length between the hand and the center of mass increased while the required control input varied in a non-monotonic, yet predictable manner. Finally, we show that the experimental results can be modelled as a PD controller with a time delay of τ = 140 ms, matching the human visuomotor delay. Our results provide evidence of the importance of vision in a control of unstable systems and serve as a proof of concept of a simulated inverted pendulum.**

I. Introduction

The inverted pendulum is a classic problem in control theory, often used as an assessment tool to test control strategies. This system is both unstable without control and contains nonlinear dynamics. The inverted pendulum is normally implemented with a pivot point mounted on a cart that moves horizontally under the control of a servo motor system. Here we implement this classic model within our virtual reality robotic system in order use this unstable system as a tool to assess human motor control and behaviour.

The use of the inverted pendulum system is also a classic approach to investigating human motor control [1]-[3]. Normally the subject is asked to control a fully mechanical inverted pendulum after training in order to investigate the delays and processes that govern this balancing control. In certain cases the inverted pendulum has been simulated in order to briefly suppress visual feedback and examine the predictive control strategies [1]. Here we fully simulate a virtual inverted pendulum in the horizontal plane in order to further investigate the processes of control. As the system is fully simulated, this allows the experimenter to control every aspect of the feedback (visual, haptic or temporal) in order to understand how each parameter is used by the sensorimotor control system.

Stability is an essential component of human motor control and learning [4]. However the mechanisms used by the human sensorimotor system vary depending on both the task and the effectors used [5], [6]. While reaching movements in an unstable environment have promoted the use of predictive co-contraction to modify the endpoint stiffness of the limb [7]-[9], balance of postural sway has strongly supported the use of feedback control for stabilization [10]-[12]. These differences have been suggested to arise due to the different timescales of the system [5], allowing for the use of delayed feedback for long lengths such as full body sway. The simulation of a virtual inverted pendulum allows for changes in the lengths of the pendulum to be investigated. The objective of this task is to actively balance the inverted pendulum by applying a force to the cart. Here we applied this one degree of freedom pendulum onto a two-dimensional virtual reality robotic system. This apparatus can apply visual and haptic feedback to the participants as they interact with the virtual cart and pendulum. Such a system allows manipulation of the supplied feedback in order to investigate the control strategies of the human subjects in this complex task.

J. Česonis and D.W. Franklin are with Neuromuscular Diagnostics, Department of Sports and Health Science, Technical University of Munich, Munich, 80992 Germany (phone: +49 89289 24536; e-mail: justinas.cesonis@tum.de, david.franklin@tum.de).

S. Franklin is with the Institute of Cognitive Systems, Department of Electrical and Computer Engineering, Technical University of Munich, Munich, 80333 Germany (e-mail: sae.franklin@tum.de).

II. Materials and Methods

A. Subjects

Six neurologically healthy, right-handed [13] human subjects (1 female) took part in the experiment (mean age 29.0 years). All subjects were naïve to the study purpose and provided written informed consent before participation. The study was approved by the institutional ethics committee at the Technical University of Munich.

B. Experimental apparatus

Participants performed a balancing task of an inverted pendulum simulated with a robotic manipulandum. Subjects were seated with their right arm resting on an airsled and grasping the endpoint handle of the vBOT robotic interface (Fig 1A). The vBOT is a custom made planar robotic interface that generates state-dependent forces on the hand at 1kHz [14] (Fig. 1A). A six-axis force transducer (ATI Nano 25; ATI Industrial Automation) measured the end-point forces applied on the robotic handle by the participants, while handle position was calculated from joint-position sensors (58SA; Industrial Encoders Direct). Position and force data were sampled at 1 kHz. Visual feedback was projected via a computer monitor and a mirror system to the plane of the movement in such a way that the direct visual feedback of the hand was prevented.

C. Experimental paradigm

The inverted pendulum was simulated in the x-y plane with the gravity acting in the negative y direction while corrective movements were performed in the x-axis. Mechanically the

978-1-5386-3646-6/18/$31.00 ©2018 IEEE

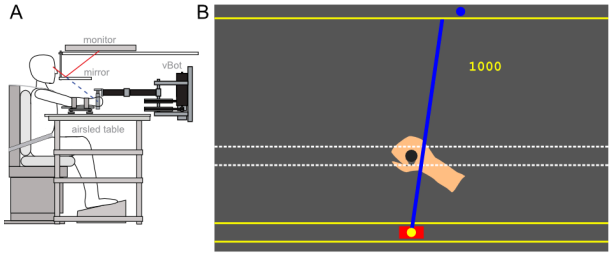


Figure 1. Experimental design. **A**, Subjects attempted to balance an inverted pendulum simulated using a planar robotic manipulandum providing both visual and haptic feedback. **B** A sample screenshot of an experimental trial. The circular cursor at the top of the screen provides visual feedback of the center of mass while the pendulum (blue line) is truncated at the top of the screen. The y-coordinate of the physical hand location (not visible to subjects) is offset with respect to the cart position.

pendulum was represented as a point mass (m = 1 kg) balanced at height (L) above a cart (M = 0.1 kg). The dynamic equations of motion describing the system are:

$$F_x = \ddot{x}(m\sin^2\theta + M) - mL\dot{\theta}^2\sin\theta + mg\sin\theta\cos\theta \quad (1)$$

$$\ddot{\theta} = (g\sin\theta - \ddot{x}\cos\theta)/L \quad (2)$$

where F_x is the lateral force applied by a pendulum on the cart, θ is the angle between the pendulum and the y-axis, x is the position of the cart and g is the gravitational acceleration constant.

The cart, controlled directly by the hand of a subject, was represented as a 1.5 cm by 3.0 cm red block. It was constrained to a single axis of motion in the x direction approximately 30 cm in front of participant's chest by a simulated mechanical channel (stiffness 4000 N/m; damping 2 Ns/m and maximum force value of 25 N). This channel was framed visually on the screen by two yellow lines of 1.0 mm thickness. Any force F_x exerted by the pendulum on the cart was applied on the subject's hand in the x direction. For safety reasons this force was saturated at the absolute value of 5 N and switched off completely when the pendulum angle exceeded 30° from the vertical. In order to maximize visual range, the visual representation of the task was shifted 13.0 cm towards the participant. The x-coordinate of the cart and the handle were always matched. The pendulum itself was represented as a blue line of 3.0 mm thickness connected to the center point of the cart (Fig. 1B). Due to the limitations of the screen size the whole pendulum could not be visualized and therefore it was truncated at the top of the screen. In addition, a blue circle (d = 1.0 cm) moving only in x direction was presented at the top of the screen. This circle represented the lateral position of the center of mass of the pendulum.

Trials were self-paced: subjects initiated each trial by moving the cart to the start position, indicated by a grey rectangle (3.0 cm by 1.5 cm). Participants were notified that they were within the home position by a yellow circle (d = 1.0 cm) appearing at the center of the cart. The trial initiation cue was a short beep followed by the pendulum starting to fall after 600 ms with initial angular velocity $\dot{\theta} = 0.01$ rad/s. The direction of the fall was randomized with equal probabilities for left and right. Subjects were required to maintain the pendulum in an upright position and with as little oscillation as possible. A trial was considered to have terminated when the angle between the pendulum and the y-axis reached 90° or when the pendulum was successfully balanced for 5.0 s. Subjects were then free to return to the start position and initiate the next trial while the feedback about the previous trial was shown.

In order to provide consistent feedback for participants a score variable (S) was introduced:

$$S = 100\ln\left(\frac{9000}{\sum_{t=0.001}^{5}\theta(t)^2}\right) \quad (3)$$

where t is the time of the sample. If the pendulum was not maintained upright for the duration of the trial, θ = 90° was used for all the remaining samples until the end of the trial.

Participants were introduced to a range of different pendulum dynamics. Specifically, participants were required to control a pendulum of mass m = 1 kg and lengths L = [0.25 m, 0.5 m, 0.75 m, 1 m, 1.5 m, 2 m, 4 m, 6 m, 8 m]. Each experimental block consisted of 20 trials of one given pendulum length. The nine different blocks were presented twice to participants in a pseudo-random order, so that every pendulum length was presented before any condition was repeated. Between blocks a short break was provided (3 s) where an illustration of a teapot was shown to notify participants that conditions would change. This resulted in 40 repetitions of each pendulum condition and a total of 360 trials per participant.

D. Data analysis

The data were analysed using Matlab R2016b. Force and kinematic time series were low-pass filtered with a fifth-order, zero-phase-lag Butterworth filter with 40 Hz cutoff. Acceleration data were obtained online by differentiating velocity data and then filtering it with an eight-order Butterworth filter (40 Hz cutoff).

III. RESULTS

A. Experimental data

Participants balanced a simulated inverted pendulum of unknown length while being provided the visual feedback of the end point, which coincided with the centre of mass of the pendulum. Participants' ability to maintain the pendulum upright increased with the length of the pendulum until the critical length, where consistent balance was achieved. Beyond this point increasing the length of the pendulum did not improve the stability (Fig. 2).

A small, but consistent, effect across participants was a decrease in the score for the longest length of the pendulum (Fig. 2A). However, a decrease was neither found for the balance time (Fig. 2B) nor angular velocity, which is a measure for the total system instability. Instead, a small increase in corrective movements was recorded for longer lengths (Fig. 2C). Such behaviour may arise due to the higher innate stability of longer pendulums. A long pendulum can be maintained upright even with only a small

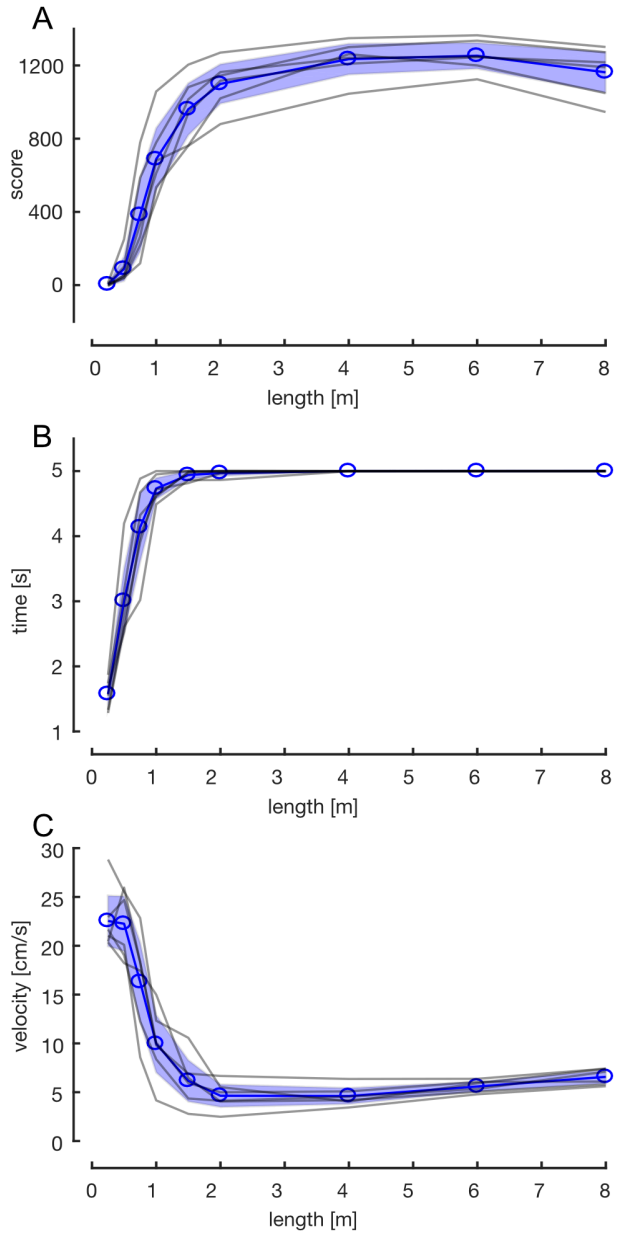


Figure 2. Effect of pendulum length on the controllability of the inverted pendulum. **A**, The score. Individual subject data is represented by grey lines, average response for all subjects is represented by the blue line. Shaded areas represent 95% confidence intervals of the mean. **B**, Average time the pendulum was maintained upright (maximum trial length 5s). **C**, Average velocity of the handle (cart). Cart velocity primarily reflects the subjects' control actions.

