Sensorimotor feedback loops are selectively sensitive to reward

Olivier Codol[1,2,3,6], Christopher J. Forgaard[1,2], Joseph M. Galea[3], Paul L. Gribble[1,2,4,5]

[1] Brain and Mind Institute, University of Western Ontario, Ontario, Canada

[2] Department of Psychology, University of Western Ontario, Ontario, Canada

[3] School of Psychology, University of Birmingham, United Kingdom

[4] Department of Physiology & Pharmacology, Schulich School of Medicine & Dentistry, University of Western Ontario, Ontario, Canada

[5] Haskins Laboratories, New Haven CT, USA

[6] Correspondence should be addressed to Olivier Codol at codol.olivier@gmail.com.

Abstract

While it is well established that motivational factors such as earning more money for performing well improve motor performance, how the motor system implements this improvement remains unclear. For instance, feedback-based control, which uses sensory feedback from the body to correct for errors in movement, improves with greater reward. But feedback control encompasses many feedback loops with diverse characteristics such as the brain regions involved and their response time. Which specific loops drive these performance improvements with reward is unknown, even though their diversity makes it unlikely that they are contributing uniformly. This lack of mechanistic insight leads to practical limitations in applications using reward, such as clinical rehabilitation, athletic coaching, and brain-inspired robotics. We systematically tested the effect of reward on the latency (how long for a corrective response to arise?) and gain (how large is the corrective response?) of eight distinct sensorimotor feedback loops in humans. Only the feedback responses known to rely on prefrontal associative cortices showed sensitivity to reward, while feedback responses that relied mainly on premotor and sensorimotor cortex did not show sensitivity to reward. Our results may have implications regarding feedback control performance in pathologies showing a cognitive decline, or on athletic coaching. For instance, coaching methodologies that rely on reinforcement or “reward shaping” may need to specifically target aspects of movement that rely on reward-sensitive feedback responses.

Introduction

If a cat pushes your hand whilst you are pouring a glass of water, a corrective response will occur that acts to minimise water spillage. This simple action is an example of a behavioural response triggered by sensing a relevant change in the environment—here, a push that perturbs the movement of your arm away from the intended movement path. This form of feedback control requires the brain to integrate sensory information from the periphery of the body, and thus suffers from transmission delays inherent in the nervous system. There is evidence that when more is at stake, we react faster to respond to demands of the task (Reddi and Carpenter, 2000). For instance, if wine was being poured instead of water and your favourite, white tablecloth covers the table, you may be faster at correcting for a perturbation that risks spilling your wine and staining the tablecloth.

In the context of human motor control, feedback-based control is not a monolithic process (Reschechtko and Pruszynski, 2020; Scott, 2016). Rather, the term encompasses a series of sensorimotor feedback loops that rely on different sensory information, are constrained by different transmission delays (Figure 1a), and are supported by different neural substrates (Figure 1b; Reschechtko and Pruszynski, 2020). The circuitry underlying the short-latency stretch reflex is entirely contained in the spinal cord (Sherrington, 1906, 1913). The long-latency reflex relies on supraspinal regions such as the primary motor and primary sensory cortices (Pruszynski et al., 2011a), and is modulated by upstream associative cortical regions (Figure 1b; Beckley et al., 1991; de Graaf et al., 2009; Omrani et al., 2016; Scott, 2012; Zonnino et al., 2021). Visuomotor feedback responses rely on visual cortex and other cortical and subcortical brain regions (Day and Brown, 2001; Desmurget et al., 2004). Due to these differences, each feedback loop is governed by different objectives such as maintenance of a limb position or reaching toward a goal (Figure 1a). Therefore, to address whether sensorimotor feedback is sensitive to motivational factors requires testing multiple distinct feedback responses. In this work, we employ rewarding outcomes (specifically, monetary reward) as a means to manipulate motivation (Codol et al., 2020a; Galea et al., 2015; Goodman et al., 2014; Hübner and Schlösser, 2010; McDougle et al., 2021).

Recent work has demonstrated that rewarding outcomes improve motor performance in many ways. Reward results in changes to the speed-accuracy trade-off, a hallmark of skilled performance (Codol et al., 2020a, 2020b; Manohar et al., 2015, 2019). It can lead to a reduction in noise in the central nervous system and at the effector to improve the control of movement (Codol et al., 2020a; Goard and Dan, 2009; Manohar et al., 2015; Pinto et al., 2013). But whether reward modulates sensorimotor feedback control specifically remains scarcely tested, although previous work in saccadic eye movements (Manohar et al., 2019) and indirect evidence in reaching (Codol et al., 2020a) suggests this may be the case. More recent studies outline a general sensitivity of feedback control to reward during reaching but does not differentiate between each distinct feedback loop that the nervous system relies on to implement this control (Carroll et al., 2019; De Comité et al., 2021; Poscente et al., 2021). However, the information to which each loop is tuned greatly varies (Reschechtko and Pruszynski, 2020; Scott, 2016), and consequently it is unlikely that they are all uniformly impacted by reward.

