bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Reward-dependent selection of feedback gains impacts rapid motor decisions

Abbreviated title: Competition between movement vigor and flexibility

Antoine De Comite, Frédéric Crevecoeur, and Philippe Lefèvre

Institute of Information and Communication Technologies, Electronics and Applied Mathematics (ICTEAM), Institute of Neuroscience (IoNS), Université catholique de Louvain, Belgium

ADC, FC and PL designed the study, analysed the data and drafted the manuscript, ADC collected the data

Corresponding author : *Philippe Lefèvre,*

4 Avenue Georges Lemaître, 1348 Louvain-la-Neuve, Belgium

Mail : *philippe.lefevre@uclouvain.be*

Phone : *+32 10 47 23 82*

Number of figures : 6

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Expected reward is known to affect planning strategies through modulation of movement vigor. Strikingly, although current theories suggest that movement planning consists in selecting a goal-directed control policy, the influence of reward on feedback control strategies remains unknown. Here we investigated this question in three human reaching experiments. First, we varied the explicit reward associated with the goal target and found an overall increase in movement vigor for higher reward targets, highlighted by larger velocities, feedback responses to external loads, and background muscle activity. Then, assuming that larger feedback gains were used to reject perturbations, we sought to investigate whether this effect hindered online decisions to switch to a new target in the presence of multiple successful goals. We indeed observed idiosyncratic switching strategies dependent on both target rewards and movement vigor, such that the more vigorous movements were less likely to switch to a new goal following perturbations. To gain further insight into a causal influence of movement vigor on rapid motor decisions, we demonstrated that biasing the baseline activity and reflex gains by means of a background load evoked a larger proportion of target switches in the direction opposite to the background load associated with lower muscle activity. Our results highlight the competition between movement vigor and flexibility to switch target during movement.

Significance statement

Humans can modulate their movement vigor based on the expected reward. However, a potential influence of reward on feedback control has not been documented. Here we investigated reaching control strategies in different contexts associated with explicit rewards for one or multiple goals, while exposed to external perturbations. We report two strategies: reward could either invigorate feedback gains, or promote flexible switches between goals. The engagement of peripheral circuits in the modulation of feedback gains was confirmed by the application of a background load that biased feedback vigor directionally, which evoked differences in switching behavior in the opposite direction. We conclude that feedback vigor and flexible changes in goal are two competing mechanisms to be selected when one interacts with a dynamic environment.

Introduction

From the toddler picking their favorite toys to the footballer selecting the best path through opponents, humans manifest the exquisite ability to plan movements. Movement planning integrates many factors related to the task goal and to the environment. Amongst them, the reward associated with the task influences movement planning. Indeed, past studies have reported an increase in movement vigor associated with increasing reward in saccadic eye movements (Manohar et al., 2015, 2017) and upper limb reaching movements (Esteves et al., 2016; Summerside et al., 2018). It was also reported that the selection of the best alternative between different options was biased toward movements associated with the highest reward (Trommershäuser et al., 2003, 2008). Similarly, when humans have to select a target, their choices are biased by parameters such as the biomechanical costs incurred when reaching to each potential option, resulting in target selection toward less effortful movements (Cos et al., 2011; Morel et al., 2017). Therefore, the commitment to an action results from a distributed consensus between low level sensorimotor representations of movement costs and higher level cognitive representations of their outcomes (Cisek, 2007, 2012; Gold & Shadlen, 2007).

Here we tested this theory in a dynamical context by probing the effect of movement reward on movement execution and online motor decisions. Indeed, recent studies have sought to investigate whether and how much the factors that characterize action selection during movement planning could also influence movement execution when the hand has already started moving. A first body of work has shown that dynamical changes in target selection can be triggered by mechanical (Nashed et al., 2014) or visual (Kurtzer et al., 2020; Michalski et al., 2020) perturbations occurring during

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

movement. More recently, a study investigated whether cognitive factors, such as the reward associated with the different options, also influenced online motor decisions when a mechanical perturbation was applied (Cos et al., 2021). They revealed that the reward distribution across the potential options influences these decisions and suggested that there could be a link between the state of the limb at perturbation onset and the outcome of the decision.

In the present work, we addressed the relationship between feedback control and online motor decisions by applying perturbations while participants performed reaching movements toward one or several targets that differed explicitly by their associated rewards.

We first reproduced previous findings of reward-related increase in velocity toward the target. Importantly, we uncovered that this modulation was associated with an increase in feedback gains and muscle activity. In a second experiment, we found out that this overall increase in movement vigor was indeed in competition with the potential selection of a new target. Our third experiment confirmed that biases in feedback gains induced experimentally were negatively correlated with the ability to switch goal during movement. Our findings demonstrate that the increase in movement vigor associated with reward was detrimental to the flexible ability to switch target during movement. They also demonstrate a dynamic interaction between sensorimotor and cognitive factors, which influences not only target selection but also the way we respond to perturbations.

Methods

Participants

A total of 53 participants were enrolled in this study and took part to one of the three experiments. The first group performed Experiment 1 and included 14 right-handed participants (7 females) ranging in age from 21 to 27. The second group performed Experiment 2 and included 20 right-handed participants (14 females) ranging in age from 20 to 46. The last group performed Experiment 3 and included 19 right-handed participants (11 females) ranging in age from 18 to 52. Participants were naïve to the purpose of the experiments and had no known neurological disorder. The ethics committee of the Université catholique de Louvain (UCLouvain) approved the experimental procedures.

Experiments

For the three experiments, participants sat on an adjustable chair in front of a Kinarm end-point robotic device (KINARM, Kingston, ON, Canada) and grasped the handle of the right robotic arm with their right hand. The robotic arm allowed movements in the horizontal plane and direct vision of both the hand and the robotic arm was blocked. Participants were seated such that at rest their elbow formed an angle of approximately 90° pointed downward and their forehead rested on a soft cushion attached to the frame of the setup. A virtual reality display placed above the handle allowed

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the participants to interact with virtual targets. A white dot of 0.5 cm radius corresponding to the position of the handle was shown on this display during the whole experiment.

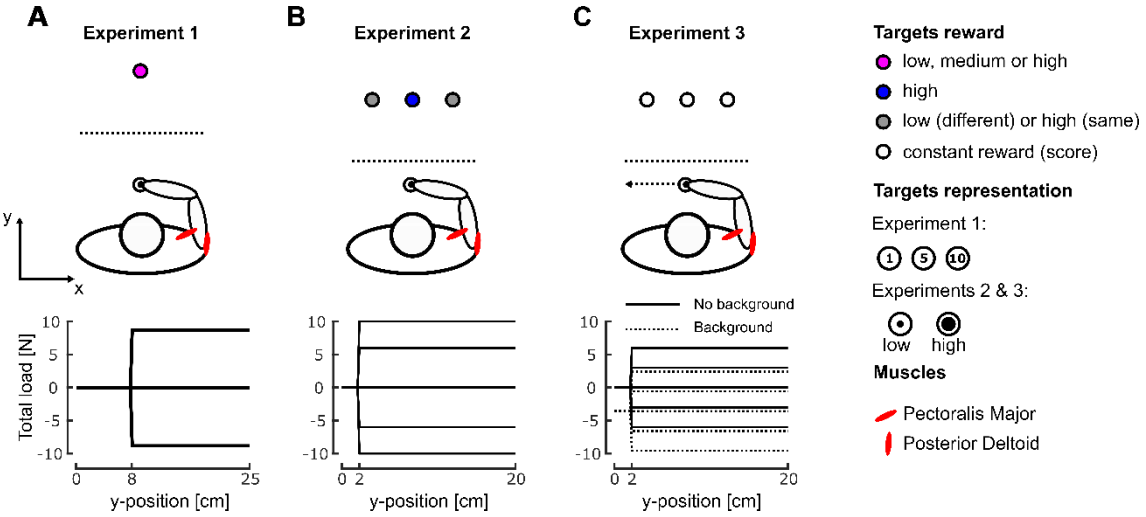

*Figure 1 : Task paradigms –. **A** Representation of the task paradigm of Experiment 1. Participants controlled a hand-aligned cursor represented by the black dot on a virtual reality display. They had to reach for the goal target, represented by the purple goal target in front of them. This goal target could have a low, medium or high reward (1, 5 or 10 points). The bottom part of the panel represents the load profiles that participants could experience. **B** Representation of the task paradigm of Experiment 2. Participants had to reach for any of the three targets presented in front of them. The central target always had a high reward whereas the two others either had a low or a high reward. The bottom part of the panel represents the perturbation load that participants could encounter during movements. **C** Representation of the task paradigm of Experiment 3. Participants had to reach for any of the three targets presented in front of them. During the second half of the trials a background load force directed leftward was applied prior and during the movement (dashed line bottom panel). The bottom part of the panel represents the possible profiles of the total load forces (perturbation load + background load). EMG data from Pectoralis Major and Posterior Deltoid were collected during all Experiments.*