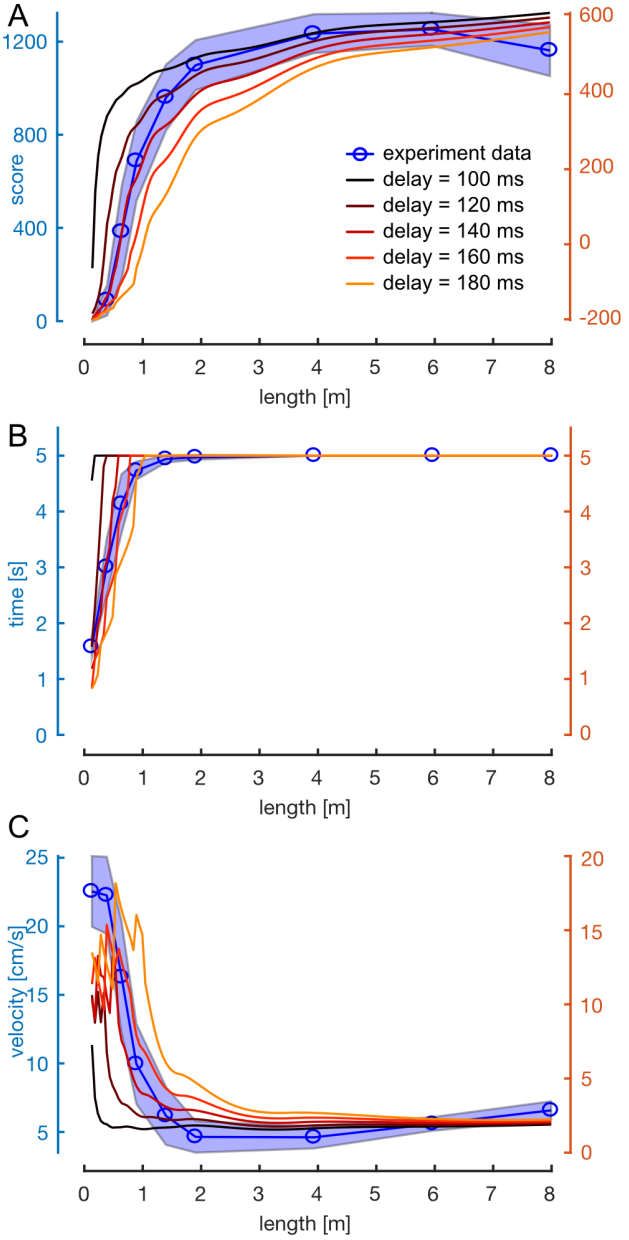


Figure 3. Comparison of experimental data with a delayed PD controller. **A-C**, Score, time balanced, and average cart velocity respectively. Individual lines represent the best fit PD controller (k_p=23, k_d=1.3) with different time delays. Secondary axes represent magnitudes of the controller output (non-normalized). Blue line and shaded region are experimental results as in Fig 2.

angular displacement, resulting in cart movement and increasing the average corrective velocity (Fig. 2C).

B. Computational model

The experimental data allowed us to compare control strategies of participants to a PD controller. We simulated a PD controlled virtual pendulum with a feedback delay in Matlab over a range of different controller parameters and compared the results with the experimental data. The performance of the controller was evaluated by comparing normalised output of the controller to the subject data in terms of score, time, angular error, and handle velocity. The fit between simulation and experimental results was then obtained by ranking each parameter set by least-squares error (LSE) for all four state variables. We then selected the best-fit PD controller by minimising the combined ranking for each parameter set resulting in parameters of $k_p = 21$, $k_d = 2.3$ and delay $\tau = 0.14$ s.

The effect of delay can be seen (Fig. 3) where the prediction of the PD controllers with the same PD parameters but different delays are shown. The model prediction was found to closely match subject data for short

lengths of the pendulum. At these lengths, due to innate instability of the system a constant control action is required making the controller output comparable to subject data. Subtle differences were observed at longer lengths, likely occurring from increased observation noise due to the higher visual motion for the human subjects. In contrast the PD controller has perfect information about the system, so longer lengths with increased stability will exhibit a monotonic increase in score and smaller control actions.

IV. Discussion

A simplified pendulum is a system consisting of a point mass, connected to a frictionless pin via a rigid weightless link. Such system is stable at the minimum energy configuration with a point mass hanging underneath the pin-joint or resting at the support. In its inverted configuration, the system can achieve marginal stability when no noise is present, however for it to be balanced consistently an online control action is required. The inverted pendulum model is therefore interesting from human motor control perspective, as it allows us to investigate the interaction of these two strategies. For example, a human could stay upright by maintaining a wider stance and co-contracting the muscles, therefore increasing the state-space of marginal stability, or engage the active control of the full-body oscillations.

Another reason why an inverted pendulum system is interesting is that its stability can be varied by changing its parameters such as length or mass. As expected our results show that human subjects have more difficulty to maintain the pendulum upright when its length is decreased (Fig. 2A). Such effects are likely caused by the delays in the human sensorimotor system; as a control action is applied to a delayed state of the pendulum, a less stable system may be too far away from the original state at the time when control input reaches the system, to be successfully balanced.

In this paper we present a simulated inverted pendulum system that could be employed to evaluate the human behavior when controlling external dynamics. We show that from a control perspective a simulated pendulum behaves similarly to a real pendulum as it is increasingly easier to maintain with an increasing length. Moreover, in our study the control input by the subjects was minimized at medium lengths (Fig. 2C, L = 2 m, 4 m). We would expect a similar behavior while controlling a real pendulum due to competing mechanisms: a decreasing control input with increasing length due to improving stability, and an increasing control input with increasing length due to the fact that same angular deviation moves center of mass of the pendulum further away for a longer system.

The results of our study can be used to examine the sensory feedback mechanisms used by the subjects. Different input modalities e.g. vision or proprioception, have different delays [15], [16] and different noise characteristics [17]. Therefore, comparison of our results with a time delayed PD controller can be used to estimate the way humans integrate sensory information to control external objects. A controller with time delay τ = 140 ms best explained the experimental data (Fig. 3). Similar delays are known to be present in a human visuomotor system [15], thus suggesting the importance of visual feedback in the control of an inverted pendulum. This may be explained by the fact that an inverted pendulum is largely controlled within a state-space of small angle deviations. At small angles the forces applied on the hand due to the pendulum drifting sideways are negligible and vision is a more reliable estimator of the state. Here we introduce a simulated pendulum for the study of control processes underlying stabilization. Future studies will investigate the relative role and adaptation of visual and haptic feedback to this control.

References

[1] B. Mehta and S. Schaal, "Forward models in visuomotor control.," *J Neurophysiol*, vol. 88, no. 2, pp. 942–953, 2002.

[2] I. D. Loram, P. J. Gawthrop, and M. Lakie, "The frequency of human, manual adjustments in balancing an inverted pendulum is constrained by intrinsic physiological factors.," *J. Physiol. (Lond.)*, vol. 577, no. 1, pp. 417–432, Nov. 2006.

[3] J. L. Cabrera and J. G. Milton, "Human stick balancing: tuning Lèvy flights to improve balance control.," *Chaos*, vol. 14, no. 3, pp. 691–698, Sep. 2004.

[4] E. Burdet, D. W. Franklin, and T. E. Milner, *Human Robotics - Neuromechanics and Motor Control*. MIT Press, 2013.

[5] P. Morasso, "'Brute force' vs. 'gentle taps' in the control of unstable loads.," *J. Physiol. (Lond.)*, vol. 589, no. 3, pp. 459–460, Feb. 2011.

[6] D. W. Franklin, "Impedance control: Learning stability in human sensorimotor control.," *Conf Proc IEEE Eng Med Biol Soc*, vol. 2015, pp. 1421–1424, Aug. 2015.

[7] E. Burdet, R. Osu, D. Franklin, T. Milner, and M. Kawato, "The central nervous system stabilizes unstable dynamics by learning optimal impedance," *Nature*, vol. 414, no. 6862, pp. 446–449, 2001.

[8] D. W. Franklin, G. Liaw, T. E. Milner, R. Osu, E. Burdet, and M. Kawato, "Endpoint stiffness of the arm is directionally tuned to instability in the environment.," *J. Neurosci.*, vol. 27, no. 29, pp. 7705–7716, Jul. 2007.

[9] D. W. Franklin, U. So, M. Kawato, and T. E. Milner, "Impedance control balances stability with metabolically costly muscle activation.," *J Neurophysiol*, vol. 92, no. 5, pp. 3097–3105, Nov. 2004.

[10] P. G. Morasso and V. Sanguineti, "Ankle muscle stiffness alone cannot stabilize balance during quiet standing.," *J Neurophysiol*, vol. 88, no. 4, pp. 2157–2162, Oct. 2002.

[11] I. D. Loram, H. Gollee, M. Lakie, and P. J. Gawthrop, "Human control of an inverted pendulum: Is continuous control necessary? Is intermittent control effective? Is intermittent control physiological?," *J. Physiol. (Lond.)*, vol. 589, no. 2, pp. 307–324, Jan. 2011.

[12] I. D. Loram and M. Lakie, "Direct measurement of human ankle stiffness during quiet standing: the intrinsic mechanical stiffness is insufficient for stability.," *J. Physiol. (Lond.)*, vol. 545, no. 3, pp. 1041–1053, Dec. 2002.

[13] R. C. Oldfield, "The assessment and analysis of handedness: the Edinburgh inventory.," *Neuropsychologia*, vol. 9, no. 1, pp. 97–113, Mar. 1971.

[14] I. S. Howard, J. N. Ingram, and D. M. Wolpert, "A modular planar robotic manipulandum with end-point torque control.," *J. Neurosci. Methods*, vol. 181, no. 2, pp. 199–211, Jul. 2009.

[15] D. W. Franklin, A. Reichenbach, S. Franklin, and J. Diedrichsen, "Temporal Evolution of Spatial Computations for Visuomotor Control.," *Journal of Neuroscience*, vol. 36, no. 8, pp. 2329–2341, Feb. 2016.

[16] J. A. Pruszynski, I. Kurtzer, T. P. Lillicrap, and S. H. Scott, "Temporal evolution of "automatic gain-scaling".," *J Neurophysiol*, vol. 102, no. 2, pp. 992–1003, Aug. 2009.

[17] R. J. van Beers, D. M. Wolpert, and P. Haggard, "When feeling is more important than seeing in sensorimotor adaptation.," *Curr. Biol.*, vol. 12, no. 10, pp. 834–837, May 2002.

3.8 Study V

Controller Gains of an Inverted Pendulum are Influenced by the Visual Feedback Position.

This study, authored by Justinas Česonis, Raz Leib, Sae Franklin and David W. Franklin was published in the proceedings of *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. In this study we tested how human participants balanced a range of pendulums with visual feedback of the pendulum that was congruent or incongruent with the actual dynamic length. As a result, we showed that participants heavily rely on the provided visual feedback of the pendulum tip, as the control behaviour deteriorated with increasing separation between the dynamics and the visual feedback.

Contributions

Justinas Česonis was the lead author and primary contributor in this research. In addition, Justinas Česonis implemented the experimental design, analysed the data, built the computational models and wrote the original draft. Sae Franklin collected the experimental data. Justinas Česonis, Raz Leib, Sae Franklin and David W. Franklin designed the study, selected the analysis methods, interpreted the results and finalised the paper together.