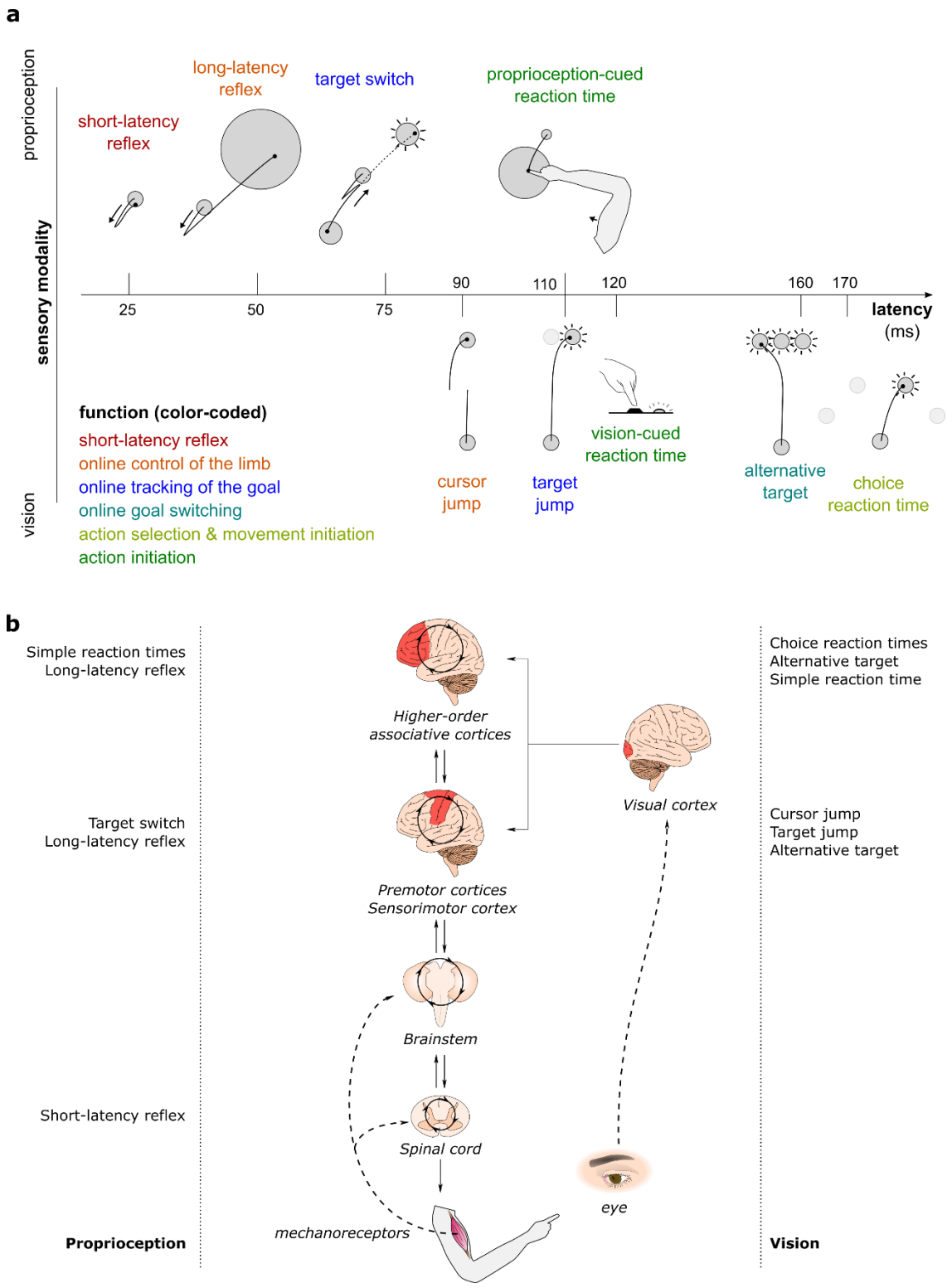

Figure 1: Sensorimotor feedback loops involved in motor control. *(a) Feedback loops can be classified along three dimensions: the sensory modality on which they rely (vertical axis), the post-perturbation latency of their response (horizontal axis), and the function they perform (color-coded). Figure adapted from* (Scott, 2016). *Note that latencies indicated here reflect the*

fastest reported values from the literature and not necessarily what was observed in this study. (b) Each feedback response relies on different anatomical substrates, which form a nested network able to integrate information to produce movement.

Systematically assessing each individual feedback loop's sensitivity to reward would provide a mechanistic, rather than descriptive, understanding of this relationship. Additionally, it would further complete the picture of how rewarding information improves motor performance (Figure 1b), a point of focus in the past decade due to applications in clinical rehabilitation (Goodman et al., 2014; Quattrocchi et al., 2017), sports coaching (Cashaback et al., 2019; Galea et al., 2015; Manley et al., 2014; Parvin et al., 2018), or studies on motor pathologies (Izawa and Shadmehr, 2011; Manohar et al., 2015, 2019; Pekny et al., 2015; Therrien et al., 2016, 2018). Teasing apart which feedback loops are modulated by reward can also provide a window into the neural substrate of reward processing during the production of movement, since they rely on different brain regions (Omrani et al., 2016; Reschechtko and Pruszynski, 2020).

Therefore, in the present study we tested how eight distinct sensorimotor feedback responses are modulated by reward. We measured feedback latency (how long does it take for a corrective response to arise?) and feedback gain (how large is the corrective response?) for each feedback response within rewarded and unrewarded conditions. Motivational factors can take different forms, such as rewarding or punishing outcomes (Chen et al., 2017, 2018a, 2018b; Codol et al., 2020a; Galea et al., 2015; Guitart-Masip et al., 2014), inhibition versus mouvement (Chen et al., 2018a; Guitart-Masip et al., 2014), contingency (Manohar et al., 2017), expectation (Lowet et al., 2020; Schultz et al., 1997), urgency (Poscente et al., 2021), or agency (Parvin et al., 2018). In this study we focused on contingent rewarding outcomes, where participants have agency over the returns they obtain, and with an expectation component since potential for returns is indicated at the start of each trial (see Results and Methods).