Experiment 1

In Experiment 1 (Figure 1A top), participants (N=14) were instructed to perform reaching movements to a small circular goal target (1.5 cm radius) located at 25 cm in the y-direction from the home target, a red disk of 1.5 cm radius. Participants had first to put the hand-aligned cursor in the home target, which turned green as they reached it. After a random time delay (anywhere from 1 to 2s), the goal target appeared as a red disk containing a number (1, 5 or 10) that corresponded to the reward participants would receive if they reached and stabilised within the target for a prescribed time window. Reaction time was not constrained and participants could start the movement whenever they wanted. Following the exit of the home target, participants had up to 600ms to reach the goal target and keep the cursor inside for at least 500ms. The goal target turned green at the end of successful trials, or remained red otherwise. During movements, a mechanical perturbation load could be applied to participants' hand (33 % of the trials). This load consisted of a lateral step force of ±9 N, with a 10ms linear build-up, aligned with the x-axis. This force was triggered when the hand-aligned cursor crossed a virtual line located at 8 cm from the center of the home target (Figure 1A bottom). Unperturbed and perturbed trials as well as trials with different rewards were randomly interleaved such that participants could not predict the occurrence or the direction of the perturbations. Participants started with a 27-trials training block in order to become familiar with the task and the force intensity of perturbation loads. After completing this training block, they performed 6 blocks of 72 trials. Each 72-trials block included: 48 unperturbed trials (16 with each target reward) and 24 trials which contained mechanical perturbations (leftward or rightward, 8 of each reward condition). Participants performed a total of 432 trials, including 24 for each perturbed condition (direction of the mechanical perturbation and value of the target reward). A total score corresponding to the cumulative sum of individual movement rewards was projected next to the goal target. Participants were compensated for their participation according to a conversion of this total

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

score. This conversion was calculated such that each participant received between 10 and 15 € as an incentive to score a maximum number of points during the experiment.

Experiment 2

Experiment 2 was designed to assess the effect of reward on online motor decisions. Instead of reaching to a single target, participants (n=20) were instructed to perform reaching movements to any of three circular targets (1.5 cm radius) located at 20 cm in the y-direction from the same home target as in Experiment 1 (Figure 1B top). As in Experiment 1, the goal targets appeared after participants stabilised the hand-aligned cursor in the home target. All three goal targets appeared in each trial, the central one being aligned along the y-axis with the home target and the other two equidistant from this central target at 9 cm along the x-axis. These targets were presented as an inner disk of radius 0.7 or 1.2 cm inside an outer circle of radius 1.5 cm. The purpose of the inner disk was to show the reward associated with the target: the larger the diameter of this disk, the higher the reward. There were two different conditions of reward: either all the targets had the same large reward (*same values* condition) or only the central target had a large reward while the other two had smaller rewards (*different values* condition). In a pilot study, we considered a third reward configuration: the central target had a small reward while the other two had larger rewards. We observed that, in this third configuration, the behavior in the absence of perturbation load was biased toward the lateral targets. We therefore decided to exclude this condition to keep the conditions in which participants spontaneously reached for the center target for the largest proportion of trials in the absence of any perturbation load. After a random time delay (anywhere between 1.5 and 3s), the inner disks of the goal targets turned white and participants had to reach any of these within 400ms to 1000ms to pass the trial. The trial was successfully completed if participants reached any goal target in the prescribed time window and stabilised the cursor in it for 500ms. The inner disks of the goal targets turned green if the trial was successful and red otherwise. As in Experiment 1, a mechanical perturbation load could be applied to participant's hand (50 % of the trials, ±6 N or ±10 N, 10 ms build-up aligned with the x-axis Figure 1B bottom). This perturbation was triggered when the hand-aligned cursor crossed a virtual line located at 2 cm from the home target. Unperturbed and perturbed trials as well as trials with different reward distributions and force intensities were randomly interleaved. Participants started with a 58-trials training block followed by 6 blocks of 80 trials. Each 80-trials block included: 40 unperturbed trials and 40 trials which contained mechanical perturbations. Participants performed a total of 480 trials including 30 trials of each perturbation condition (reward condition and mechanical perturbation condition). Participants were compensated for their participation using the same conversion rule as in Experiment 1.

Experiment 3

The third experiment was a variant of Experiment 2 and was designed to test the possible impact of muscle activity on online motor decisions by applying a background force laterally to the reach path (Figure 1C top). Participants had to perform reaching movements to any of the three targets, located as in Experiment 2. These targets were identical to the large reward target of Experiment 2 and the time course of events in the trial was similar as well except that a leftward background mechanical load of 4N was applied as volunteers reached the home target and remained on throughout the trials. As in the Experiment 2, a mechanical perturbation load could be applied to participant's hand during movement (33% of the trials). This load consisted of a ±3 N or ±6 N with a 10ms build-up triggered when the hand-aligned cursor crossed a line located at 2 cm from the home target (Figure 1C bottom) and was added to the background load. Participants first performed a 21-trials training block which did not involve background load. After completing this training, participants performed 4 blocks of 60 trials which did not include the background load. Each 60-trials block included 40 unperturbed trials and 20 trials with mechanical perturbations. After these 60-trials blocks,

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

participants performed a second 21-trials training block which included the background load. Once this second training block was completed, participants performed a second set of 4 blocks of 60 trials which included the background load. They thus performed a total of 480 trials amongst which 24 of each condition (with different perturbation loads and background load on or off). To motivate participants, a score corresponding to their number of successful trials was projected next to the goal targets. Participants were compensated a fixed amount for their participation.

Data collection and analysis

Raw kinematics data was sampled at 1kHz and low-pass filtered using a 4th order double-pass Butterworth filter with cut-off frequency of 20 Hz. Hand velocity along the y-axis was computed from numerical differentiation of the position data using a 4th order centered finite difference.

Surface EMG electrodes (Bagnoli surface EMG sensor, Delsys INC. Natick, MA,USA) were used to record muscles activity during movements. We measured the Pectoralis Major (PM) and the Posterior Deltoid (PD) based on previous studies (Crevecoeur et al., 2019; Lowrey et al., 2017) that showed that in this configuration they are stretched by the application of forces opposite to their action, and therefore largely recruited for the feedback responses. Before applying the electrodes, the skin of the participant was cleaned and abraded with cotton wool and alcohol. Conduction gel was applied on the electrodes to improve the quality of the signals. The EMG data were sampled at a frequency of 1kHz and amplified by a factor of 1000. A reference electrode was attached to the right ankle of the participant. Raw EMG data from the pectoralis major (PM) and posterior deltoid (PD) were band-pass filtered using a 4th order double-pass Butterworth filter (cut-offs: 20 and 250 Hz), rectified, aligned to force onset and averaged across trials or time windows as specified in the Results section.

EMG data were normalized for each participant to the average activity collected when participants maintained postural control at the home target against a constant force of 9 N. Data from the pectoralis major were normalized by the EMG activity in the same muscle while performing postural control against a rightward force whereas data from the posterior deltoid were normalized by the EMG activity in the same muscle while performing postural control against a leftward force. This calibration procedure was applied after the second and the fourth blocks in the first two experiments and after the first, third, fifth and seventh blocks in the third experiment. Data processing and parameters extractions were performed using Matlab 2019a.

In Experiment 1, we fitted linear mixed models to determine the effect of the target reward on the kinematics and EMG activity. These models were fitted using the *fitlme* function of Matlab and the formula used was the following:

$$Parameter = \beta_0 + \beta_1 * Reward + \alpha_i \qquad (1)$$

The fixed predictors were the intercept (β_0) and the reward condition (β_1) while the participants were included as a random offset (α_i). For all linear mixed model analyses that we performed, we reported the estimate for β_1, the t-statistics for this estimate as well as the corresponding p-value and the $r^2$of the model. One-tailed paired t-tests were used for post-hoc analyses where we collapsed data across trials and participants to compare the different conditions. Effect size for these tests were reported using Cohen's d defined as the difference between the means of the two populations divided by the standard deviation of the whole sample.