Abstract

In this study we experimentally test and model the control behavior of human participants when controlling inverted pendulums of different dynamic lengths, and with visual feedback of varying congruence to these dynamic lengths. Participants were asked to stabilize the inverted pendulum of L = 1 m and L = 4 m, with visual feedback shown at various distances along the pendulum. We fit a family of linear models to the control input (cart velocity) applied by participants. We further tested the models by predicting this control input for a pendulum with dynamic length L = 2 m and comparing the prediction to the experimental data. We show that the sum of proportional error correction and a term inversely proportional to visual feedback gain can well describe the control in human participants.

Controller Gains of an Inverted Pendulum are Influenced by the Visual Feedback Position

Justinas Česonis, Raz Leib, Sae Franklin, and David W. Franklin, *Member, IEEE*

***Abstract*—In this study we experimentally test and model the control behavior of human participants when controlling inverted pendulums of different dynamic lengths, and with visual feedback of varying congruence to these dynamic lengths. Participants were asked to stabilize the inverted pendulum of $L = 1$ m and $L = 4$ m, with visual feedback shown at various distances along the pendulum. We fit a family of linear models to the control input (cart velocity) applied by participants. We further tested the models by predicting this control input for a pendulum with dynamic length $L = 2$ m and comparing the prediction to the experimental data. We show that the sum of proportional error correction and a term inversely proportional to visual feedback gain can well describe the control in human participants.**

I. Introduction

Humans are regularly exposed to tasks where control of unstable dynamics is required, such as walking, cycling, slicing an apple or balancing a broom on the fingertips [1]. The performance in these tasks improves with learning (i.e. cycling) or deteriorates with suppression of feedback [2, 3]. Many studies have previously looked at learning (or predictive control) in the unstable environments [4, 5], however feedback control in those conditions was investigated much less [6, 7]. While control engineering approaches, such as PD control or LQG, have successfully been applied to model the feedback control of human movement in stable environments with congruent feedback [8, 9], the control with incongruent feedback has not yet been modelled.

The overall importance of visual feedback in human motor learning and control has been broadly investigated. We have previously experimentally shown that the control of a virtual pendulum with variable visual feedback was the most stable when this feedback was congruent with the dynamics of the pendulum [10]. Various hypotheses, such as the limited resources hypothesis, could be used to explain this paradigm qualitatively, but they do not quantify the development of the control stability with a change in visual feedback. Here we suggest a model that may describe the development of control input applied by humans when controlling an inverted pendulum with visual feedback incongruent with the pendulum dynamics. We further use the model to normatively simulate the control input for a novel condition and compare the results with experimentally collected data.

II. Methods

A. Participants

Six right-handed [11], neurologically healthy human participants (1 female, mean age 24.7 years), naïve to the purpose of this study, participated in the experiment. The participants were drawn from a pool of our previous inverted pendulum studies [10, 12], and therefore were familiar with the setup. All participants provided a written informed consent before participating in this study. The study was approved by the Ethics Committee of the Medical Faculty of the Technical University of Munich.

B. Experimental apparatus

Participants performed a balancing task of an inverted pendulum in a robotic manipulandum. Participants were seated in an adjustable chair in front of a robotic rig, with their shoulder movement restrained by a seatbelt. The subject's right arm rested on an airsled and their right hand grasped the handle of the vBOT robotic interface [13]. All hand movements were performed in an x-y plane parallel to the ground. Position and force data was sampled at 1 kHz. Visual feedback was projected via a computer monitor and a mirror system to the plane of the movement in such a way that the direct visual feedback of the hand was prevented.

C. Experimental paradigm

The inverted pendulum was simulated in the x-y plane, with the gravity acting in the negative y direction (towards the participant) and corrective movements executed by participants along the x axis (parallel to participant's chest). Pendulum kinematics and the interface were simulated and presented as described in [12] (Fig. 1A).

Trials were self-paced: participants initiated each trial by moving a cart to the start position, indicated by a grey rectangle (3.0 cm by 1.5 cm) and positioned in the middle of the control channel. The start of the trial was then cued via a short beep, followed by a perturbation of $\dot{\theta} = 0.01$ rad/s on the pendulum 600 ms later. The direction of this perturbation was pseudo-randomised, with equal number of trials for left and right. During each trial participants were instructed to maintain the pendulum upright and with as little of angular movement as possible for 5 seconds. A trial was considered over when the pendulum was successfully maintained for these 5 seconds, or when the angular deviation between the pendulum and the y axis reached 90°. Participants were then

J. Česonis, D.W. Franklin and R. Leib are with Neuromuscular Diagnostics, Department of Sport and Health Sciences, Technical University of Munich, Munich, 80992 Germany (phone: +49 89289 24536; e-mail: david.franklin@tum.de, justinas.cesonis@tum.de, raz.leib@tum.de).

S. Franklin is with the Institute of Cognitive Systems, Department of Electrical and Computer Engineering, Technical University of Munich, Munich, 80333 Germany (e-mail: sae.franklin@tum.de).

978-1-5386-1311-5/19/$31.00 ©2019 IEEE

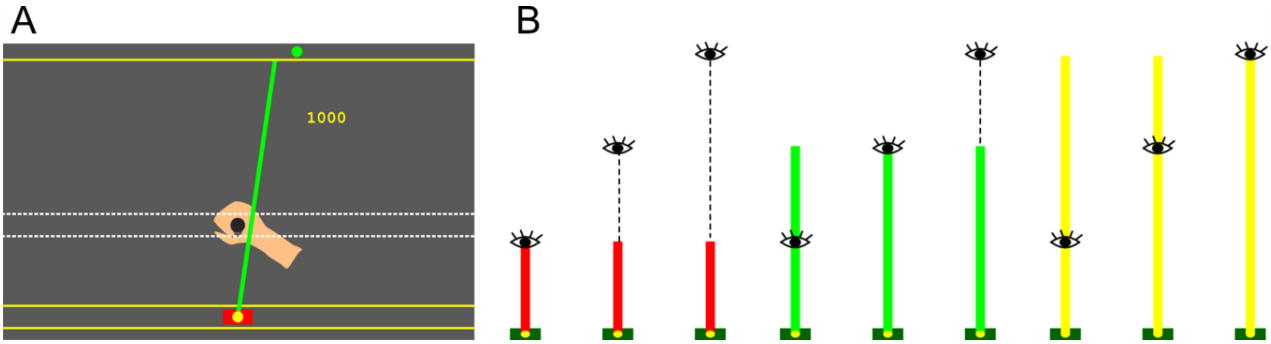


Figure 1. **A.** A sample snapshot of the experimental design. Participants controlled a cart (1.5 cm by 3 cm rectangular red block) directly, by moving the robotic handle along a mechanical channel (white lines and hand, not visible to participants; position dependent force field; stiffness 4000 N/m, damping 2 Ns/m, maximum force of 25 N). This channel constrained participants to move in x-axis only, at a distance approximately 30 cm in front of participant's chest, and was framed in the visual workspace by two yellow lines of 1.0 mm thickness. From participant's perspective, the physical hand location did not match with its visual representation (cart), but was shifted 13.0 cm forward in order to maximize the amount of visual feedback and the range of motion. However, the x-coordinate of the cart always matched the x-coordinate of the hand. **B.** Experimental paradigm schematic. Participants were introduced to pendulums of two different dynamic lengths (1 m, red and 4 m, yellow), and had previously participated in a similar experiment of a different dynamic length (2 m, green). Participants were provided with visual feedback at different locations. These locations were the same for all three dynamic conditions. As the horizontal displacement of the visual feedback point at the same pendulum angle is proportional to the visual feedback distance, this distance can be treated as the visual feedback gain.

provided with their score, indicating their task success [12], and were free to initiate the next trial.

Participants were required to control two different pendulums of dynamic length L = 1 m and L = 4 m. Each of the lengths was presented in a blocked fashion, with three participants starting with L = 1 m and then L = 4 m, and three participants with this order reversed. For each of the dynamic lengths participants were provided visual feedback equivalent to nine different visual distances from the cart: L_v=[0.25 m, 0.5 m, 0.75 m, 1 m, 1.5 m, 2 m, 4 m, 6 m, 8 m]. Each length was also presented in a blocked fashion, with 20 trials in each block and block order randomized (Fig. 1B). Each of these blocks was repeated twice, so that every pendulum condition was repeated for 40 trials in total, resulting in 720 trials per participant. Short breaks (5 s) were provided after every block, indicating participants that the condition would change.

D. Data analysis

Data collected in this experiment was analyzed and compared to the data of [10]. We used MATLAB 2017b for the data analysis. Kinematic time series were low-pass filtered with a zero-phase-lag, fifth-order Butterworth filter with 40 Hz cutoff frequency. Linear acceleration was obtained by differentiating the velocity data online and filtering it with eight-order low-pass Butterworth filter (40 Hz cutoff)

E. Modelling

Previously we have suggested a theoretical proportional feedback control model that could explain the development of the pendulum control input (cart velocity) [10]. In this study we test our theoretical model, as well as compare this model with alternative models. In order to quantify our model performance, we fit the model coefficients on the dynamic pendulum lengths L = 1 m and L = 4 m, and test the model by calculating the residual sum of squares (RSS) between the normative prediction of dynamic length L = 2 m and the respective data, collected in our previous study [10].

Our previously proposed model follows the mathematical expression:

$$M_0: v_x = A \cdot \frac{1}{L_v} + B \cdot e_x, \quad (1)$$

where v_x is the average cart velocity applied by participant, e_x is the visual feedback error, and A and B are model constants. However, in this study we examine a family of models of the general form

$$M: v_x = A \cdot \frac{1}{(L_v - C)} + B \cdot e_x + D, \quad (2)$$

where C and D are also model constants. Model constants may be dependent on the dynamic length, in which case they will be denoted with the subscript L (eg. A_L). We will further refer to separate models by these model constants *[A, B, C, D]*.

III. Results

A. Experimental data

Performance of six participants was compared across three different pendulum lengths in a balancing task. In order to test the effect of the dynamic length (L) and the visual feedback length (L_v) on the stability of the control, we performed a two-way repeated-measures ANOVA on the balance score as the dependent variable, with dynamic length (3 levels) and visual feedback length (9 levels) as within-subject independent factors. The analysis showed a significant main effect in both factors and their interactions (L: $F_{2,10}$=51.94, p<0.001; L_v: $F_{8,40}$=32.11, p<0.001; L*L_v: $F_{16,80}$=12.28, p<0.001). Post-hoc analysis (Holm-Bonferroni) revealed significant pairwise differences across all three dynamic lengths, with L = 4 m being most stable, and L = 1 m being the least stable, indicating an increase in stability for each increasing dynamic pendulum length (Fig. 2, top).

In order to better understand the interaction effects we also performed three one-way repeated-measures ANOVAs with visual feedback lengths (L_v) as factors (9 levels), and where dynamic pendulum length (L) was constant. For each of the dynamic lengths we found significant main effects in visual feedback length (L=1 m: $F_{8,40} = 17.32$, $p<0.001$; L=2 m: $F_{8,40} = 13.74$, $p<0.001$; L=4 m: $F_{8,40} = 53.77$, $p<0.001$). Moreover, for each pendulum, participants exhibited the most stable control when presented with a visual feedback length matching its dynamic length (Fig. 2, middle), with decay in controllability away from this point. Such a result is consistent with our previous findings [10], as well as with optimal feedback control models of an inverted pendulum where the observer and the plant have conflicting dynamic models.

B. Modelling

We fit a family of linear models (2) to our experimental data and evaluated their performance by comparing the model predictions for dynamic length $L = 2$ m to our previously collected data [10]. Here we present two best-fit models and compare them to the baseline models.