Results

We first assessed feedback gain and latency for the short- and long-latency stretch reflex (SLR and LLR), which are the fastest feedback responses observed in human limb motor control. Participants were seated in front of a robotic device that supported their arm against gravity and allowed for movement in a horizontal plane. They positioned their index fingertip at a starting position while countering a +2 Nm background load (dashed arrows in Figure 2a, top panel) to activate the elbow and shoulder extensor muscles. We recorded electromyographic signals (EMG) using surface electrodes placed over brachioradialis, triceps lateralis, pectoralis major (clavicular head), posterior deltoid, and biceps brachii (short head). After participants held their hand in the starting position for 150-200 ms, a 10 cm target appeared at 20 degrees either inward (closer to the chest) or outward (away from the chest) with respect to the elbow joint. Next, a ±2 Nm torque perturbation was generated by the robot about the elbow and shoulder joints (solid arrows in Figure 2a-c). A positive or negative torque signifies an inward or

an outward perturbation from the starting position, respectively. Participants were instructed to move their fingertip into the target as soon as the perturbation occurred.

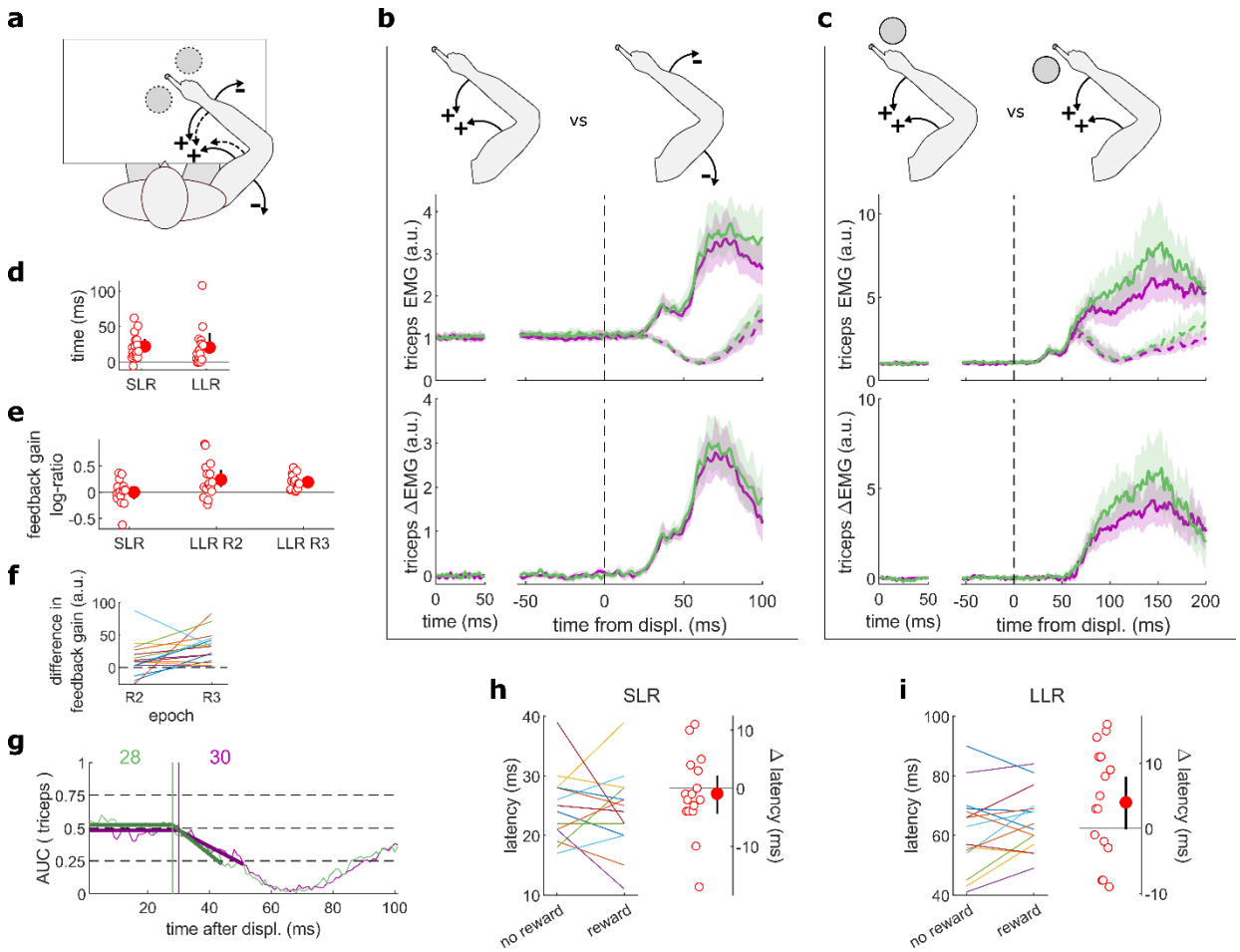

Figure 2: results for the SLR and LLR contrasts. *(a) top view schematic of the apparatus. Participants could be moved in a horizontal plane. Background forces were applied to pre-activate the extensor muscles (dashed arrows) and a mechanical perturbation, either two positive or negative torques, was applied to elicit the SLR and LLR (solid arrows). The dashed circles indicate the two possible target positions. (b) top panel: contrast used to observe the SLR. The control and manipulation conditions were on the right and left, respectively. Background loads are not drawn here for clarity. middle panels: average triceps EMG signal from the SLR contrast, with the dashed and solid lines representing control and manipulation conditions, respectively; bottom panels: difference between the manipulation and control condition. The left panels show EMG at trial baseline (see methods). Green and purple lines indicate EMGs for rewarded and non-rewarded trials, respectively. Shaded areas indicate 95% CIs. (c) same as (b) but for the LLR contrast. (d) Difference in time spent in target between rewarded and non-rewarded trials. A larger value indicates more time spent in the target for rewarded trials. (e) feedback gains following SLR and LLR response onset, expressed as a log-ratio of rewarded over non-rewarded feedback gain. R2 and R3 epochs were taken as the first 25 ms and the following*

25 ms after each participant's LLR latency, respectively. Larger values correspond to higher gain in rewarded trials compared to non-rewarded trials. (f) Difference in LLR feedback gain between reward and no-reward conditions for each participant, in the R2 and R3 epochs. (g) Example area under curve (AUC) to obtain response latency for one participant and for the SLR contrast. Green and purple lines indicate rewarded and non-rewarded trials, respectively. Thick lines indicate line-of-best-fit for a two-step regression (see methods). (h) left panel: absolute response latency after the perturbation for the SLR; right panel: difference between rewarded and non-rewarded trials. (i) same as (h) but for the LLR. In all panels with a red filled dot and black error bars, the filled dot indicates the group mean and error bars indicate 95% CIs.