To analyse the data from Experiments 2 and 3, we designed a multilinear logistic regression model to infer the effect of reward distribution and background load on target choice as the dependent variable, respectively. Considering that the dependent variable was a discrete variable (the chosen target), we use the following logistic regression model:

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$$\log\left(\frac{P(Lateral\ target)}{P(Central\ target)}\right) = \beta_0 + \beta_1 * Parameter_1 + \beta_2 * Parameter_2 \quad (2)$$

where the first effect (β_1) was the reward condition (Experiment 2) or the presence of a background load (Experiment 3) and the second effect (β_2) was the intensity of the perturbation load. For these logistic regressions, we reported the estimates for β_1 and β_2, their corresponding t-statistics as well as their p-value. For post-hoc analyses in Experiment 2, we used a one tailed Wilcoxon signed ranked test for which we reported the ranksum, the z-statistics when provided, the p-value as well as the effect size given by the Cohen's d as defined above. In order to investigate the asymmetry in the parameters β_1 obtained in Experiment 3, we used bootstrap resampling on the individual data to generate 1000 estimates of the β_1 parameter for each condition (leftward perturbation versus rightward perturbation). We then used two samples t-test to compare the bootstrapped distributions of these two parameters.

In order to determine the effect of the background load on the baseline muscle activity in Experiment 3, we fitted a linear mixed model with interaction terms following this equation:

$$EMG = \beta_1 * Background + \beta_2 * Muscle + \beta_{12} * Muscle{:}Background + \alpha_i \quad (3)$$

Where the first term (β_1) refers to the background condition, the second (β_2) to the muscle, the third one (β_{12}) to the interaction term and the last one (α_i) to the random offset of participants. For all these β, we reported their estimated value as well as their t-statistics, associated p-value and the r^2 of the model. Significance was considered at the level of p=0.05 even though we decide to exactly report any p-value that was larger than p=0.005 as previously proposed (Benjamin & Berger, 2018). In the figures, we reported significant differences for the level p=0.05 (*), p=0.01(**) and p=0.005(***).

Results

Experiment 1 : Influence of the target reward on feedback corrections during movement

To determine whether target reward influences feedback corrections during movement, participants were instructed to perform reaching movements to a goal target associated with a reward that could change across trials (see Methods). During movements, mechanical perturbation loads could be applied to reveal feedback corrections. The occurrence of feedback corrections was assessed by looking at movement kinematics and EMG responses of the muscles stretched by the perturbations.

The mean hand path trajectories across participants are represented in Figure 2A for the different perturbations and reward conditions. Consistent with previous work (Shadmehr et al., 2016; Summerside et al., 2018) we observed a significant increase in forward peak velocity with increasing reward values. Figure 2B shows the differences in the forward velocities between the high (dash-dot lines) or medium (full lines) and low reward conditions unperturbed (top) and perturbed (bottom) trials. The peak forward velocity increased with increasing reward value both for unperturbed (Figure 2D top, linear mixed model: β_1=0.013, t=6.51, p<0.005, r^2=0.76) and perturbed (Figure 2D bottom, linear mixed model, right: β_1=0.014, t=3.76, p<0.005, r^2=0.78, left: β_1=0.018, t=4.68, p<0.005, r^2=0.79) trials. Post-hoc comparisons between low and high reward conditions revealed a significant increase of peak velocity with reward for all perturbation conditions (one-tailed paired t-tests,

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

unperturbed: t=-7.48, p<0.005, d=0.12, left: t=-5.37, p<0.005, d=0.16 and right: t=-3.99, p<0.005, d=0.13).

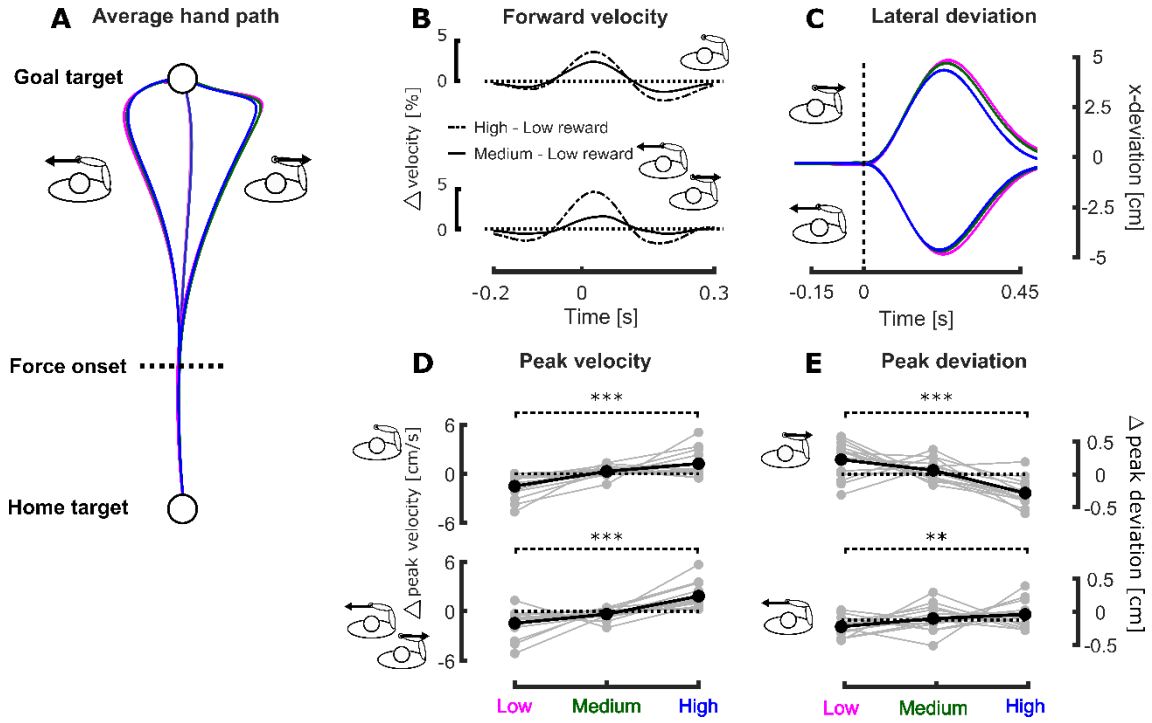

*Figure 2: Experiment 1, kinematics – **A** Mean hand path across participants for the different conditions of the first experiment. The purple, green and blue traces respectively correspond to the low, medium and high reward conditions. The dashed line represents the onset of the mechanical perturbation. **B** Mean difference in forward velocity between the high and low (dash-dot line) and medium and low (full line) for the unperturbed (top) and perturbed (bottom) trials. The time axis is aligned on the force onset. **C** Mean hand deviation across participants for the perturbed trials. The hand deviation has been obtained by subtracting the mean hand path to the perturbed hand path in the same reward condition for every subjects. The top part of the graph represents the trials perturbed to the right whereas the bottom part of the graph represents the trials perturbed to the left. **D** Group mean (black) and individual means (grey) of the differential forward peak velocity for the unperturbed trials (top) and perturbed trials (bottom) as a function of the reward condition with respect to average forward peak velocity. **E** Group mean (black) and individual means (grey) of the difference in hand deviation with respect to the mean hand deviation for leftward (top) and rightward (bottom) perturbation in the three reward conditions with respect to the average hand deviation.*

The effect of the mechanical perturbation on the movement kinematics was also dependent on the reward value. Indeed, the lateral hand deviation induced by the mechanical perturbation (Figure 2C), computed as the difference between the hand paths in the perturbed conditions and the mean hand path in the corresponding unperturbed reward condition for each participant, depended significantly on the reward condition. For both perturbation directions (Figure 2E top for leftward and bottom for rightward perturbations), we observed a significant decrease in the maximal hand deviation along the x-axis with increasing reward value (linear mixed models, right: β_1=-0.0025, t=-4.98, p<0.005, r^2=0.38 and left : β_1=-0.0009, t=-2.25, p=0.024, r^2=0.35). Post-hoc comparisons between low and high reward conditions revealed a significant decrease for both perturbation directions (one-tailed paired t-tests, right: t=5.31, p<0.005, d=0.14 and left: t=2.34, p=0.009, d=0.3).

Based on these kinematics analyses and previous studies showing that faster movements and smaller hand deviations induced by perturbations are correlated with high EMG activity (Crevecoeur et al., 2019) we hypothesised that the EMG activity in PM and PD during movement scaled with increasing

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

reward. We investigated this effect both for baseline activity measured during unperturbed trials and for feedback responses to perturbation loads.