The two baseline models were chosen of the form *[A, 0, C, D]* and *[0, B, 0, D]*. The former model represents the control strategy where only the visual feedback location (L_v) influences the controller input. The best fit model of this form showed $RSS_{fit} = 1899.2$ and $RSS_{test} = 862.6$. The latter model represents the control strategy where only the visual endpoint error (Fig. 2, bottom) is corrected, with no estimate of the pendulum dynamics based on visual feedback location. The best fit model of this form showed $RSS_{fit} = 1491.4$ and $RSS_{test} = 873.0$. These both models show only a marginal improvement over the constant model *[0,0,0,D]*, with $RSS_{fit} = 2040.5$, and $RSS_{test} = 1034.9$, suggesting that neither of the two mechanisms are enough to represent the control system in our human participants.

Our two best-fit models were of the form *[A,B,C,0]* and *[A_L,B,C,0]*. The former model represents the control strategy where both inverse term and visual error are combined in a linear manner (Fig. 3, left). The best fit model of this form showed $RSS_{fit} = 800.7$ and $RSS_{test} = 494.6$, a significant improvement over any of the baseline models. The latter model assumes additional modulation of the inverse term with dynamic pendulum length. The best fit model of the latter form showed $RSS_{fit} = 726.1$, $RSS_{test} = 489.7$, a significant improvement (Fig. 3, right). Although more complex, the model *[A_L,B,C,0]* better describes our experimental data, and therefore is selected as a best-fit model.

IV. Discussion

In this study we tested and modelled a control behavior of human participants when controlling inverted pendulums of different mechanical properties under different visual feedback conditions. Our participants showed increasingly stable control behavior the closer visual feedback location was to the dynamic center of the pendulum, a result matching our previous study [10]. In addition to replicating this result for two new dynamic lengths (L = 1 m and L = 4 m), we also tested our previously proposed model of the control behavior.

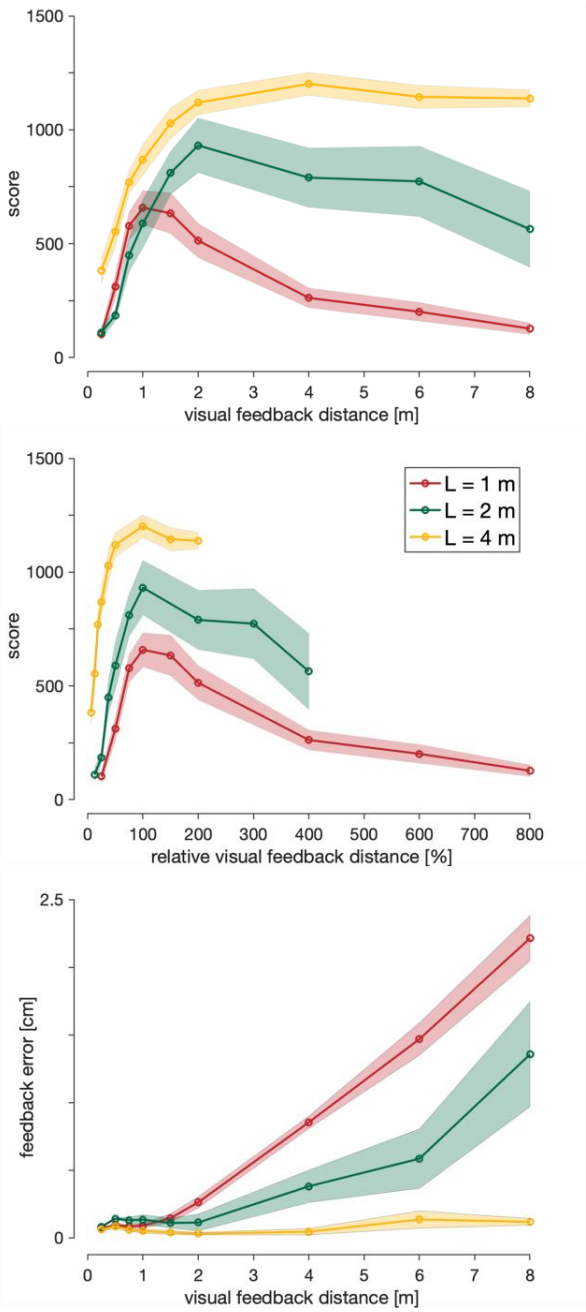


Figure 2. **Top.** Mean score across participants for different dynamic conditions and visual feedback distances. **Middle.** Score across all participants, as a function of visual feedback distance normalised by the dynamic pendulum length. Participants show the highest stability when controlling a pendulum with visual feedback congruent to the dynamic length. **Bottom**. Visual feedback point errors with respect to cart position. Participants successfully mitigate errors for visual feedback distances shorter than the dynamic lengths, but errors increase proportionally at longer than dynamic visual feedback lengths.

Our normative prediction of this control behavior generated results comparable to those recorded in human participants, with the model predictions within 1SEM from the human data (Fig. 3).

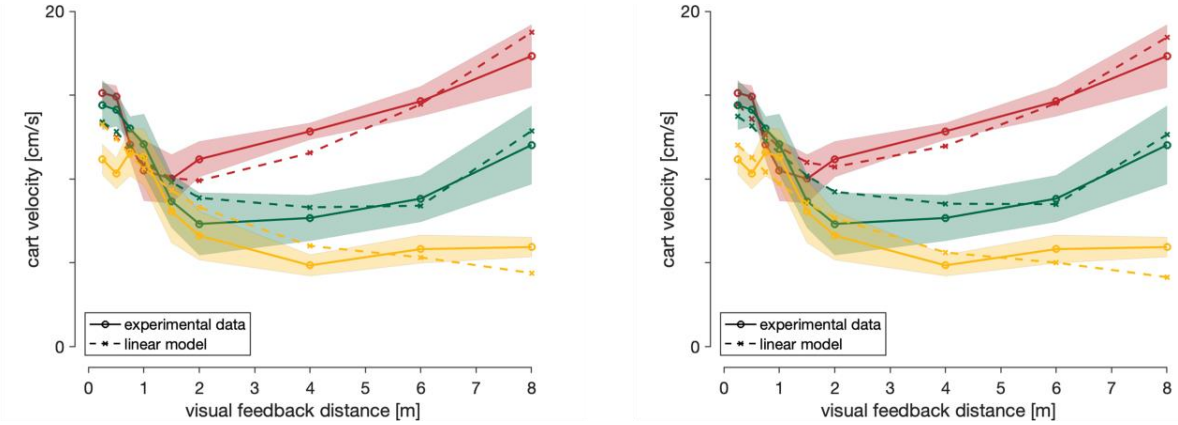


Figure 3. Two best-fit cart velocity models (solid lines), overlaid with the experimental data (dashed lines). Shaded areas represent 1SEM across participants. Both models capably represent our experimental data, with model predictions within 1SEM from the experimental data. **Left**. Model *[A,B,C,0]* (2nd best model). **Right.** Model *[A_L,B,C,0]* (best-fit model).

Our proposed model qualitatively captures the human control behavior of an inverted pendulum by combining mechanisms of proportional error correction and a term inversely proportional to visual feedback length. While the idea of error correction is widely accepted in human motor control, the purpose of a hyperbolic term is yet unclear and could have multiple explanations. One explanation for the presence of this term could be that humans experience an illusionary effect of reduced inertia of the pendulum with increasing visual length. Assuming a prior of Newton's 2nd law, the inertia of the pendulum would be estimated by a ratio between applied force and the acceleration of the visual feedback point. However, as the acceleration of this point scales with L_v, the perceived inertia would scale inversely. As a result, conditions with short visual feedback would feel "heavy" and would invoke stronger corrective responses, while the opposite is true for the long visual feedback locations.

Our participants exhibited significantly better stability when controlling a pendulum of a dynamic length L = 4 m, compared to L = 2 m and L = 1 m, at a visual feedback length $L_v = L$, while stability *of* L = 2 m was not different from L = 1 m in the same conditions. Previously we [12] have shown that virtual pendulums of L = 4 m and L = *2* m, simulated in a similar environment as in our experiment, were significantly more stable than a pendulum of L = 1 m. We believe that this difference stems from learning effects – all of our participants started with dynamic length L = 2 m, followed by the other two conditions in a balanced order, while in the previous study all three lengths were presented in a random order. Therefore, our participants managed to significantly improve their stability by performing the control task on different conditions. As a result, this may mean that the control strategies observed in our study have not yet converged to the optimal control. Further studies may compare how the controller changes with experience within this inverted pendulum environment.

In this article, we proposed a normative model of how humans may try to stabilize an inverted pendulum. More importantly, we showed the importance of a controller modality in which the control gains are inversely proportional to the visual feedback gains, suggesting that similar processing may occur in humans when controlling an unstable system. This result may allow us to better understand what control strategies or cost functions are used by humans, leading to a better understanding of human brain as well as a possibility to develop more efficient control algorithms by mimicking it.

REFERENCES

[1] D. W. Franklin and D. M. Wolpert, "Computational mechanisms of sensorimotor control," *Neuron,* vol. 72, pp. 425-442, 2011.

[2] F. R. Sarlegna, N. Malfait, L. Bringoux, C. Bourdin, and J.-L. Vercher, "Force-field adaptation without proprioception: can vision be used to model limb dynamics?," *Neuropsychologia,* vol. 48, pp. 60-67, 2010.

[3] H. Z. Lefumat, R. C. Miall, J. D. Cole, L. Bringoux, C. Bourdin, J.-L. Vercher, *et al.*, "Generalization of force-field adaptation in proprioceptively-deafferented subjects," *Neuroscience Letters,* vol. 616, pp. 160-165, 2016.

[4] D. W. Franklin, G. Liaw, T. E. Milner, R. Osu, E. Burdet, and M. Kawato, "Endpoint stiffness of the arm is directionally tuned to instability in the environment," *J Neurosci,* vol. 27, pp. 7705-7716, 2007.

[5] E. Burdet, R. Osu, D. Franklin, T. Milner, and M. Kawato, "The central nervous system stabilizes unstable dynamics by learning optimal impedance," *Nature,* vol. 414, pp. 446-449, 2001.

[6] I. D. Loram and M. Lakie, "Human balancing of an inverted pendulum: position control by small, ballistic-like, throw and catch movements," *J Physiol,* vol. 540, pp. 1111-1124, 2002.

[7] I. D. Loram, H. Gollee, M. Lakie, and P. J. Gawthrop, "Human control of an inverted pendulum: Is continuous control necessary? Is intermittent control effective? Is intermittent control physiological?," *J Physiol,* vol. 589, pp. 307-324, 2011.

[8] D. Liu and E. Todorov, "Evidence for the flexible sensorimotor strategies predicted by optimal feedback control," *J Neurosci,* vol. 27, pp. 9354-9368, 2007.

[9] L. Rigoux and E. Guigon, "A Model of Reward- and Effort-Based Optimal Decision Making and Motor Control," *PLOS Computational Biology,* vol. 8, p. e1002716, 2012.

[10] S. Franklin, J. Česonis, and D. W. Franklin, "Influence of visual feedback on the sensorimotor control of an inverted pendulum," *Conf Proc IEEE Eng Med Biol Soc,* vol. 2018, pp. 5170-5173, 2018.

[11] R. C. Oldfield, "The assessment and analysis of handedness: the Edinburgh inventory," *Neuropsychologia,* vol. 9, pp. 97-113, 1971.

[12] J. Česonis, S. Franklin, and D. W. Franklin, "A Simulated Inverted Pendulum to Investigate Human Sensorimotor Control," *Conf Proc IEEE Eng Med Biol Soc,* vol. 2018, pp. 5166-5169, 2018.

[13] I. S. Howard, J. N. Ingram, and D. M. Wolpert, "A modular planar robotic manipulandum with end-point torque control," *J Neurosci Meth,* vol. 181, pp. 199-211, 2009.