This yielded a 2x2 factorial design, in which an inward or outward perturbation is associated with an inward or outward target. Different contrasts allowed us to assess the SLR and LLR within the same task (Figure 2b-c). For extensor muscles, the SLR arose in trials with an inward perturbation but not in trials with an outward perturbation, regardless of the target displayed (Figure 2b). Feedback latency and gain were assessed by measuring when the EMG signal diverges between those two conditions, and the magnitude of the EMG signal in the 25 ms time period following this divergence, respectively. Conversely, the LLR arose if the direction of the perturbation conflicted with the task goal. For example, a LLR was observed in extensor muscles when contrasting an inward perturbation with an outward target (LLR occurring) but not for an inward perturbation with an inward target (LLR not occurring, Figure 2c).

Finally, we introduced monetary reward as a third factor in our task design to assess its impact on feedback responses. Rewarded and non-rewarded trials were indicated at the beginning of each trial by displaying "$$$" and "000" symbols, respectively, on the display screen in front of participants. These symbols were replaced by each trial's actual monetary value once the target was reached (always 0 ¢, CAD, in the case of non-rewarded trials). For rewarded trials, the monetary gains were proportional to the time spent inside the end target, therefore promoting faster reaches (see methods) because trial duration was fixed.

<u>Reward altered feedback responses as early as 50 ms post-perturbation</u>

In all conditions with rewarded outcomes, behavioral performance, measured as time in target, improved compared to conditions without reward (SLR contrast: W=135, r=0.99, p=5.31e-4; LLR contrast: W=130, r=0.95, p=1.34e-3, Figure 2d). We quantified feedback gains as a log-ratio of integrals of EMG activity between rewarded and non-rewarded conditions (see Methods), meaning a positive number indicates an increase in feedback gain for the rewarded conditions. We observed no difference at baseline and in a 25 ms window after the onset of the SLR (W=74, r=0.54, p=0.75, Figure 2e), that is in the R1 window of the SLR (Pruszynski et al., 2008). Next, we assessed the time of divergence of each participants' EMG activity between the reward and no-reward conditions using a Receiver Operating Characteristic (ROC) signal discrimination method (Figure 2g). We performed this analysis on the rewarded and non-rewarded trials separately, yielding two latencies per participant. Latencies for the SLR in the non-rewarded conditions

were in the 25 ms range post-perturbation, and latencies in the rewarded conditions were in a similar range as well, with no significant difference observed (W=68.5, r=0.50, p=0.64, Figure 2h). Therefore, rewarding outcomes affected neither feedback latency nor feedback gains of the SLR.

Turning to the LLR feedback response, we observed that feedback gains were greater in the rewarded condition in a 25 ms window following LLR onset, which corresponds to the R2 epoch (W=114, r=0.84, p=0.017, Figure 2e). We also assessed feedback gains in the R3 epoch, which is the 25 ms window following the R2 epoch, because slower processes that contribute to the LLR also occur in that timeframe (Lee and Tatton, 1975; Pruszynski et al., 2011b; Tatton et al., 1975). In the R3 epoch, we also observed an effect of reward on feedback gains (W=136, r=1, p=4.37e-4, Figure 2e). To test whether the effect of reward is different in the R2 and R3 epoch, we assessed the interaction between reward and epoch in a repeated-measure analysis of variance (rm-ANOVA). For this analysis, we did not use a log-ratio, because EMG activity was not similar in R2 and R3 epochs (Figure 2c), and that difference would lead to a mismatched ratio normalisation across epoch, hindering comparisons over epochs. The rm-ANOVA revealed a small but reliable interaction, indicating that reward increased feedback gains more in the R3 epoch than it did in the R2 epoch (Figure 2f). Finally, ROC analysis showed that LLR latencies were similar in the rewarded condition compared to the non-rewarded condition (W=102.5, r=0.75, p=0.074, Figure 2i).

In summary, while the prospect of reward did not alter the SLR, it led to increases in feedback gains as early as the LLR, that is, about 50 ms post-perturbation, which is much earlier than the increase in latencies with reward reported in previous work (Carroll et al., 2019; De Comité et al., 2021). This increase in LLR gains with reward was stronger in the R3 epoch than in the R2 epoch.

Slower, goal-tracking feedback responses were unaltered by reward

In addition to the SLR and LLR, slower feedback responses also exist that control for higher level aspects of movement, such as tracking the task goal. We tested the effect of reward on goal-tracking feedback responses in two different settings. In a first “target switch” task, participants were instructed to adjust their movements to switch to a second target when a mechanical perturbation occurred during the movement (Figure 3a). In a second task, the switch to a second target (goal) was signalled by a visual jump of the target about halfway through the movement (Figure 3g). This allowed us to assess goal-tracking feedback responses relying on either proprioceptive or visual information, which are distinct and display their own characteristics. Notably, the proprioceptive feedback response displays faster latencies than the visual target jump response. Note that the target jump task was performed on an endpoint KINARM robot (Figure 4a, see methods).