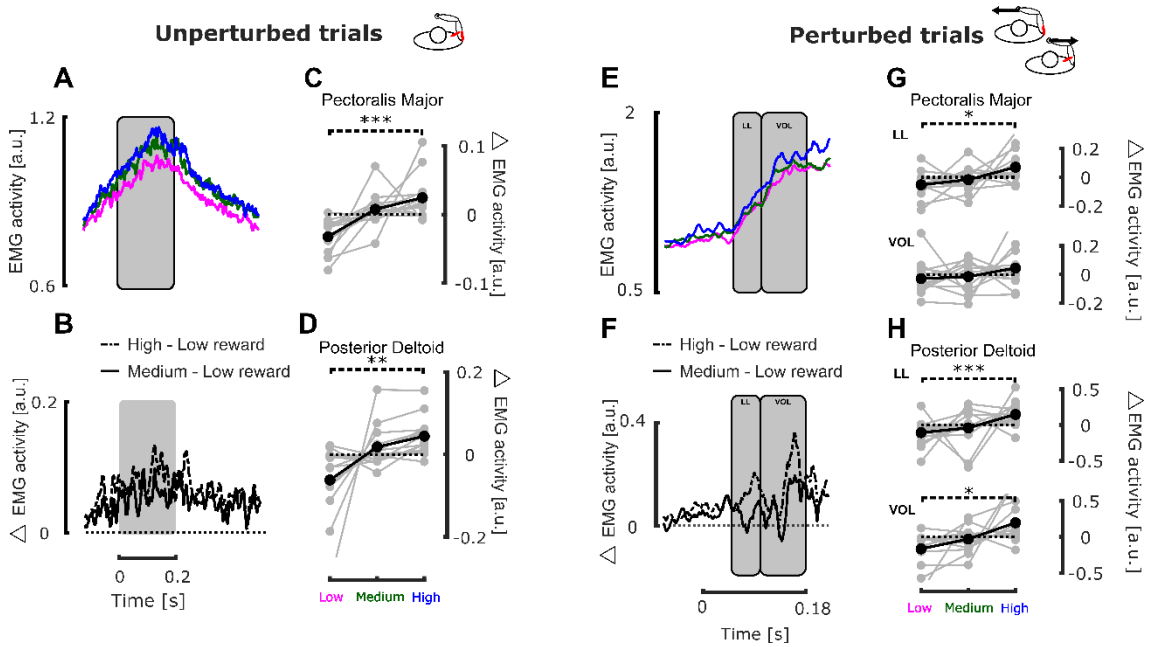

*Figure 3 : Experiment 1, EMG activity – **A** Mean EMG activity collapsed across muscles and participants for unperturbed trials. The time axis is aligned on force onset. **B** Mean differences in EMG activity collapsed across muscles and participants between high and low (dash-dot line) and medium and low (full line) reward conditions for unperturbed trials. **C** Group mean (black) and individual means (grey) Pectoralis Major EMG activity binned between 0 and 200 ms after force onset for unperturbed trials. **D** Group mean (black) and individual means (grey) Posterior Deltoid EMG activity binned between 0 and 200 ms after force onset for unperturbed trials. **E** Mean EMG activity collapsed across muscles and participants while responding to perturbations. Only the activities of the agonist muscles (ie. PM for rightward perturbation and PD for leftward perturbation) were used. **F** Mean differences in EMG activity collapsed across muscles and participants between high and low (dash-dot line) and medium and low (full line) reward conditions for agonist muscles in presence of perturbation load. **G** Group mean (black) and individual means (grey) differential EMG activity in PM binned in the long latency (50-100 ms, top) and voluntary epochs (100-180 ms, bottom) as a function of the reward condition. **H** Group mean (black) and individual means (grey) of the differential EMG activity in PD binned in the long latency (50-100 ms, top) and voluntary epochs (100-180 ms, bottom) as a function of the reward condition.*

We observed a positive relationship between the EMG activity during unperturbed trials and the value of the target reward. Figure 3A represents the mean EMG activity collapsed across muscles and participants for unperturbed trials while the differences between these collapsed EMG activities in the high (dash-dot line) or medium (full line) and the low reward condition are represented in Figure 3B. We binned the EMG activity of each trial in a time bin ranging from 0 to 200 ms after perturbation onset (gray rectangle in Figure 3A and 3B) and fitted a linear mixed model (see Methods) on these binned values to determine whether reward had an influence on the EMG activity (deviations from the mean binned EMG activity in the different reward conditions are represented in Figure 3C and 3D for PM and PD respectively). We observed an increase in EMG activity with the reward in both muscles (PM: β_1=0.028, t=4.603, p<0.005, $r^2 = 0.68$, PD: β_1= 0.053, t=5.98, p<0.005, r^2=0.66). Post-hoc analyses performed on individual data showed that EMG activity was larger in the high reward condition than in the small one for both muscles (one-tailed paired t-tests: pectoralis, t=4.14, p<0.005, d=-0.118 and deltoid, t=-2.92, p=0.0059, d=0.1653).

The EMG response to mechanical perturbation in the agonist muscles was also modulated by the reward value. Indeed, linear mixed model analyses performed on the responses measured in PM and PD, when respectively a rightward or leftward perturbation occurred, showed a significant increase of EMG activity with increasing reward in the long-latency epochs (50-100 ms). We reported the EMG

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

activities collapsed across muscles and participants in Figure 3E as well as the difference in these activities between the high (dash-dot line) or medium (full line) and the low reward condition in Figure 3F. For each perturbation direction, we binned the EMG activity of the stretched muscle in the long latency (LL 50-100 ms after force onset) and voluntary (VOL 100-180 ms after force onset) epochs. Figure 3G and 3H respectively represent the deviation from the mean binned EMG activity in these two time bins (LL top and VOL bottom) for PM and PD in the different reward conditions. In PM we observed a significant increase in the LL window (mixed model: β_1=0.0615, t=2.89, p<0.005, r^2=0.64) but even though a positive tendency emerged in the VOL window, no significant increase was observed (mixed model: β_1=0.036, t=1.616, p=0.106, r^2=0.69). Individual pairwise post-hoc comparisons between low and high conditions confirmed these findings (one-tailed paired t-tests: LL, t=-2.48, p=0.0137, d=0.1592 and Vol, t=-1.18, p=0.128, d=0.08). The same holds for PD in which we found a significant increase of EMG activity in LL window with the reward (mixed model: β_1=0.216, t=2.12, p=0.034, r^2=0.63) but no significant effect in the VOL window (mixed model: β_1=0.181, t=1.887, p=0.059, r^2=0.77). In this case however, the individual pairwise comparisons between low and high conditions revealed a significant increase in both time windows (one-tailed paired t-tests: LL, t=-3.68, p<0.005, d=0.11 and VOL, t=-2.57, p<0.01, d=0.07).

Therefore, the results of Experiment 1 revealed that the value of the target reward influenced both the movement kinematics and the EMG activity recorded during movement. Indeed, we showed that the hand deviation induced by mechanical perturbations decreased with increasing reward value for both rightward and leftward perturbations. Moreover, the forward peak velocity of reaching movement increased with increasing reward value. Finally, EMG activity in both PM and PD increased with increasing reward value for unperturbed trials and in the long-latency response window for perturbed movement when the muscles were stretched by the perturbation. The modulation of forward hand velocity and baseline EMG activity which also produced increases in feedback responses to perturbation loads is consistent with an increase in control gains previously observed in uncertain dynamical contexts and interpreted as a robust control strategy (Crevecoeur et al., 2019).

Experiment 2: Influence of the reward of the different options on online motor decisions

In Experiment 1, we showed that reward of the goal target has an influence on the way humans perform reaching movements to this target similarly to other task parameters such as target shape, presence of obstacles, etc. Moreover, previous studies have shown that these task parameters that modify the control strategies could also influence online motor decisions (Nashed et al., 2014). We therefore designed a second experiment to determine whether reward could also influence online motor decisions. In this second experiment, participants had to reach to any of three potential targets aligned orthogonally to the main reaching direction (see Methods). The central target always had a large reward while the two lateral targets could either have lower reward or a reward equal to that of the central target. We assessed the effect of the difference between central and lateral rewards on online motor decisions by investigating the frequencies of reaching for the lateral targets. Mechanical perturbations that could occur during movement were used to evoke changes in goal target. Because perturbations were unpredictable, a change in reaching frequency for the lateral targets dependent on the perturbation load was indicative of a perturbation-mediated change in goal that occurred during movement. The biomechanical and EMG states at perturbation onset in the

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

same condition were also investigated to determine whether they had an influence on the future decision.