4 Discussion

This section provides a summary and discussion of main findings across all studies included in this thesis. The reader should approach this section as an addition to, and a summary of, individual discussions presented with each of the five studies. Even though all the studies in this thesis focus on various aspects of visuomotor control, major findings from studies I - III are summarised separately from studies IV - V due to different approaches and methodology. In addition, this section also discusses methodological suggestions for future work, that stem from the findings throughout this thesis, as well as statistical methods that may not yet be popularised within the current scientific practices, but are extremely effective. Finally, here we also discuss limitations that our studies were subjected to, and suggest future directions for studies that can build upon this presented work.

4.1 Summary of the visuomotor perturbation studies

The primary goal when designing studies I, II and III was to answer the question of how visuomotor feedback responses are regulated in humans. As a foundation to this work, it has been known that these responses are much faster than voluntary [11, 16]. To achieve this, it was suggested that visuomotor gains could be regulated via a simple constant-gain scaling, so avoiding cortical processing [82]. Yet, these responses are also flexible enough to demonstrate modulation on timescales of a single trial [22], implying some non-trivial processing. In addition, qualitatively this modulation also appeared optimal, as the response intensities increased when the perturbation happened closer to the target (leaving less time to correct), up until the point where they decreased again, indicating a giving-up (where perturbation happens so close to the target that correcting in time is physically impossible due to delays). As most of the known computational implementations of the OFC, particularly in presence of control-dependent noise, are computationally heavy [31, 43, 63–66, 83], the question remains whether the visuomotor feedback responses are controlled by OFC, or another variable-gain scaling could explain the results.

In Study I we have analysed the dependence of the visuomotor feedback intensities on the time-to-target, by dissociating various kinematic variables from each other. Specifically, even though the vision is a critical input for visuomotor responses to even occur, visual cursor trajectory or velocity provided no significant modulation of these responses. In contrast, modulating physical movement

kinematics provided significant regulation, even for perturbations that matched in onset location and velocity (Figure 3 in Study I). However, as such modifications to movement kinematics naturally altered movement durations and times-to-target, our experimental design allowed us to uncover the modulation where the responses produce matching intensities, as long as the time-to-target and the task goal, but not necessarily movement kinematics match. Finally, through additional OFC simulations we established a simple variable-gain tuning function that provides a simple mapping between time-to-target and visuomotor feedback intensities for a given feedback controller, in line with earlier hypothesis that visuomotor gain regulation could be precomputed before the movement [52].

Even though Study I demonstrated that we can reliably predict the visuomotor feedback response regulation via the time-to-target, the whole such control depends on the premise that time-to-target is an available variable to such a controller. On the one hand, time-to-contact studies in ecological physiology have long since demonstrated that such estimation is possible, as long as the moving object follows natural laws of motion [84–86]. However, it is unclear how such estimation happens on an algorithmic level. Some recent studies have suggested approaches that use utility functions, where an optimal movement duration maximises a utility – a trade-off between effort and reward [48, 49]. While in few of these studies accurate movement durations can be estimated for given movements, such approaches are difficult to generalise to novel or perturbed movements due to a non-trivial change in one or multiple parameters. In contrast, as we demonstrated in simulations of Study I, infinite-horizon or receding-horizon OFC can accurately predict the movement durations of unseen perturbed movements. As a result, in Study II we expand on this idea and propose a mixed-horizon control that first determines the expected movement duration and then generates the movement that produces variable feedback gains.

In Study II we provide a computational model that allows us to simulate complete perturbed movements and bypasses individual limitations of infinite-horizon OFC, and of finite-horizon OFC. To achieve this computationally, we combined an infinite-horizon and a finite-horizon OFCs in series, which we likened to motor selection and motor planning stages respectively. While this interpretation is a subject of ongoing scientific discussions [87, 88], its usefulness in computationally modelling previously unexplained behaviours is evident from case studies, presented Study II. For example, the mixed-horizon OFC allows us to simulate perturbed movements that were not performed experimentally due to experimental limitations [18]. In turn, such simulations can not only replicate the behaviour results at the measured data points, but also provide context around them by simulating unseen perturbations, in turn helping explain interesting phenomena, such as an unexpected saturation of visuomotor feedback intensities. Note that this would not be possible without the

infinite-horizon component, as the expected movement duration would not be known to simulate such movements in finite-horizon.

Indirectly, the mixed-horizon OFC further reinforces the importance of time-to-target in the optimal feedback control of various human movements. Specifically, from the functional perspective an infinite-horizon component simply provides a reliable source of time-to-target to the finite-horizon controller, which then does the heavy lifting in producing actual movements. As these simulations do provide a better fit to the data compared to just the finite-horizon simulations (see Figure 7B in Study II), this emphasises that the time-to-target is a key variable in controlling the feedback intensities. However, in order to test this regulation directly we can now use this model to make behavioural predictions and test them experimentally.

In Study III we demonstrated two major outcomes. First, we definitively verified the time-to-target based feedback intensity regulation by first generating a prediction through a finite-horizon OFC (which due to our experimental design, matched the predictions of the mixed-horizon OFC), and then observing the predicted regulation in the experimental data. Second, we also tested this control for two different tasks, first in blocked schedule, and then in a mixed schedule, to demonstrate that humans can reliably switch between multiple feedback controllers when cued by context – similarly to the contextual switching of the feedforward control. This latter finding was also facilitated by our earlier findings of the time-to-target control: as time-to-target regulates the feedback response intensities through a characteristic function, we can establish these characteristic functions for each movement condition from experimental data. As a result, comparing the two functions then allows us to conclude whether the two controllers can be similar or distinct.

Overall, the primary take-away from studies I - III is that the time-to-target seems to be a critical independent variable in the control of the visuomotor feedback gains (and intensities). In addition, we have also uncovered the characteristic relationship between the two variables, expressed as a time-to-target tuning function (see Study I for more details). This function is unique for the task goal and is independent from movement kinematics, such as peak velocity, general velocity profiles or movement distance, but will differ for movements with different goals (such as stop or hit movements in Study III). Moreover, this difference across different movement goals stems specifically from different feedback controllers, and thus the time-to-target tuning curves can be used to verify whether the two separate types of movements share the same feedback controller, or use different feedback controllers. Indeed, in such a fashion we compared the feedback controllers for hit and stop movements in Study III, not only showing that these controllers were different, but also that they can be rapidly switched on a trial-by-trial schedule. Finally, while it is known from ear-

lier work that humans seem to have the ability to estimate the time-to-target in their movements, infinite-horizon or receding-horizon OFC provide a well-functioning algorithmic implementations for this variable in computations, enabling a powerful, mixed-horizon OFC approach for predicting and modelling human-like behaviour.

4.2 Summary of the pendulum studies

While the author of this thesis was the lead author and primary contributor to Studies IV and V, he also contributed to additional studies, lead by other lab members, that also investigated the control of the inverted pendulum. Hence, this section summarizes the outcomes of Studies IV and V within the broader context of all inverted pendulum studies.

Study IV is primarily a methodological publication, where we tested the viability of the simulated inverted pendulum, and how its control compares to the control of a similar, fully physical pendulum. Here we presented a set of dynamic equations that govern the behaviour of a cart-pendulum system, as well as produce lateral forces that the pendulum dynamics exert on the cart. The results of the experiment, where participants controlled various such pendulums of different lengths, matched well with our hypotheses, informed by the understanding of the physical control of such pendulums – longer pendulums lead to longer balance, smaller oscillations and reduced corrective movements. In addition, we also compared human-controlled pendulum kinematics with the same pendulum, where the control was simulated using a PD controller with various delays. Here the best-fit delay of 150 ms, similar to the delay in the visuomotor feedback loop in humans [11], provided additional support that visual feedback is critical in stabilising such a system.

In [79] and Study V we modulated the feedback of the virtual pendulums so that the visual information was not anymore directly matching with pendulum dynamics. Here we tested pendulums of different dynamic (physical) lengths, and for each length the visual feedback was provided either at the centre of mass, or at various positions away from the centre of mass. In both these studies we recorded the most stable control when the visual information provided direct feedback on the centre of mass, as opposed to other feedback locations, indicating the importance of consistent information for efficient multisensory integration. Indirectly, the results also demonstrate that for control of such pendulum humans used the feedback of a single point – the provided “tip” of the pendulum (relative to the cart position) – and not any angular information. This is because any angular information would remain invariant to the modulation of the visual feedback length, as such modulation is simply equivalent to looking at a different length on the same rod, which does not affect its angular properties. However, the changes in produced control as this feedback length changed suggest that the

angular information is not the primary visual feedback source.

The results of these foundational studies have lead to further research that is of interest to our research group. First, fully virtual environment allowed us to test the effect of feedback delays to human control behaviour, demonstrating that additional delays gradually reduce the controllability of the system [10]. Second, in addition to PD control, which is commonly used to explain the motor performance in balance, we tested alternative, LQG based models, showing a good fit between data and the model [45]. Finally, as the singular feedback point was used as a source of primary feedback in balancing tasks, we also tested the contribution and behaviour of visuomotor feedback responses in the balancing task, showing that these responses are present not only in reaching, but also in balancing tasks, and likely contribute to the feedback control in stabilisation tasks [89].

4.3 Outlook on the methods

Over the published experimental studies, as well as multiple pilot designs that were tested in the process of this thesis we raised a few methodological considerations for future studies. These are summarised in this chapter.

4.3.1 Channel trials and maintained perturbations

Many recent studies on visuomotor feedback responses have included perturbations that are induced within channel trials, where the movements are laterally constrained in a force channel and perturbations only last for a short duration [18, 22, 25, 26]. Such channel trials are extremely useful in recording the feedback responses, as the force can be measured directly, as opposed to obtaining it through double differentiation of velocity, which reduces accuracy. On the other hand, due to task irrelevance of perturbations in the channels participants tend to reduce their responses over the duration of the experiment, which has given rise to an inclusion of maintained perturbations where the cursor or target remained perturbed throughout the trial after the initial perturbation onset, so requiring an active correction. Typically, any unperturbed movement with or without force channel is very similar, with matching movement duration and kinematics. However, in trials where reaching a target is critical, the movement durations are significantly shorter in channel trials than in free trials if a lateral perturbation is induced (Figure 6 in Study I). As a result, having a mixture of both channel trials and maintained perturbation trials lead to inconsistent times-to-target for otherwise matching trials.

Even if a mixture of maintained and channel perturbation was not a problem for participants to successfully complete the trials, researchers are forced to make assumptions on what the perceived movement duration is. In Study I we worked around this issue by averaging the durations of matching

free and channel trials, weighed by the density of each type of trial. However, to properly control the time-to-target in Study III we instead opted to not have any maintained perturbations and instead only used the channel trials. As a downside, we observed reduced feedback intensities, which in turn required a larger experimental sample to maintain the power of the study. Thus, while we do not advocate for a particular design regarding a combination of maintained perturbation and channel trials, the effects of each of them should be considered before designing an experimental study.

4.3.2 Statistical approach

Frequentist statistical analysis is widely spread and accepted in scientific literature. However, such analysis is significantly limited by its single decision boundary (null hypothesis or alternative hypothesis), and a binary outcome (H_0 accepted or H_0 rejected). In contrast, Bayesian Factor analysis considers each available datapoint as evidence towards either a null hypothesis, or an alternative hypothesis, effectively providing two decision boundaries with three outcomes (H_0 accepted, H_1 accepted, or not enough evidence to accept either hypothesis). Moreover, Bayesian evidence can be interpreted in a continuous scale, which is helpful to observe tendencies in the results, even if no significance is determined. As a result, in addition to traditional frequentist methods, we supplemented all our statistical analysis with equivalent Bayesian tests, easily implemented via JASP software [90].