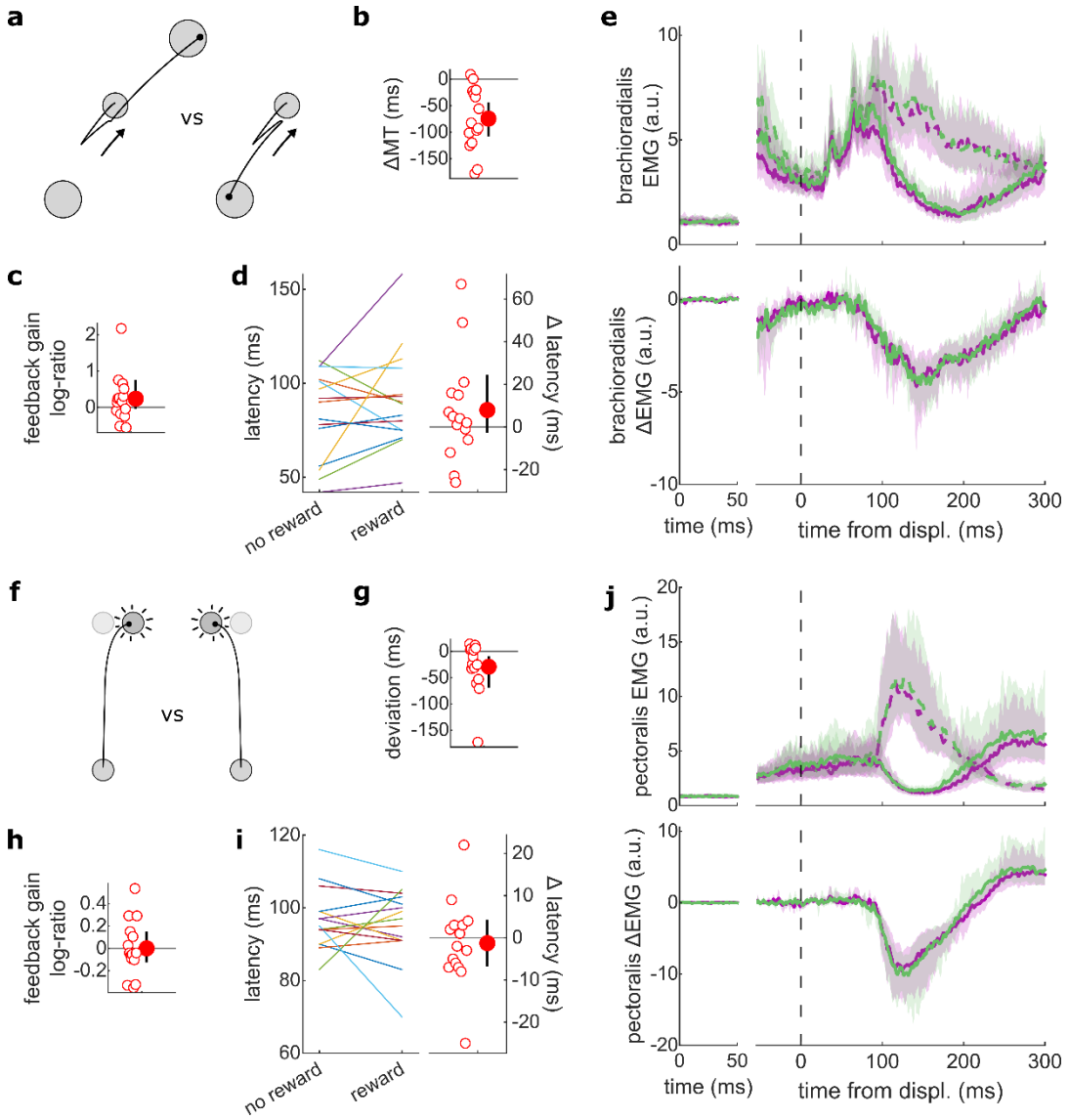

Figure 3: results for the goal-sensitive feedback responses. *(a) contrast used to observe the target switch response. The control and manipulation conditions were on the right and left, respectively. Background loads are not drawn here for clarity. The solid arrow indicates the mechanical perturbation. When two targets are displayed, the perturbation signals the need to switch to the second (top) target. (b) Difference in movement times (ΔMT) between rewarded and non-rewarded trials. A more negative value indicates shorter movement times for rewarded trials. (c) feedback gains following response onset, expressed as a log-ratio of rewarded over non-rewarded feedback gain. Larger values mean more gain in rewarded trials. (d) left panel: absolute response latency after the perturbation for the target switch response; right panel: difference between rewarded and non-rewarded trials. (e) top panels: average brachioradialis EMG signal, with the dashed and solid lines representing control and manipulation conditions, respectively; bottom panels: difference between the manipulation and control condition. The left panels show EMG at trial baseline (see methods). Green and purple lines indicate EMGs for rewarded and non-rewarded trials, respectively. Shaded areas indicate 95% CIs. (f) contrast used*

to observe the target jump response. (g) difference in deviation from the desired movement time between rewarded and non-rewarded trials. A more negative value indicates movement times closer to the desired value for rewarded trials. (h-j) same as (c-e) for the target jump task. In all panels with a red filled dot and black error bars, the filled dot indicates the group mean and the error bars indicate 95% CIs.