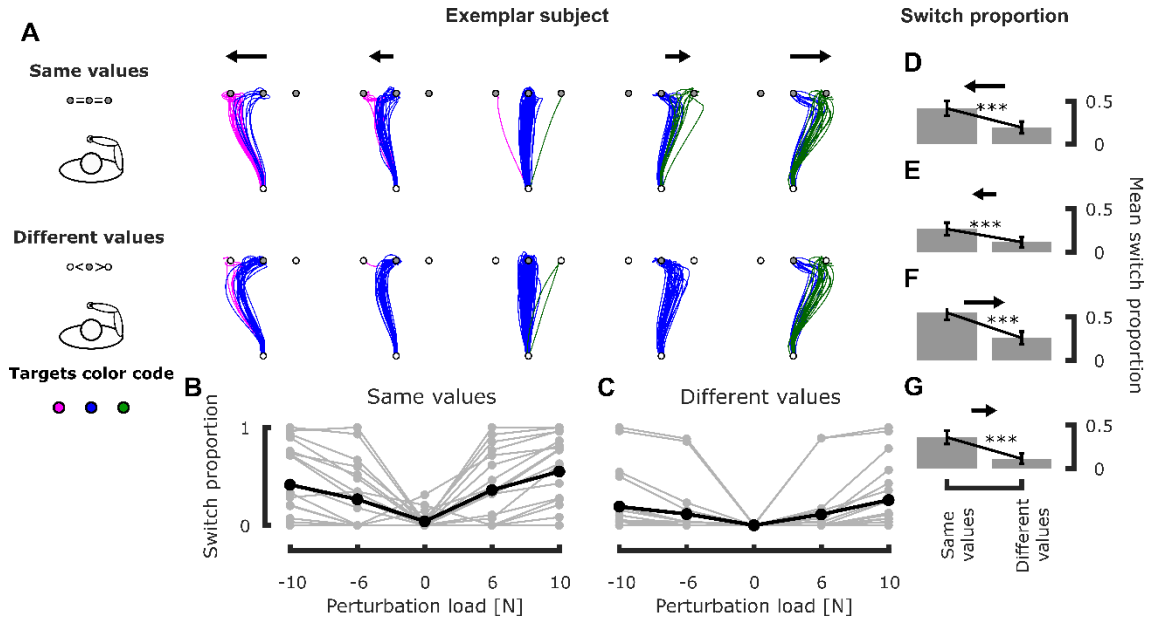

*Figure 4 : Experiment 2, Kinematics – **A** Representation of hand path of individual trials for a representative subject in the same condition on top (all the targets had the same reward) and in the different condition on bottom (the central target had a higher reward than the other two). The different columns represent the different force levels (from right to left, large leftward perturbation to large rightward perturbation). Magenta, blue and green paths respectively represent the paths that reached the left, central and right targets. **B** Group mean (black) and individual means (grey) of the switch proportion (i.e. fraction of trials that reached either the left or right targets) as a function of the applied load for the same condition. **C** Group mean (black) and individual means (grey) of the switch proportion (i.e. fraction of trials that reached either the left or right targets) as a function of the applied load for the different condition. **D-G** Comparison of the switch proportion for the same (left) and different (right) conditions for the trials with large leftward force, small leftward force, small rightward force and large rightward force respectively.*

First, we observed that the reward of the lateral targets had a clear effect on the frequency of trials that ended on these targets. Figure 4A represents the hand paths of a typical participant toward the different targets in various conditions. In general, in the absence of perturbation, subjects tend to reach to the central target except for some trials (<1% in the *different values* condition and 8 % in the *same values* condition)**.** In all cases, the frequency of lateral target increased with the magnitude of the perturbation (top and bottom part of Figure 4A for *same* and *different values* conditions respectively). In addition, there was a significant effect of the lateral targets reward on the frequency of lateral target reach: lower frequencies for different rewards. In order to determine the significance of these effects, we identified for each trial the target that was reached at the end of the movement and fitted a multilinear logistic regression on these data to determine whether the reward condition and the force had an influence on the target reached (see Methods). We observed a significant effect of the reward for both the left ($\beta_1 = 1.103, t = 10.84, p < 0.005$) and right ($\beta_1 = 1.666, t = 18.45, p < 0.005$) targets versus the central one. These positive values indicate that the reach proportion to the lateral targets is larger in the same than in the different condition (see Figure 4B and 4C for the *same values* and *different values* conditions respectively). The intensity of the perturbation loads also had a significant effect for both lateral targets versus the central one (left: β_1=-1.33, t=-26.82, p<0.005 and right: β_1=1.25, t=29.07; p<0.005). Due to the sign of the force kept in the regression model, in both cases the frequency of lateral target reach increased with the force magnitude in absolute value. Post-hoc analyses performed at fixed force levels showed significant effect of the reward condition on the reaching proportion for all the perturbed conditions. We observed a smaller reach proportion to the left target in the *different values* condition compared to

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the *same values* condition for both perturbation directions (one tailed Wilcoxon signed-rank test: ranksum=3, p<0.005, d=0.61 - see Figure 4D - and ranksum=1, p<0.005, d=0.49 – see Figure 4E - for loads of -10 and -6N respectively). The mirror effect was observed for the right target: a decrease in the reach proportion in the *different values* condition for both perturbation directions (one-tailed Wilcoxon signed-rank test: ranksum=0, p<0.005, d=0.74 – see Figure 4F - and ranksum=3, p<0.005, d=0.78 – see Figure 4G - for loads of 6 and 10 N respectively).

These results showed that participants took the reward distribution of the options offered by the three targets into account while deciding which target they should reach. The next question that we will address is whether any parameters linked to the current state of the limb could modify the decision between the different motor outcomes.

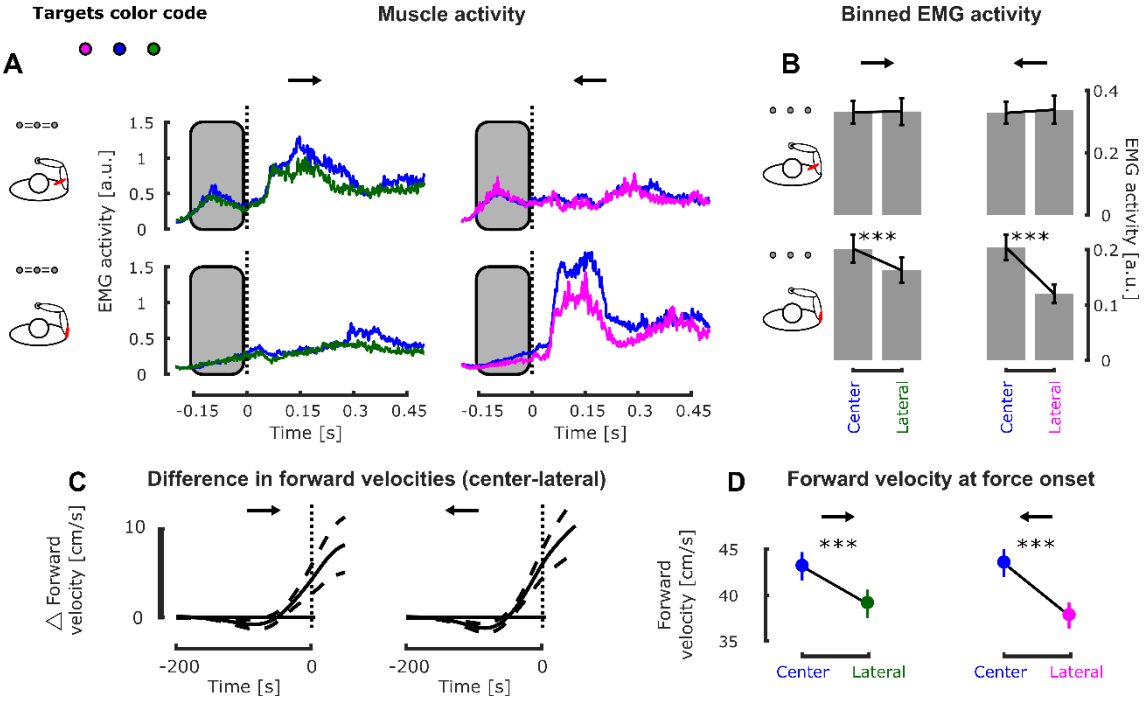

*Figure 5 : Experiment 2, EMG activity- **A** Mean EMG activity in Pectoralis Major (top) while responding to rightward (first column) and leftward perturbations (second column) and in Posterior Deltoid (bottom) while responding to rightward (first column) and leftward (second column) perturbations in the second experiment. The magenta, blue and green traces represent the mean EMG activity measured when participants reached the left, center or right target respectively. **B** Binned EMG activity before force onset in Pectoralis Major (top) and Posterior Deltoid (bottom) for the leftward and rightward perturbation loads, for the trials that reached the central (left bin) and lateral (right bin) targets. **C** Group mean and SEM of the differences in forward velocities across participants between the center and lateral trials for trials with rightward (left) and leftward (right) perturbations. **D** Comparison of the forward velocity at force onset for the trials that reached the central (blue) and lateral (green or purple) targets with a rightward or leftward perturbation load.*

Interestingly, we observed some differences in the kinematics and EMG activity before force onset between the trials that ended up in the central target and the ones that reached one of the two lateral targets. Figure 5A represents the mean EMG activity recorded in PM (top) and PD (bottom) in presence of mechanical perturbations (rightward, first column and leftward second column) across participants for the different targets (purple: left, blue: center and green: right) in the *same values* condition. No significant differences were observed in PM prior to force onset (-150ms to 0 ms, grey rectangle in Figure 5A) between the trials that reached the center target and the ones that reached the lateral targets (Figure 5B top) for both force directions (left: linear mixed model β_1=-0.019, t=-1.76, p=0.0782, r^2=0.62 and right: linear mixed model β_1=0.0054, t=0.89, p=0.3758, r^2=0.64). However, we observed an increase in the EMG activity of PD prior to perturbation onset for the trials that reached the center target compared to the ones that reached the lateral targets for both force