4.3.3 Modelling of noise

As noise is one of the key problems that human motor control system has to deal with, previous literature presented models that account for such noise [31, 38, 63, 65, 66, 83]. Specifically, two types of noise – additive and multiplicative – are considered, as they both influence the control system in different ways. On the one hand, additive noise simply corrupts the output of the system (i.e. the state of the controlled object) by adding a zero-mean random variable, sampled from a Normal distribution. In experimental paradigms where only mean effect is of interest, the effects of such noise are diminished by recording repeated samples (usually ranging between 25 and 40 in our studies). As a result, when modelling mean behaviours, the additive noise term can simply be ignored. On the other hand, multiplicative noise systematically affects the control system, as it usually scales with the control signal (or a state variable, if such noise is affecting the sensory system). It was demonstrated by earlier work [31, 38, 39] that inclusion of multiplicative noise that scales with a control signal can capture additional variance of human movements, compared to similar models without multiplicative noise. However, as shown in Study II, where we compared models with and without multiplicative noise in simulating visuomotor feedback intensities, even models without multiplicative noise capture the temporal evolution of the intensities well enough to

separate different movement conditions or even nearby perturbations. While more complex models with control dependent noise indeed improve on the overall accuracy of the simulations compared to the data, this difference is much smaller than the effect of the simulated perturbations. Hence, for simplicity we did not include multiplicative noise in our simulations in Study III.

4.4 Limitations

All research, presented in this thesis is to some extent subject to limitations, arising primarily from study logistics, study design, or methodological constraints. While some of these limitations can be addressed by specifically testing them in future work, others need to be critically considered when evaluating the results of this thesis. Overall, all five studies, as well as other studies where the author contributed, only focus on neurologically healthy younger adults. As a result, without further testing the outcomes from this work should not be generalised to older populations or to neurological patients.

4.4.1 Visuomotor feedback response studies

Few earlier visuomotor feedback response studies have looked at response time windows not only between 180-230 ms, but also later, beyond 230 ms [5, 18, 19]. As the threshold at 230 ms signifies the onset of voluntary responses [16], these later responses contain a mixture of voluntary and involuntary control, which may manifest in different behaviour compared to purely involuntary control. Indeed, [19] have demonstrated differences between earlier and later responses, for example, in successfully avoiding obstacles, which require a more complex control behaviour. Even though we expect that our results would hold for later time windows, in our work we only experimentally evaluate human behaviour in the involuntary, early visuomotor window, and thus can not make strict conclusions about later control.

Another limitation of our work pertains to the model-based support for our experimental data. Specifically, due to an infinite set of theoretically available model parameters it is likely possible that fitting a control model to the data post-study would eventually provide a good fit. In order to avoid such overfitting, we deliberately aimed to maintain a fixed set of model parameters wherever possible and compare the models to the data qualitatively in terms of common features, such as relative intensity regulation across conditions or shapes of response profiles. While we maintain that such comparisons are extremely valuable, they ultimately are subject to interpretation, as they do not provide objective, statistically-backed hard evidence.

4.4.2 Pendulum studies

While vBot environment enables interesting manipulations for our pendulum balance studies, the control of long pendulums (up to 8 m) is limited by the workspace size and orientation. First, as we could not fit the entire lengths of pendulum into the screen, we cropped most of the length outside of the screen and provided a cursor, representing the tip of the pendulum. Second, the whole motion was simulated in the plane parallel to the ground, and therefore the gravity was experienced in the direction towards participant's chest, with haptic feedback only available in one axis. Such a design might have biased participants into more visuomotor control simply out of necessity, as the haptic feedback or any angular information was scarce, compared to the visual feedback of the cart and the tip of the pendulum.

Finally, our pendulum studies aimed to conceptually test the paradigms for future experimental studies, as well as to verify the viability of our hypotheses. As a result, these studies were performed by a small sample of experienced participants, rather than a larger samples from the general population. In turn, while such a sample should not have significantly reduced the effect sizes (due to generally lower variance in experienced participants), it may have introduced biases to the results.

4.5 Future work

Studies I - III in this thesis all provide some direct or indirect evidence into time-to-target based control of feedback responses. However, as a whole, there is yet not much experimental evidence for this phenomenon outside of these studies. Further work could test additional predictions of our models. An interesting control would be to test the regulation of responses in matching movements with maintained perturbations and channel trials, as earlier studies combining the two did not assume any differences between them. Importantly, the forward kinematics in the two types of movements would be matched, and thus perturbation onset times, kinematics and locations would also be consistent, meaning that the major difference between the two conditions would arise from extended time-to-target for maintained perturbations. Moreover, this difference can be easily predicted using a mixed-horizon OFC model.

In Study III we demonstrated that different task goals act as a strong contextual cue for different feedback controllers. However, stop and hit movements were kinematically different, not ruling out the possibility that the context was cued by different movement kinematics or dynamics. A possible extension of this work could compare in a similar design the control between hit and long-stop movements, where we expect to see matching kinematics along the entire movement, yet our simulations suggest entirely different feedback intensity profiles. If such model predictions were experimentally

confirmed, this would demonstrate that task context can simply be cued by instruction, rather than any change in observable kinematics.

While most of our work was experimental, our results were also reinforced by simulations of a similar task within OFC framework. However, all our studies only focused on the early visuomotor feedback response (before 230 ms), and none of our models contained any assumptions, specific to the early visuomotor response, but not to the late visuomotor response. Yet new evidence emerges, suggesting that the regulation of early response may differ from the regulation of later response [19]. Similarly, while lateral target jumps seem to simply evoke a feedback response by the current controller, other transformations such as target jump away or towards the hand seem to require a recomputation of the feedback control gains [22]. Further work could investigate the origins of these differences, which will lead to further increased understanding of the feedback processing in humans.

Studies IV and V have laid the foundation for further tests in the control of balance of unstable objects. In addition of analysing specific algorithmic strategies for such balance in humans, future work can test relative contributions of vision and haptic information by modifying uncertainty of the sensory inputs or by inducing feedback delays on all or some modalities. Finally, as humans seem to combine optimal control and robust control (i.e. through modulating limb stiffness) in various motor actions, we can test how each of these systems are modulated individually.

4.6 Summary, conclusions and outlook

This thesis suggests a new theoretical mechanism of visuomotor feedback control, grounded in the relation between feedback gains and the time-to-target. As this mechanism stems primarily from optimal feedback control theory and computationally contains no assumptions about the effector, similar effects could be tested for other feedback mechanisms where optimal control is suspected (i.e. eye saccades or stretch reflexes). Moreover, we have successfully used the time-to-target model, as well as mixed-horizon OFC, the idea of which is grounded in the time-to-target based control, to model a vast array of human behaviours. Thus, our results could also be considered as methodological contributions to the field of motor control, and will hopefully inspire novel hypotheses for new studies in the future.

Numerous studies, both with visual perturbations [16, 18, 19, 22, 23, 25, 26, 33], as well as mechanical [20, 42, 91, 92] often cue perturbations based on position, which may not necessarily lead to consistent time-to-target across the compared perturbations. Our work demonstrates that outcomes of such comparisons could need to be re-evaluated, as even the same controller may produce dif-

ferent responses to perturbations cued at the the same location if the times-to-target do not match. Similarly, while in recent practice multiple visuomotor feedback response studies combined maintained perturbations and channel trials in order to keep responses high, such designs may have induced unforeseen and unwanted inconsistencies between measured responses and similar responses outside the lab. Taken together, our results indicate that various different interventions should be controlled with respect to the time-to-target, rather than conventionally used variables of position or time from the start of the movement.

Finally, the balance of an inverted pendulum is a fundamental motor skill that keeps humans standing on two feet. While locomotion and gait have received broad interest and research focus in the past many years, pendulum balancing remains relatively under-studied. I hope, that our proofs of concept of experimental design in virtual environment, as well as first results that match earlier hypotheses for such control, will lay new ground for more similar studies in the future.

5 Bibliography

[1] P. Cisek and J. F. Kalaska, “Neural mechanisms for interacting with a world full of action choices,” *Annual Review of Neuroscience*, vol. 33, pp. 269–298, 2010.

[2] J. P. Gallivan, K. S. Barton, C. S. Chapman, D. M. Wolpert, and J. R. Flanagan, “Action plan co-optimization reveals the parallel encoding of competing reach movements,” *Nature Communications*, vol. 6, no. May, pp. 1–9, 2015.

[3] M. J. Wagner and M. A. Smith, “Shared internal models for feedforward and feedback control,” *Journal of Neuroscience*, vol. 28, no. 42, pp. 10663–10673, 2008.

[4] M. A. Ahmadi-Pajouh, F. Towhidkhah, and R. Shadmehr, “Preparing to reach: Selecting an adaptive long-latency feedback controller,” *Journal of Neuroscience*, vol. 32, no. 28, pp. 9537–9545, 2012.

[5] S. Franklin, D. M. Wolpert, and D. W. Franklin, “Rapid visuomotor feedback gains are tuned to the task dynamics,” *Journal of Neurophysiology*, p. jn.00748.2016, 2017.

[6] R. S. Maeda, T. Cluff, P. L. Gribble, and J. A. Pruszynski, “Feedforward and feedback control share an internal model of the arm’s dynamics,” *Journal of Neuroscience*, vol. 38, no. 49, pp. 10505–10514, 2018.

[7] S. Franklin and D. W. Franklin, “Feedback Gains modulate with Motor Memory Uncertainty,” *Neurons, Behavior, Data analysis, and Theory*, vol. 5, no. 2, pp. 1–28, 2021.

[8] D. M. Wolpert and M. Kawato, “Multiple paired forward and inverse models for motor control,” *Neural Networks*, vol. 11, no. 7-8, pp. 1317–1329, 1998.

[9] F. Crevecoeur, D. P. Munoz, and S. H. Scott, “Dynamic Multisensory Integration: Somatosensory Speed Trumps Visual Accuracy during Feedback Control,” *Journal of Neuroscience*, vol. 36, no. 33, pp. 8598–8611, 2016.

[10] S. Franklin, J. Česonis, R. Leib, and D. W. Franklin, “Feedback delay changes the control of an inverted pendulum,” in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2019, Berlin, Germany, July 23-27, 2019*, pp. 1517–1520, IEEE, 2019.

[11] C. Prablanc and O. Martin, “Automatic control during hand reaching at undetected two-dimensional target displacements,” *Journal of Neurophysiology*, vol. 67, pp. 455–469, feb 1992.

[12] J. A. Saunders and D. C. Knill, “Humans use continuous visual feedback from the hand to control fast reaching movements,” *Experimental Brain Research*, vol. 152, no. 3, pp. 341–352, 2003.

[13] M. Jeannerod, *The neural and behavioural organization of goal-directed movements*. Clarendon Press/Oxford University Press., 1988.

[14] B. L. Day and I. N. Lyon, “Voluntary modification of automatic arm movements evoked by motion of a visual target,” *Experimental Brain Research*, vol. 130, no. 2, pp. 159–168, 2000.

[15] F. R. Sarlegna and P. K. Mutha, “The influence of visual target information on the online control of movements,” *Vision Research*, vol. 110, no. PB, pp. 144–154, 2015.

[16] D. W. Franklin and D. M. Wolpert, “Specificity of Reflex Adaptation for Task-Relevant Variability,” *Journal of Neuroscience*, vol. 28, no. 52, pp. 14165–14175, 2008.

[17] J. A. Pruszynski, I. Kurtzer, and S. H. Scott, “Rapid Motor Responses Are Appropriately Tuned to the Metrics of a Visuospatial Task,” *Journal of Neurophysiology*, vol. 100, no. 1, pp. 224–238, 2008.

[18] D. W. Franklin, A. Reichenbach, S. Franklin, and J. Diedrichsen, “Temporal Evolution of Spatial Computations for Visuomotor Control,” *Journal of Neuroscience*, vol. 36, no. 8, pp. 2329–2341, 2016.

[19] K. P. Cross, T. Cluff, T. Takei, and S. H. Scott, “Visual Feedback Processing of the Limb Involves Two Distinct Phases,” *Journal of Neuroscience*, vol. 39, no. 34, pp. 6751–6765, 2019.