Participants were rewarded for short movement times and small deviations from desired movement time in the target switch and target jump task, respectively. Performance improved in the rewarding condition in both tasks (target switch: W=117, r=0.97, p=3.05e-4, Figure 3b; target jump: W=97, r=0.80, p=0.035, Figure 3h), but this was not due to an increase in feedback gains (target switch: W=82, r=0.68, p=0.23, Figure 3c; target jump: W=66, r=0.55, p=0.76, Figure 3i) nor to a change in feedback latency (target switch: W=79.5, r=0.66, p=0.28, Figure 3d-e; target jump: W=72.5, r=0.60, p=0.49, Figure 3j-k).

Therefore, our results indicate expectation of reward did not affect the response of feedback responses that adjust movements for task goals, regardless of which sensory modality it relies on (proprioceptive or visual). While superficially, this stands in contrast to the improvement in behavioural performance observed in these tasks, this merely suggests that other aspects of movement could have shown improvement, including other feedback responses that may not be apparent with the task contrast used here.

<u>Online visual control of limb position was also unaltered by reward</u>

Next, we assessed feedback response due to a cursor jump rather than a target jump. This feedback response is sensitive to position of the limb like the LLR, but it displays longer latencies (Dimitriou et al., 2013). The task design was identical to that of the target jump task, except that the cursor, rather than the target, visually jumped sideways (Figure 4a-c). Again, behavioural performance improved in the rewarded conditions, as measured by deviation from desired movement time (W=116, r=0.96, p=4.27e-4, Figure 4d), but feedback latencies and feedback gains were similar with or without reward (gains: W=79, r=0.66, p=0.29, Figure 4e; latencies: W=75.5, r=0.63, p=0.15, Figure 4f-g).

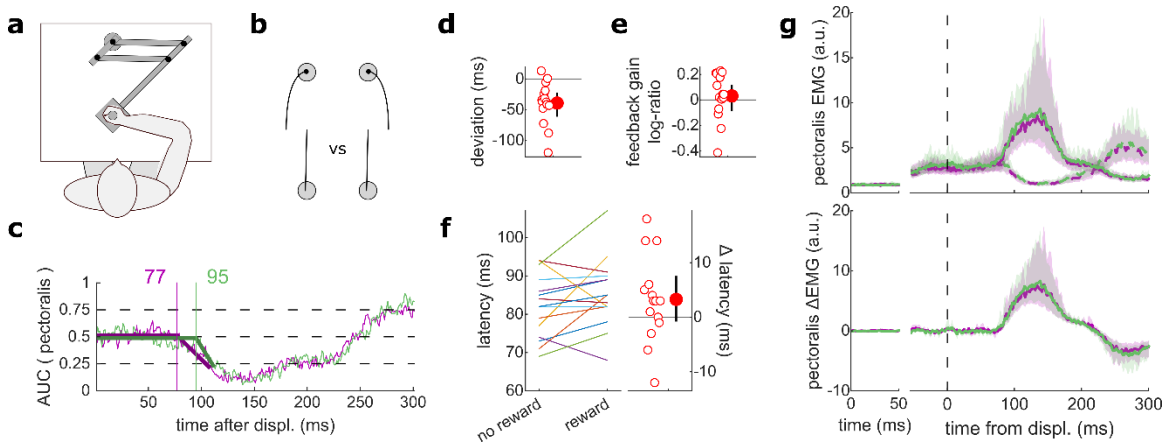

Figure 4: results for the cursor jump task. *(a) top view schematic of the apparatus. (b) contrast used to observe the cursor jump response. The discontinuity in the trajectory (solid black line) indicates the cursor jump. (c) Example area under curve (AUC) to obtain response latency for one participant. Green and purple lines indicate rewarded and non-rewarded trials, respectively. The thick lines indicate line-of-best-fit for a two-step regression (see methods). (d) difference in deviation from the desired movement time between rewarded and non-rewarded trials. A more negative value indicates movement times closer to the desired value for rewarded trials. (e) feedback gains following response onset, expressed as a log-ratio of rewarded over non-rewarded feedback gain. Larger values mean more gain in rewarded trials. (f) left panel: absolute response latency after the perturbation; right panel: difference between rewarded and non-rewarded trials. (g) top panels: average pectoralis EMG signal, with the dashed and solid lines representing control and manipulation conditions, respectively; bottom panels: difference between the manipulation and control condition. The left panels show EMG at trial baseline (see methods). Green and purple lines indicate EMGs for rewarded and non-rewarded trials, respectively. Shaded areas indicate 95% CIs. In all panels with a red filled dot and black error bars, the filled dot indicates the group mean and the error bars indicate 95% CIs.*

<u>Reaction times improve with reward</u>

Reaction times have been measured in many different settings that include rewarding feedback (Douglas and Parry, 1983; Steverson et al., 2019; Stillings et al., 1968). The consensus is that reaction times are reduced when reward is available. However, previous work always considered reaction times triggered by non-proprioceptive cues, such as auditory (Douglas and Parry, 1983) or visual cues (Stillings et al., 1968). Here, we assessed participants' reaction times triggered by a proprioceptive cue, which for arm movement tasks produce faster response latencies than visual cues (Pruszynski et al., 2008).