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

directions (Figure 5B bottom, left: linear mixed model $\beta_1 = 0.022$, t=3.78, p<0.005, r^2=0.68 and right: β_1=-0.051, t=-4.804, p<0.005, r^2=0.60). This increase in EMG activity for trials that ended at the central target was correlated with larger forward velocities at force onset. We reported in Figure 5C the differences in forward velocities between the center and the lateral trials for both perturbation directions. In presence of a leftward perturbation (Figure 5C and D, right panels), we observed a larger forward velocity at force onset for trials that end up at the center target compared to those that reached the lateral target (linear mixed model: β_1=-0.013, t=-3.347, p<0.005, r^2=0.54). The same holds for trials with rightward perturbations (Figure 5C and D, left panels - linear mixed model: β_1=0.040, t=9.476, p<0.005, r^2=0.57). Similar observations were reported in the *different values* conditions. Indeed, we observed an increase in EMG activities in both muscles for the trials that ended up at the center target compared to those that reached the lateral target (PM linear mixed model: β_1=-0.051, t=-4.81, p<0.005, r^2=0.60 and PD linear mixed model: β_1=0.022, t=3.78, p<0.005, r^2= 0.68). Moreover, some tendencies were observed in the forward speed for trials with rightward (β_1=-0.011, t=-2.019, p=0.0436, r^2=0.52) and leftward (linear mixed model β_1=0.012, t=1.95, p=0.0505, r^2=0.57). These results collected in the *different values* conditions have to be analysed with caution because of the low number of trials that ended up at one of the two lateral targets (7.5% in the *different values* and 23.5% in the *same values* condition).

We also tested whether the reward condition (ie. *same values* and *different values*) modified movement vigor by comparing the forward velocities and muscle activities at force onset between both reward conditions. We did not observe any difference in forward velocities between both reward conditions at perturbation onset as reported by mixed effect models (t=0.60, p=0.54, r^2=0.03). Similarly, we did not observe any differences in EMG activities averaged during the 50ms preceding perturbation onset as revealed by mixed model analyses (PM: t=-0.3962, p=0.69, r^2=0.28 and PD: t=0.07, p=0.93, r^2=0.06). This absence of correlation between the reward condition and movement vigor was interesting as it confirmed that we did not introduce any experimentally induced modulation of vigor in our paradigm. The differences in switching frequencies observed between the *same* and *different values* conditions are therefore attributable to the reward distribution and to vigor variability within both reward conditions.

This second experiment showed that humans take reward into account to respond to perturbations and potentially change target goal during movement. More specifically, participants will tend to reduce their frequency of reaching toward targets that have a lower reward. We also showed that some parameters measured before the perturbation onset are correlated with the target that is reached at the end of the trial.

Experiment 3: Effect of the pre-activation of muscle on the motor decision

An outstanding question is when was the decision made to switch target. Did participants decide to change after the perturbation, or did they plan to change prior to movement? On the one hand, in this experiment as in previous reports (Nashed et al., 2014), changes in goal target depend on the occurrence and magnitude of the force so it is at least partially determined by sensory information collected during movement. On the other hand, the observation that the switch also depended on the baseline activity suggests that there could be an influence of the state of the limb from the beginning of the movement on the decision. We wanted to investigate this possibility in Experiment 3. This experiment was specifically designed to investigate whether the pre-activation of PD prior to movement onset could bias the frequency of target switches. Participants had to reach any of the three targets located at the same position as in Experiment 2. All targets had the same reward in this experiment. During movement, mechanical perturbation loads could push participant's hand orthogonally to the main reaching decision. During half of the trials, a leftward background load was applied to participant's hand throughout movement evoking a background activation to counter the

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

background load (see Methods). We assessed the effect of pre-activation of PD by investigating the reach proportions to the lateral targets as a function of force intensity and background condition.

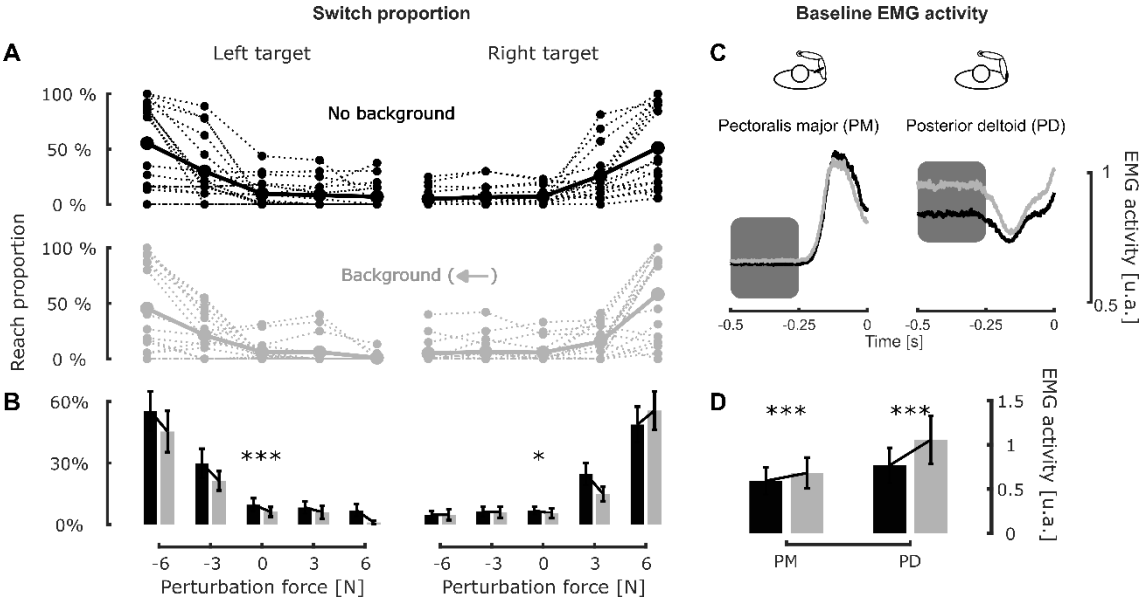

*Figure 6 : Experiment 3 – **A** Group mean (full line) and individual means (dashed lines) of the reach proportion to the left and right targets for the conditions without (top row, black) and with (bottom row, grey) the leftward background load as a function of perturbation load. **B** Comparison of the reach proportion of the left and right targets (left and right columns respectively) with (grey boxes) and without (black boxes) the leftward background force. **C** Group mean of the EMG activity in Pectoralis Major (left) and Posterior Deltoid (right) prior to movement onset for trials without (black) and with (grey) a background load. **D** Comparison of the binned EMG activity between 500 and 300ms (corresponds to the grey box in panel C) before force onset in Pectoralis Major and Posterior Deltoid for the conditions with and without background.*

The application of a leftward background force induced an increase in both PM and PD baseline EMG activity (Figure 6C). We found a significant effect of the background load in both muscles (main effect of the linear mixed model on both muscles: β_1=-0.11, t=-6.73, p<0.005, r^2=0.91) as represented in Figure 6D. Moreover, we also observed an interaction effect between the background load and the muscle: baseline activity in PD increased more than PM activity (β_{12}=0.20, t=19.067, p<0.005, r^2=0.91).

We found that the leftward background load modified the reach proportion to the left target for all kind of online mechanical loads. Figure 6A represents the reach proportions to the left and right targets (respectively left and right column) as a function of the intensity of the perturbation load for the trial with (bottom) or without (top) background. In order to show the effect of the background load on the reach proportion to the lateral targets, we fitted a multilinear logistic regression (see Methods) that inferred the effect of perturbation and background load on the reached target. This multilinear logistic regression revealed a significant effect of the background and perturbations loads on both the left and right targets reaching proportion. Concerning the perturbation load, we observed an increase of the reach proportion to the left target with increasing leftward force (β_1=-0.9223, t=-22.16, p<0.005) and the mirror effect for the right target (β_1=1.0204, t=23.86, p<0.005). The background load also had a significant effect on the reach proportion for these two targets. The reach proportion to the left target decreased when the background load was applied (β_1=-0.4611, t=-6.1759 and p<0.005, Figure 6B left panel). Intuitively an increase in force toward a target could bias the choice for that target but it was not the case. A slight decrease in reach proportion for the right target was also revealed by this regression (β_1=-0.1544, t=-1.9972, p=0.0458, Figure 6B right panel). The intensity of this effect on the two lateral targets was compared using bootstrap resampling: this effect was larger for the left than for the right target (one tailed t-test: t=92.08, p<0.005, d=-1.79).