[20] A. de Comite, F. Crevecoeur, and P. Lefèvre, “Online modification of goal-directed control in human reaching movements,” *Journal of neurophysiology*, vol. 125, no. 5, pp. 1883–1898, 2021.

[21] L. Oostwoud Wijdenes, E. Brenner, and J. B. J. Smeets, “Fast and fine-tuned corrections when the target of a hand movement is displaced,” *Experimental Brain Research*, vol. 214, no. 3, pp. 453–462, 2011.

[22] M. Dimitriou, D. M. Wolpert, and D. W. Franklin, “The Temporal Evolution of Feedback Gains Rapidly Update to Task Demands,” *Journal of Neuroscience*, vol. 33, no. 26, pp. 10898–10909, 2013.

[23] D. C. Knill, A. Bondada, and M. Chhabra, “Flexible, Task-Dependent Use of Sensory Feedback to Control Hand Movements,” *Journal of Neuroscience*, vol. 31, no. 4, pp. 1219–1237, 2011.

[24] J. Izawa and R. Shadmehr, "On-Line Processing of Uncertain Information in Visuomotor Control," *Journal of Neuroscience*, vol. 28, no. 44, pp. 11360–11368, 2008.

[25] A. J. de Brouwer, T. Jarvis, J. P. Gallivan, and J. R. Flanagan, "Parallel Specification of Visuomotor Feedback Gains during Bimanual Reaching to Independent Goals," *Eneuro*, vol. 4, no. 2, pp. ENEURO.0026–17.2017, 2017.

[26] A. J. de Brouwer, J. P. Gallivan, and J. R. Flanagan, "Visuomotor feedback gains are modulated by gaze position," *Journal of Neurophysiology*, vol. 120, pp. 2522–2531, sep 2018.

[27] J. A. Saunders and D. C. Knill, "Humans use continuous visual feedback from the hand to control both the direction and distance of pointing movements," *Experimental Brain Research*, vol. 162, no. 4, pp. 458–473, 2005.

[28] F. Sarlegna, J. Blouin, J. P. Bresciani, C. Bourdin, J. L. Vercher, and G. M. Gauthier, "Target and hand position information in the online control of goal-directed arm movements," *Experimental Brain Research*, vol. 151, no. 4, pp. 524–535, 2003.

[29] F. Sarlegna, J. Blouin, J. L. Vercher, J. P. Bresciani, C. Bourdin, and G. M. Gauthier, "Online control of the direction of rapid movements," *Experimental Brain Research*, vol. 157, no. 4, pp. 468–471, 2004.

[30] A. Reichenbach, D. W. Franklin, P. Zatka-Haas, and J. Diedrichsen, "A dedicated binding mechanism for the visual control of movement," *Current Biology*, vol. 24, no. 7, pp. 780–785, 2014.

[31] D. Liu and E. Todorov, "Evidence for the Flexible Sensorimotor Strategies Predicted by Optimal Feedback Control," *Journal of Neuroscience*, vol. 27, no. 35, pp. 9354–9368, 2007.

[32] L. Oostwoud Wijdenes, R. J. Van Beers, and W. P. Medendorp, "Vestibular modulation of visuomotor feedback gains in reaching," *Journal of Neurophysiology*, vol. 122, no. 3, pp. 947–957, 2019.

[33] Y. Zhang, E. Brenner, J. Duysens, S. Verschueren, and J. B. Smeets, "Postural responses to target jumps and background motion in a fast pointing task," *Experimental Brain Research*, vol. 236, no. 6, pp. 1573–1581, 2018.

[34] E. Brenner and J. B. Smeets, "Fast corrections of movements with a computer mouse," *Spatial Vision*, vol. 16, no. 3-4, pp. 365–376, 2003.

[35] N. Saijo, I. Murakami, S. Nishida, and H. Gomi, "Large-Field Visual Motion Directly Induces an Involuntary Rapid Manual Following Response," *Journal of Neuroscience*, vol. 25, no. 20, pp. 4941–4951, 2005.

[36] T. Flash and N. Hogan, “The coordination of arm movements: an experimentally confirmed mathematical model.,” *Journal of Neuroscience*, vol. 5, no. 7, pp. 1688–1703, 1985.

[37] Y. Uno, M. Kawato, and R. Suzuki, “Formation and control of optimal trajectory in human multijoint arm movement,” *Biological Cybernetics*, vol. 61, no. 2, pp. 89–101, 1989.

[38] C. M. Harris and D. M. Wolpert, “Signal-dependent noise determines motor planning,” *Nature*, vol. 394, p. 780, 1998.

[39] E. Todorov and M. I. Jordan, “Optimal feedback control as a theory of motor coordination,” *Nature Neuroscience*, vol. 5, no. 11, pp. 1226–1235, 2002.

[40] F. Crevecoeur, I. Kurtzer, T. Bourke, and S. H. Scott, “Feedback responses rapidly scale with the urgency to correct for external perturbations,” *Journal of Neurophysiology*, vol. 110, no. 6, pp. 1323–1332, 2013.

[41] I. L. Kurtzer, J. A. Pruszynski, and S. H. Scott, “Long-Latency Reflexes of the Human Arm Reflect an Internal Model of Limb Dynamics,” *Current Biology*, vol. 18, no. 6, pp. 449–453, 2008.

[42] I. Kurtzer, J. A. Pruszynski, and S. H. Scott, “Long-latency responses during reaching account for the mechanical interaction between the shoulder and elbow joints.,” *Journal of neurophysiology*, vol. 102, no. 5, pp. 3004–15, 2009.

[43] J. Y. Nashed, F. Crevecoeur, and S. H. Scott, “Influence of the behavioral goal and environmental obstacles on rapid feedback responses,” *Journal of Neurophysiology*, vol. 108, no. 4, pp. 999–1009, 2012.

[44] J. A. Pruszynski and S. H. Scott, “Optimal feedback control and the long-latency stretch response,” *Experimental Brain Research*, vol. 218, no. 3, pp. 341–359, 2012.

[45] R. Leib, J. Česonis, S. Franklin, and D. W. Franklin, “Lqg framework explains performance of balancing inverted pendulum with incongruent visual feedback,” in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2019, Berlin, Germany, July 23-27, 2019*, pp. 1940–1943, IEEE, 2019.

[46] A. F. D. C. Hamilton and D. M. Wolpert, “Controlling the statistics of action: Obstacle avoidance,” *Journal of Neurophysiology*, vol. 87, no. 5, pp. 2434–2440, 2002.

[47] R. Shadmehr, J. J. Orban de Xivry, M. Xu-Wilson, and T.-Y. Shih, “Temporal Discounting of Reward and the Cost of Time in Motor Control,” *Journal of Neuroscience*, vol. 30, no. 31, pp. 10507–10516, 2010.

[48] R. Shadmehr, H. J. Huang, and A. A. Ahmed, "A Representation of Effort in Decision-Making and Motor Control," *Current Biology*, vol. 26, no. 14, pp. 1929–1934, 2016.

[49] L. Rigoux and E. Guigon, "A Model of Reward- and Effort-Based Optimal Decision Making and Motor Control," *PLoS Computational Biology*, vol. 8, no. 10, 2012.

[50] J. E. Colgate, "Robust control of dynamically interacting systems," *International Journal of Control*, vol. 48, no. 1, pp. 65–88, 1988.

[51] A. Kuo, "An optimal control model for analyzing human postural balance," *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 1, pp. 87–101, 1995.

[52] D. W. Franklin, "Rapid Feedback Responses Arise From Precomputed Gains," *Motor Control*, vol. 20, pp. 171–176, apr 2016.

[53] K. Hidenori and Y. Jiang, "A PID model of human balance keeping," *IEEE Control Systems*, vol. 26, no. 6, pp. 18–23, 2006.

[54] K. Masani, A. H. Vette, and M. R. Popovic, "Controlling balance during quiet standing: Proportional and derivative controller generates preceding motor command to body sway position observed in experiments," *Gait and Posture*, vol. 23, no. 2, pp. 164–172, 2006.

[55] A. M. Haith and J. W. Krakauer, "Theoretical models of motor control and motor learning," *Routledge handbook of motor control and motor learning. London: Routledge*, pp. 7–28, 2013.

[56] J. Česonis and D. W. Franklin, "Mixed-horizon optimal feedback control as a model of human movement," *Neurons, Behavior, Data analysis, and Theory*, pp. 1–36, 2021.

[57] R. Oldfield, "The assessment and analysis of handedness: The Edinburgh inventory," *Neuropsychologia*, vol. 9, no. 1, pp. 97–113, 1971.

[58] I. S. Howard, J. N. Ingram, and D. M. Wolpert, "A modular planar robotic manipulandum with end-point torque control," *Journal of Neuroscience Methods*, vol. 181, no. 2, pp. 199–211, 2009.

[59] R. A. Scheidt, D. J. Reinkensmeyer, M. A. Conditt, W. Z. Rymer, and F. A. Mussa-Ivaldi, "Persistence of motor adaptation during constrained, multi-joint, arm movements.," *Journal of neurophysiology*, vol. 84, no. 2, pp. 853–862, 2000.

[60] J. Česonis and D. W. Franklin, "Contextual cues are not unique for motor learning: Task-dependant switching of feedback controllers," *PLOS Computational Biology*, p. forthcoming, 2022.

[61] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems. 1st Edition*. Wiley-Interscience, Hoboken., 1972.

[62] M. H. A. Davis and R. B. Vinter, *Stochastic Modelling and Control*. Springer Dordrecht, 1985.

[63] E. Todorov, "Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system.," *Neural Computation*, vol. 17, no. 5, pp. 1084–1108, 2005.

[64] E. Todorov and W. Li, "A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems," *Proceedings of the 2005, American Control Conference, 2005.*, pp. 300–306, 2005.

[65] Y. Jiang, Z. P. Jiang, and N. Qian, "Optimal control mechanisms in human arm reaching movements," *Proceedings of the 30th Chinese Control Conference, CCC 2011*, pp. 1377–1382, 2011.

[66] N. Qian, Y. Jiang, Z.-P. Jiang, P. Mazzoni, and D. Forster, "Movement duration, Fitts's law, and an infinite-horizon optimal feedback control model for biological motor systems," *Neural Computation*, vol. 25, pp. 697–724, mar 2013.

[67] J. Česonis and D. W. Franklin, "Time-to-target simplifies optimal control of visuomotor feedback responses," *eNeuro*, vol. 7, no. 2, pp. 1–17, 2020.

[68] E. Guigon, O. Chafik, N. Jarrassé, and A. Roby-Brami, "Experimental and theoretical study of velocity fluctuations during slow movements in humans," *Journal of Neurophysiology*, vol. 121, no. 2, pp. 715–727, 2019.

[69] J.-Y. Lee and N. Schweighofer, "Dual Adaptation Supports a Parallel Architecture of Motor Memory," *Journal of Neuroscience*, vol. 29, no. 33, pp. 10396–10404, 2009.

[70] I. S. Howard, J. N. Ingram, D. W. Franklin, and D. M. Wolpert, "Gone in 0.6 seconds: The encoding of motor memories depends on recent sensorimotor states," *Journal of Neuroscience*, vol. 32, no. 37, pp. 12756–12768, 2012.

[71] I. S. Howard, D. M. Wolpert, and D. W. Franklin, "The effect of contextual cues on the encoding of motor memories," *Journal of Neurophysiology*, vol. 109, no. 10, pp. 2632–2644, 2013.

[72] I. S. Howard, D. M. Wolpert, and D. W. Franklin, "The value of the follow-through derives from motor learning depending on future actions," *Current Biology*, vol. 25, no. 3, pp. 397–401, 2015.

[73] I. S. Howard, S. Franklin, and D. W. Franklin, "Asymmetry in kinematic generalization between visual and passive lead-in movements are consistent with a forward model in the sensorimotor system," *PLoS ONE*, vol. 15, no. 1, pp. 1–21, 2020.