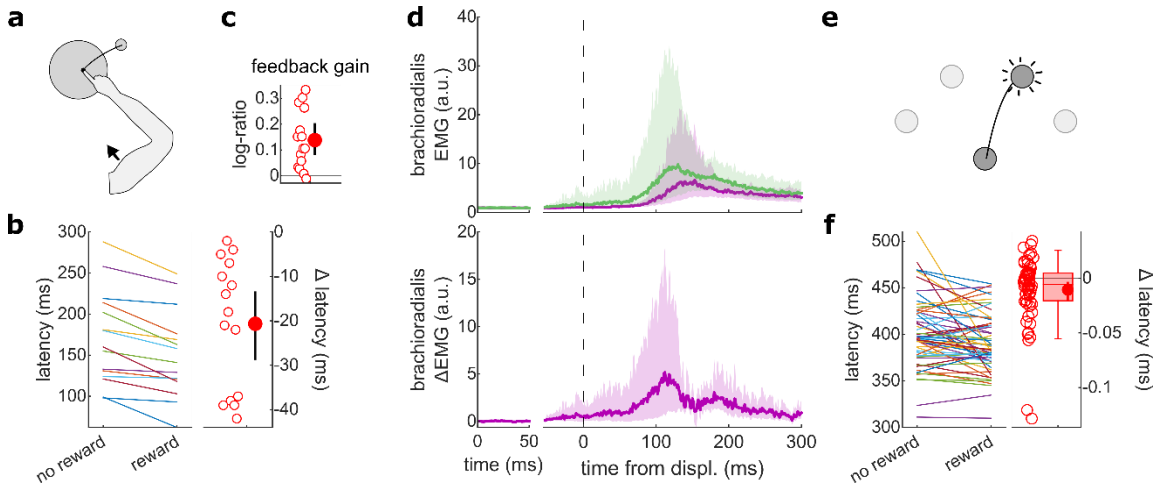

Figure 5: results for the reaction time tasks. *(a) schematic of task design for proprioception-cued reaction times. Participants were informed to initiate an elbow flexion by a small mechanical perturbation at the shoulder (black arrow). Background loads are not drawn here for clarity. (b) left panel: absolute response latency after the perturbation; right panel: difference between rewarded and non-rewarded trials. (c) feedback gains following response onset, expressed as a log-ratio of rewarded over non-rewarded feedback gain. Larger values mean more gain in rewarded trials. (d) top panels: average brachioradialis EMG signal, with the dashed and solid lines representing control and manipulation conditions, respectively; bottom panels: difference between the manipulation and control condition. The left panels show EMG at trial baseline (see methods). Green and purple lines indicate EMGs for rewarded and non-rewarded trials, respectively. Shaded areas indicate 95% CIs. (e) schematic of task design for vision-cued task design. (f) left panel: absolute vision-cued reaction times; right panel: difference between rewarded and non-rewarded trials. The boxplot indicates the 5-25-50-75-95 percentiles. In all panels with a red filled dot and error bars, the filled dot indicates the group mean and the error bars indicate 95% CIs.*

Participants held their arm so that the tip of the index finger was positioned at a starting location and the arm was stabilized against background loads that pre-loaded the forearm and upper arm flexor muscles (Figure 5a). A go cue was provided in the form of a very small extension perturbation at the shoulder, that led to less than 1 degree of shoulder or elbow rotation (Pruszynski et al., 2008). Participants were instructed to perform a fast elbow flexion toward a 10 cm target when they detected the go cue. Reaction times were defined based on brachioradialis EMG rising 5 standard deviations above baseline level (Pruszynski et al., 2008). In line with the literature on reaction times triggered by other sensory modalities, proprioception-triggered reaction times were greatly reduced under reward, reducing on average by 20.6 ms, from 170.8 to 150.2 ms (W=120, r=1, p=6.10e-5, Figure 5b). Feedback gains also increased significantly (W=118, r=0.98, p=1.83e-4, Figure 5c-d).

Finally, we assessed reaction times in a choice reaction time task by re-analysing a dataset available online (Codol et al., 2020a). In this dataset, participants (N=60) reached to one of four

targets displayed in front of them in an arc centred on the starting position (Figure 5f). Participants could obtain monetary reward for initiating their movements quicker once the target appeared (reaction times) and for reaching faster to the target (movement times). In line with the current study, reaction times were shorter in the rewarded than in non-rewarded condition, from 400.8 to 390.2 ms on average (W=1241, r=0.67, p=0.016, Figure 5g). Of note, because EMG recordings were not available for the online dataset, only kinematic data were available, which explains the slower absolute reaction times than reported in other studies (Haith et al., 2015; Summerside et al., 2018).

Discussion

In this study we tested whether reward affected eight different kinds of sensorimotor feedback responses. Results indicate a heterogeneous sensitivity, both in terms of which feedback response and which characteristics of the response were modulated by reward (Figure 6). The earliest effect was observed during the R2 epoch of the LLR, that is about 50 ms post-perturbation. This effect was constrained to the gain of the feedback response and did not extend to its latency. Following this, feedback responses slower than 110 ms were affected by reward (Figure 6a), with both an increase in gain and a reduction in response latency with reward. The SLR and two goal-tracking feedback responses (target switch and visual target jump) showed no modulation by reward.

A feedback loop that we could not assess is the cortico-cerebellar feedback loop (Becker and Person, 2019; Chen-Harris et al., 2008; Manohar et al., 2019). This loop is the only feedback loop contributing to saccadic eye movements (Chen-Harris et al., 2008), which show performance improvements under reward as well (Manohar et al., 2015, 2019). Electrophysiological evidence in mice (Becker and Person, 2019) and non-invasive manipulation in humans (Miall et al., 2007) suggest this loop also contributes to reaching movement, but behavioural assessment remains challenging.

While SLR circuitry is contained within the spinal cord, it does receive supraspinal modulation and displays functional sensitivity to higher-order task goals as well (Crone and Nielsen, 1994; Nielsen and Kagamihara, 1992, 1993; Weiler et al., 2019). Therefore, while unlikely, central modulation of SLR circuitry for rewarding outcomes could not be *a priori* ruled out. However, we observed no such modulation in the present experiments.