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We generated 1000 bootstrap datasets from the original dataset used to fit the multilinear logistic regression and fitted the multilinear regression on each of these bootstrap datasets. We extracted bootstrap estimates of the main effect of background on the target reached for both lateral targets. Normality of the distributions of these bootstrap estimates was assessed using Kolmogorov-Smirnov tests ($p<0.005$ for both targets).

Post-hoc analyses performed on the individual reach proportion to lateral targets confirmed this asymmetry between left and right target (see Figure 6B). We observed a significant decrease of the individual reach proportion to the left target induced by the background load across participants and force levels (one tailed Wilcoxon signed-rank test : $z=2.83$, ranksum=999.5, $p<0.005$, $d=0.21$). No similar effect was observed for the right target (Wilcoxon signed-rank test: $z=1.23$, ranksum = 1015, $p=0.2154$, $d=0.03$).

An interesting question is whether this background force also modulated forward velocity. We address this question by using a linear mixed model to compare forward speed at force onset in the conditions with and without background load. No modulation of movement speed between these conditions was observed (linear mixed model β_1=-0.009±0.007, t=-1.2528, p=0.210, r^2=0.44). This result is important as it discards the eventuality that the modulation of flexibility to switch to a new target goal was induced by movement velocity.

The results of this last experiment showed that the tendency to switch observed in Experiment 2 depended on the biomechanical state of the limb. Importantly, the application of a background load in a direction reduces the tendency to switch in this direction in a larger amount than the tendency to switch in the opposite direction.

Discussion

We conducted three experiments to investigate whether humans favor movement vigor or flexibility for rapid motor decisions by leveraging the influence of explicit target reward and muscle activity. In Experiment 1, we demonstrated that target reward did not only increase movement vigor as reported in previous studies (Summerside et al., 2018; Yoon et al., 2018), but it also increased feedback gains and muscle activity. We observed that perturbation-related lateral hand deviations were smaller when they reached for a target associated with larger reward. Moreover, we also observed an increase in the baseline EMG activity as well as an increase in the EMG response to perturbation loads with increasing reward. Altogether, these results suggest that the feedback gains used to perform movements scaled with the value of the reward. We then uncovered that an increase in feedback gains, whatever the reason, was detrimental to the ability to switch target during movement. Indeed, we observed in Experiment 2 and 3 that participants were less likely to switch target during movement in the trials associated with a higher muscle activity. The modulation in muscle activity introduced experimentally in Experiment 3 induced a directional bias in the ability to switch target online, showing a causal influence of the state of the peripheral motor system on one's ability to flexibly switch movement goal.

The increase in movement vigor and feedback gains associated with reward that we observed in Experiment 1 was coherent with the use of a robust control strategy (Bian et al., 2020; Crevecoeur et al., 2019). A robust controller consists in an alternative to stochastic optimal control (Todorov & Jordan, 2002) that has the property to consider unmodelled disturbances (Basar & Bernhard, 1991) which results in better responses to mechanical perturbations during movements (Bian et al., 2020; Crevecoeur et al., 2019). Reward is known to invigorate movements as revealed in saccadic eye movements where faster movements were observed toward higher monetary rewards (Manohar et al., 2015, 2017) or toward targets associated with higher implicit rewards (Xu-Wilson et al., 2009). Similar observations were made for upper limb reaching movements that exhibited higher peak velocities toward more rewarding targets (Esteves et al., 2016; Sackaloo et al., 2014; Summerside et al., 2018; Yoon et al., 2018). This was taken as evidence for reward-dependent selection of

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

movement time (Haith et al., 2012; Shadmehr et al., 2010). Here we postulate that another mechanism is also at play: a higher reward produced a more robust strategy that revealed participants' will to render their movements less sensitive to perturbations, thereby reducing the risk to miss the goal. In this framework the reduction in movement time results from the robustness of the control that impacts movement velocity through larger goal-feedback gains. A parallel and complementary argument was recently developed in a study demonstrating that an increase in internal feedback gains is responsible for reward-related in endpoint accuracy (Manohar et al., 2019).

In this framework, the modulation of the robustness of control has a clear limitation that we were able to establish empirically: a robust control strategy is meant to reject disturbances indistinguishably, thus in principle it is clear that this strategy is not compatible with a flexible change in movement goal online, which requires a reduction in feedback response to let the perturbation redirect one's hand toward the new goal. The competition between robust control and flexible online decisions was clearly borne out in our results. Indeed, we observed in Experiment 2 a correlation between higher muscle activities during movement and the reduced probability of switching to a new goal.

Our results revealed a dynamic interaction between sensorimotor and cognitive factors, which influences not only target selection but also the way we respond to perturbations. Theoretical models of decision-making between different reach options suggest that the target selection occurs through a competition between the different options that integrate the motor costs incurred to each option (Cos et al., 2011; Morel et al., 2017; Shadmehr et al., 2016) as well as their respective outcome (Trommershäuser et al., 2003, 2008) through a distributed consensus (Cisek, 2007, 2012; Gold & Shadlen, 2007). Recently, a body of work investigated whether the factors that were shown to influence decision-making during planning could also influence target selection while the hand has already start moving. Biomechanical factors such as target positions or intensity of unexpected mechanical perturbations have been shown to influence such decisions (Kurtzer et al., 2020; Michalski et al., 2020; Nashed et al., 2014). More recently, a study revealed that cognitive factors such as reward distribution amongst the different options could also influence online decision-making (Cos et al., 2021; Marti-Marca et al., 2020). In Experiment 2, we revealed that both the reward distribution and the state of the peripheral motor systems had an effect on online decision-making confirming therefore that both sensorimotor and abstract cognitive factors influence target selection during a movement. Importantly, Experiment 3 revealed a causal link between higher muscle activities and the reduction in switching probability and additionally dismissed any causal influence of movement speed on this reduction in flexibility contrary to previous paradigms that largely modulated the movement speed (Wong & Haith, 2017).

Physiologically, the robust control strategy, characterized partly by higher overall muscle activity of the antagonist pair of shoulder flexor and extensor muscles, is known to recruit spinal circuits and increase the gain of stretch reflexes. It has been shown that the short-latency stretch response (20-45 ms) is affected by automatic gain scaling (Bedingham & Tatton, 1984; Marsden et al., 1976; Matthews, 1986; Pruszynski et al., 2009; Stein et al., 1995; Verrier, 1985). Automatic gain scaling is also known to affect long latency stretch responses (Pruszynski et al., 2009) while decreasing with time and had no effect on the voluntary epoch. The difference in muscle activity we observed between the switch and the no-switch trials can therefore induce a difference in the amplitudes of the short and long latency responses between both conditions. We did not measure any difference in the short-latency epochs but observed some reward-dependency in the long-latency ones that could support this hypothesis. Moreover, our results support the hypothesis that an increase in co-contraction does not only increase the intrinsic limb stiffness (Burdet et al., 2001; Hogan, 1984) but

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

that it can also modulate the robustness of the neural controller through feedback gains (Crevecoeur et al., 2019).

To conclude, our study highlights that multiple mechanisms underlie reward-dependent planning and control of movement. One the one hand, we suggest that there is a robust control strategy that involves peripheral circuits by means of increases in baseline activity and gain scaling of the feedback responses. This strategy associated with robust control is likely selected to reject perturbations and reduce the risk of missing the reward suggesting that there could be a cost incurred to reward. On the other hand, there exists a more flexible control strategy able to switch target during movement. It is conceivable that this second strategy, which requires some inhibition of muscle activity and response, be mediated by higher level inhibitory circuits and response modulation (Scott, 2016; Shadmehr & Krakauer, 2008). Both strategies take into account explicit target rewards and depend on the state of peripheral control loops.

Future work could study whether and how much individuals can modulate their strategy or whether the differences in strategy reflect individual traits. Indeed, individual differences have been shown in movement vigor (Reppert et al., 2018), and their possible effect on the modulation of feedback control is an exciting open question. Such ability is potentially central to understand planning and control in complex environments.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Basar, T., & Bernhard, P. (1991). H-infinity optimal control and related minimax design problems. *Birkhaüser, Boston, Massachussets*.

Bedingham, W., & Tatton, W. G. (1984). *Dependence of EMG Responses Evoked by Imposed Wrist Displacements on Pre-existing Activity in the Stretched Muscles*. *May*, 272-280.

Benjamin, D. J., & Berger, J. O. (2018). Redefine statistical significance. *Nature Human Behaviour*, *2*(1), 6-10. https://doi.org/10.1038/s41562-017-0189-z

Bian, T., Wolpert, D. M., & Jiang, Z. P. (2020). Model-free robust optimal feedback mechanisms of biological motor control. *Neural Computation*, *32*(3), 562-595. https://doi.org/10.1162/neco_a_01260

Burdet, E., Osu, R., Franklin, D. W., Milner, T. E., & Kawato. (2001). The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature*, *414*(6862), 446-449.