[74] M. Forano and D. W. Franklin, "Timescales of motor memory formation in dual-adaptation," *PLoS Computational Biology*, vol. 16, no. 10, pp. 1–33, 2020.

[75] M. Forano, R. Schween, J. A. Taylor, M. Hegele, and D. W. Franklin, “Direct and indirect cues can enable dual adaptation, but through different learning processes,” *Journal of Neurophysiology*, vol. 126, no. 5, pp. 1490–1506, 2021. PMID: 34550024.

[76] B. Mehta and S. Schaal, “Forward Models in Visuomotor Control,” *Journal of Neurophysiology*, vol. 88, no. 2, pp. 942–953, 2002.

[77] I. D. Loram, P. J. Gawthrop, and M. Lakie, “The frequency of human, manual adjustments in balancing an inverted pendulum is constrained by intrinsic physiological factors,” *Journal of Physiology*, vol. 577, no. 1, pp. 417–432, 2006.

[78] J. L. Cabrera and J. G. Milton, “Human stick balancing: Tuning lèvy flights to improve balance control,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 14, no. 3, pp. 691–698, 2004.

[79] S. Franklin, J. Česonis, and D. W. Franklin, “Influence of visual feedback on the sensorimotor control of an inverted pendulum,” in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 5170–5173, 2018.

[80] J. Česonis, S. Franklin, and D. W. Franklin, “A simulated inverted pendulum to investigate human sensorimotor control,” *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 5166–5169, 2018 - © 2018 IEEE. Reprinted, with permission.

[81] J. Česonis, R. Leib, S. Franklin, and D. W. Franklin, “Controller gains of an inverted pendulum are influenced by the visual feedback position,” *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 5068–5071, 2019 - © 2019 IEEE. Reprinted, with permission.

[82] R. F. Reynolds and B. L. Day, “Direct visuomotor mapping for fast visually-evoked arm movements,” *Neuropsychologia*, vol. 50, no. 14, pp. 3169–3173, 2012.

[83] F. Crevecoeur, R. J. Sepulchre, J. L. Thonnard, and P. Lefèvre, “Improving the state estimation for optimal control of stochastic processes subject to multiplicative noise,” *Automatica*, vol. 47, no. 3, pp. 591–596, 2011.

[84] J. McIntyre, M. Zago, A. Berthoz, and F. Lacquaniti, “Does the brain model Newton’s laws?,” *Nature Neuroscience*, vol. 4, no. 7, pp. 693–694, 2001.

[85] M. Zago, G. Bosco, V. Maffei, M. Iosa, Y. P. Ivanenko, and F. Lacquaniti, “Internal Models of Target Motion: Expected Dynamics Overrides Measured Kinematics in Timing Manual Interceptions,” *Journal of Neurophysiology*, vol. 91, no. 4, pp. 1620–1634, 2004.

[86] N. Benguigui, H. Ripoll, and M. P. Broderick, “Time-to-Contact Estimation of Accelerated Stimuli Is Based on First-Order Information,” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 29, no. 6, pp. 1083–1101, 2003.

[87] J. Kodl, G. Ganesh, and E. Burdet, “The CNS stochastically selects motor plan utilizing extrinsic and intrinsic representations,” *PLoS ONE*, vol. 6, no. 9, pp. 1–10, 2011.

[88] J. P. Gallivan, L. Logan, D. M. Wolpert, and J. R. Flanagan, “Parallel specification of competing sensorimotor control policies for alternative action options,” *Nature Neuroscience*, vol. 19, no. 2, pp. 320–326, 2016.

[89] D. W. Franklin, J. Česonis, S. Franklin, and R. Leib, “A technique for measuring visuomotor feedback contributions to the control of an inverted pendulum,” in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2019, Berlin, Germany, July 23-27, 2019*, pp. 1513–1516, IEEE, 2019.

[90] JASP Team, “JASP (Version 0.14.1)[Computer software],” 2021.

[91] C. R. Lowrey, J. Y. Nashed, and S. H. Scott, “Rapid and flexible whole body postural responses are evoked from perturbations to the upper limb during goal-directed reaching,” *Journal of Neurophysiology*, vol. 117, no. 3, pp. 1070–1083, 2017.

[92] S. V. Poscente, R. M. Peters, J. G. Cashaback, and T. Cluff, “Rapid Feedback Responses Parallel the Urgency of Voluntary Reaching Movements,” *Neuroscience*, 2021.

6 Appendix

6.1 Permission to publish

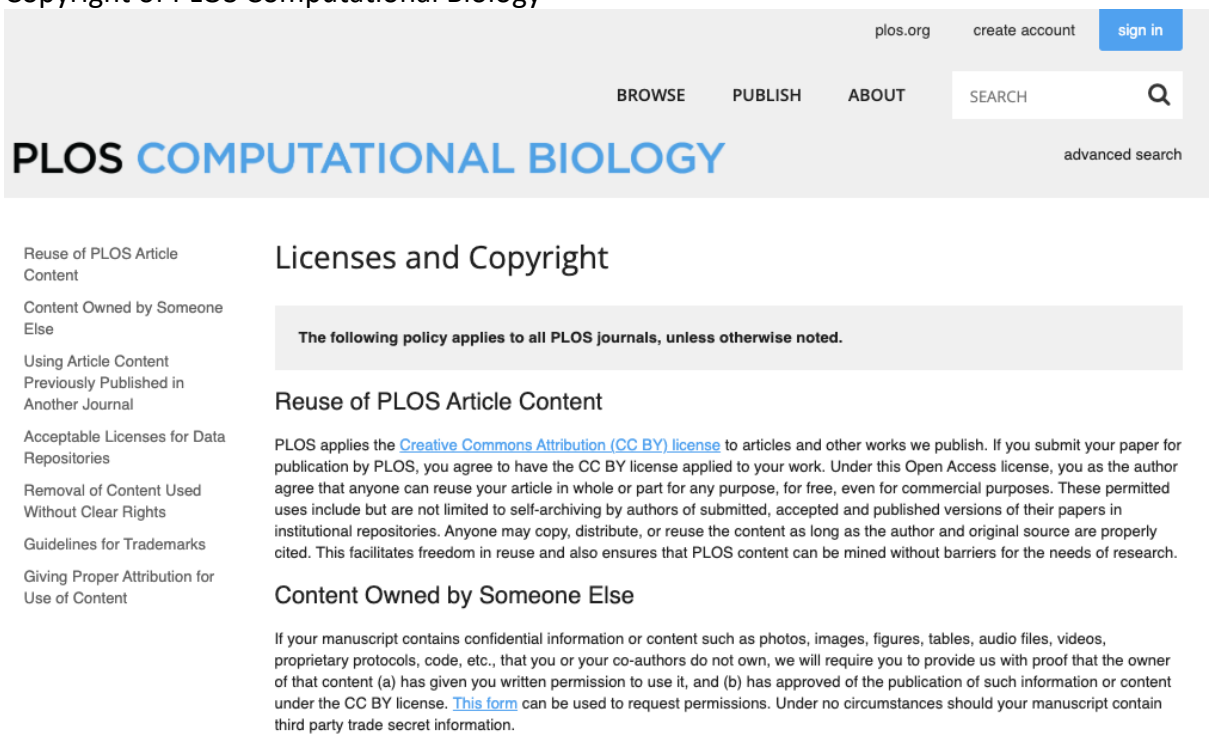
Copyrights of publishing journals:

1. eNeuro, Neurons, Behavior, Data Analysis and Theory (NBDT), and PLOS Computational Biology are OpenAccess journals. All articles published in these journals are subject to CC-BY-4.0 licence, meaning that articles can be freely distributed, provided a proper Attribution is given:

Copyright of eNeuro:

Accessed on 30.5.2022 (https://www.eneuro.org/content/7/2/ENEURO.0514-19.2020/tab-article-info)

Copyright of PLOS Computational Biology



Accessed on 30.5.2022 (https://journals.plos.org/ploscompbiol/s/licenses-and-copyright)

Copyright of NBDT:

About the Journal

Neurons, Behavior, Data Analysis and Theory (NBDT) publishes good computational neuroscience. It is a community journal - its editorial board decides what is in scope. We are especially interested in data, behavior, theory, analysis, modeling. This explicitly includes behavior-only studies, software development, systems neuroscience, theory-only studies, and data-analysis studies. All NBDT articles are peer reviewed.

This is an open access journal which means that all content is freely available without charge to the user or his/her institution. Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the articles, or use them for any other lawful purpose, without asking prior permission from the publisher or the author. This is in accordance with the BOAI definition of open access.

Authors need to choose CC-BY as license (also see https://arxiv.org/help/license) and should indicate the license chosen in the manuscript. As NBDT is an arxiv overlay journal that choice is needed in any case.

NBDT is a true community journal, open in every way.
Open science. Open data. And, we are openly asking for self-nominations for potential reviewers:

Accessed on 30.5.2022 (https://nbdt.scholasticahq.com/about)

2. **The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:**

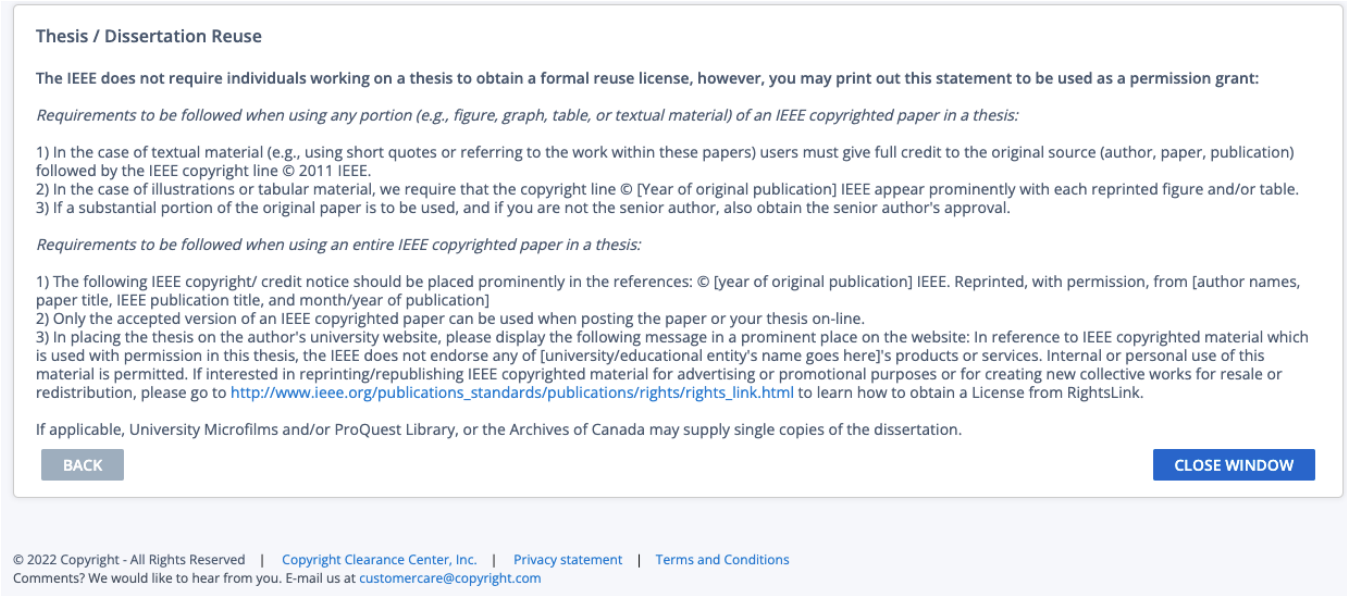
   Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

   1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
   2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
   3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

   Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

   1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
   2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
   3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

Copyright statement IEEE:



Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

BACK CLOSE WINDOW

Accessed on 30.5.2022 (https://s100.copyright.com/AppDispatchServlet#formTop via https://ieeexplore.ieee.org/document/8513434 and https://ieeexplore.ieee.org/document/8857886)