Cisek, P. (2007). Cortical mechanisms of action selection : The affordance competition hypothesis. *Philosophical Transactions of the Royal Society B*, *362*(April), 1585-1599. https://doi.org/10.1098/rstb.2007.2054

Cisek, P. (2012). Making decisions through a distributed consensus. *Current Opinion in Neurobiology*, *22*(6), 927-936. https://doi.org/10.1016/j.conb.2012.05.007

Cos, I., Bélanger, N., & Cisek, P. (2011). The influence of predicted arm biomechanics on decision making. *Journal of Neurophysiology*, *105*(6), 3022-3033. https://doi.org/10.1152/jn.00975.2010.

Cos, I., Pezzulo, G., & Cisek, P. (2021). Changes of mind after movement onset : A motor-state dependent decision-making process. *bioRxiv*, 6.

Crevecoeur, F., Scott, S. H., & Cluff, T. (2019). Robust Control in Human Reaching Movements : A Model-Free Strategy to Compensate for Unpredictable Disturbances. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *39*(41), 8135-8148. https://doi.org/10.1523/JNEUROSCI.0770-19.2019

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Esteves, P. O., Oliveira, L. A. S., Nogueira-Campos, A. A., Saunier, G., Pozzo, T., Oliveira, J. M., Rodrigues, E. C., Volchan, E., & Vargas, C. D. (2016). Motor planning of goal-directed action is tuned by the emotional valence of the stimulus : A kinematic study. *Scientific Reports*, *6*(March), 1-7. https://doi.org/10.1038/srep28780

Gold, J. I., & Shadlen, M. N. (2007). The Neural Basis of Decision Making. *Annual Review of Neuroscience*, *30*, 535-574. https://doi.org/10.1146/annurev.neuro.29.051605.113038

Haith, A. M., Reppert, T. R., & Shadmehr, R. (2012). *Evidence for Hyperbolic Temporal Discounting of Reward in Control of Movements*. *32*(34), 11727-11736. https://doi.org/10.1523/JNEUROSCI.0424-12.2012

Hogan, N. (1984). Adaptive control of mechanical impedance by coactivation of antagonist muscles. *IEEE Transactions on Automatic Control*, *29*(8), 681-690.

Kurtzer, I., Muraoka, T., Singh, T., Prasad, M., Chauhan, R., & Adhami, E. (2020). Reaching movements are automatically redirected to nearby options during target split. *Journal of Neurophysiology*, *124*(4), 10313-11028.

Lowrey, C. R., Nashed, J. Y., & Scott, S. H. (2017). Rapid and flexible whole body postural responses are evoked from perturbations to the upper limb during goal-directed reaching. *Journal of Neurophysiology*, *117*(3), 1070-1083. https://doi.org/10.1152/jn.01004.2015

Manohar, S. G., Chong, T. T. J., Apps, M. A. J., Batla, A., Stamelou, M., Jarman, P. R., Bhatia, K. P., & Husain, M. (2015). Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control. *Current Biology*, *25*(13), 1707-1716. https://doi.org/10.1016/j.cub.2015.05.038

Manohar, S. G., Finzi, R. D., Drew, D., & Husain, M. (2017). Distinct Motivational Effects of Contingent and Noncontingent Rewards. *Psychological Science*, *28*(7), 1016-1026. https://doi.org/10.1177/0956797617693326

Manohar, S. G., Muhammed, K., Fallon, S. J., & Husain, M. (2019). Motivation dynamically increases noise resistance by internal feedback during movement. *Neuropsychologica*, *123*, 19-29.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Marsden, C. D., Merton, P. A., & Morton, H. B. (1976). SERVO ACTION IN THE HUMAN THUMB. *Journal of Physiology*, *257*(1), 1-44.

Marti-Marca, A., Deco, G., & Cos, I. (2020). Visual-reward driven changes of movement during action execution. *Scientific Reports*, *10*, 15527.

Matthews, P. B. (1986). Observations on the automatic compensation of reflex gain on varying the pre-existing level of motor discharge in man. *Journal of Physiology*, *374*(May), 73-90.

Michalski, J., Green, A. M., & Cisek, P. (2020). Reaching decisions during ongoing movements. *Journal of Neurophysiology*, *123*(3), 1090-1102. https://doi.org/10.1152/jn.00613.2019

Morel, P., Ulbrich, P., & Gail, A. (2017). What makes a reach movement effortful ? Physical effort discounting supports common minimization principles in decision making and motor control. *PLOS Biology*, *15*(6), 1-23.

Nashed, J. Y., Crevecoeur, F., & Scott, S. H. (2014). Rapid Online Selection between Multiple Motor Plans. *Journal of Neuroscience*, *34*(5), 1769-1780. https://doi.org/10.1523/JNEUROSCI.3063-13.2014

Pruszynski, J. A., Kurtzer, I., Lillicrap, T. P., & Scott, S. H. (2009). Temporal Evolution of " Automatic Gain-Scaling ". *Journal of Neurophysiology*, *102*(2), 992-1003. https://doi.org/10.1152/jn.00085.2009.

Reppert, T. R., Rigas, I., Herzfeld, D. J., Sedaghat-nejad, E., Komogortsev, O., & Shadmehr, R. (2018). Movement vigor as a traitlike attribute of individuality. *Journal of Neurophysiology*, *120*(2), 741-757. https://doi.org/10.1152/jn.00033.2018

Sackaloo, K., Otr, L., Strouse, E., Otr, L., Rice, M. S., & Otr, L. (2014). Degree of Preference and Its Influence on Motor Control When Reaching for Most Preferred , Neutrally Preferred , and Least Preferred Candy. *OTJR: Occupation, Participation and Health*, *35*(2), 81-88. https://doi.org/10.1177/1539449214561763

Scott, S. H. (2016). A Functional Taxonomy of Bottom-Up Sensory Feedback Processing for Motor Actions. *Trends in Neurosciences*, *39*(8), 512-526. https://doi.org/10.1016/j.tins.2016.06.001

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Shadmehr, R., Huang, H. J., & Ahmed, A. A. (2016). A Representation of Effort in Decision-Making and Motor Control. *Experimental Brain Research*, *26*(14), 1929-1934. https://doi.org/10.1016/j.cub.2016.05.065

Shadmehr, R., Orban de Xivry, J-J., Xu-Wilson, M., & Shih, T. (2010). Temporal Discounting of Reward and the Cost of Time in Motor Control. *Journal of Neuroscience*, *30*(31), 10507-10516. https://doi.org/10.1523/JNEUROSCI.1343-10.2010

Shadmehr, R., & Krakauer, J. W. (2008). A computational neuroanatomy for motor control. *Experimental Brain Research*, *185*(3), 359-381. https://doi.org/10.1007/s00221-008-1280-5

Stein, R. B., Hunter, I. W., Lafontaine, S. R., & Jones, L. A. (1995). Analysis of short-latency reflexes in human elbow flexor muscles. *Journal of Neurophysiology*, *73*, 1900-1911.

Summerside, E. M., Shadmehr, R., & Ahmed, A. A. (2018). Control of Movement Vigor of reaching movements : Reward discounts the cost of effort. *Journal of Neurophysiology*, *119*(6), 2347-2357. https://doi.org/10.1152/jn.00872.2017

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, *5*(11), 1226-1235. https://doi.org/10.1038/nn963

Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2003). Statistical decision theory and trade-offs in the control of motor response. *Spatial Vision*, *16*(3-4), 255-275.

Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2008). Decision making , movement planning and statistical decision theory. *Trends in Cognitive Sciences*, *12*(8), 291-297. https://doi.org/10.1016/j.tics.2008.04.010

Verrier, M. C. (1985). Alterations in H reflex magnitude by variations in baseline EMG excitability. *Electroencephalogram Clin Neurophysiology*, *60*, 492-499.

Wong, A. L., & Haith, A. M. (2017). Motor planning flexibly optimizes performance under uncertainty about task goals. *Nature Communications*, *8*, 1DUMMY. https://doi.org/10.1038/ncomms14624

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.25.453678; this version posted July 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Xu-Wilson, M., Zee, D. S., & Shadmehr, R. (2009). The intrinsic value of visual information affects saccade velocities. *Experimental Brain Research*, *196*(4), 475-481. https://doi.org/10.1007/s00221-009-1879-1

Yoon, T., Geary, R. B., Ahmed, A. A., & Shadmehr, R. (2018). Control of movement vigor and decision making during foraging. *PNAS*, *115*(44), E10476-E10485. https://doi.org/10.1073/pnas.1812979115