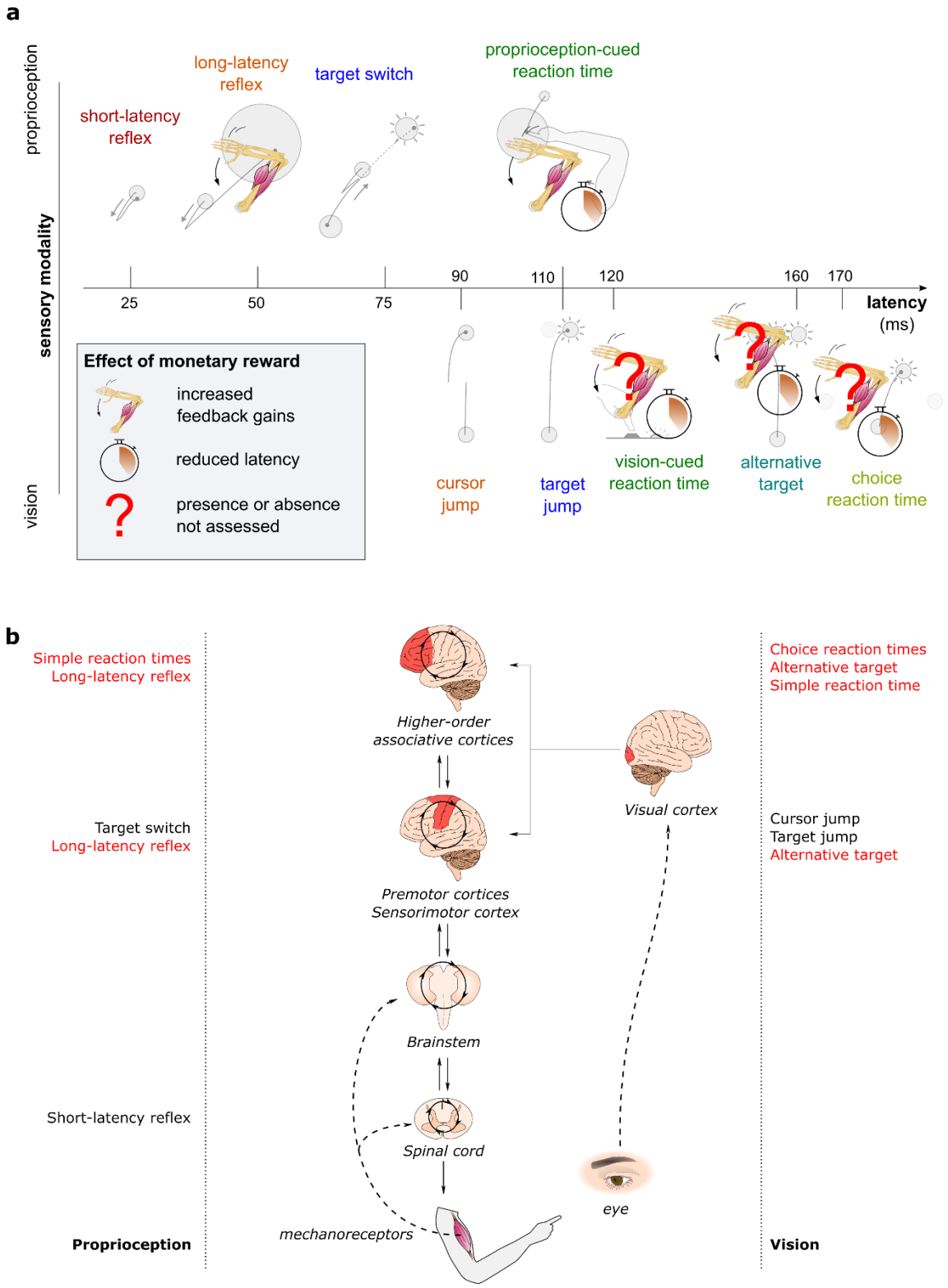

Figure 6: overview of rewarding feedback's impact on sensorimotor feedback responses. *(a) Reward can impact a feedback loop response by increasing feedback gains or reducing latency.*

The colour code indicates function and is identical to the one in Figure 1a. (b) Reward exclusively affected feedback responses that rely on cortical associative areas or on the basal ganglia.

Our results indicate that goal-tracking feedback responses are not affected by reward. This is surprising because task goal is often considered as a form of intrinsic reward (McDougle et al., 2021). By extension, one may propose that extrinsic rewards like monetary return would increase the rewarding value of the task goal and make it more salient (McDougle et al., 2021). A key factor in our task is that no decision-making process takes place. Instead, even in the target switch task, once the perturbation occurs there is no choice to be made about whether to switch or not—the participant must switch to the second target to finish the trial. Recently, Carroll et al. (2019) showed that when a choice with different reward values is available, reward information can improve the likelihood and latency of a goal switch. The slower latencies at which this switch occurs (~160 ms) compared to what we observe in our target switch task (~90 ms), and target and cursor jump tasks (~80-110 ms) is in line with the presence of an additional, decision-making component to the task in Carroll et al. (2019) because this decision-making component would add to the feedback loop's processing time. Unfortunately, the alternative target task design used does not allow us to determine whether this reward modulation takes place in high-order associative cortices or from premotor and sensorimotor areas (Figure 6b). However, the absence of reward modulation we observed in the target jump task in our study would be consistent with a scheme in which reward modulation of the sort reported by Carroll et al. (2019) occurs in associative areas rather than sensorimotor or premotor areas.

Our proprioception-cued and choice reaction time tasks, combined with previous published work, show that feedback responses slower than 110 ms are modulated by monetary reward. Additionally, the LLR was also modulated by reward. A distinction between reaction time loops and the LLR loop is that reward shortens latencies and increases feedback gains in the former and increases feedback gains alone in the latter. What causes these latencies in the first place? Regarding the LLR, latencies observed are mainly driven by transmission delays. Electrophysiological recordings show that proprioceptive information takes about 25 ms to reach the primary motor cortex, and a response travelling back downward would take an additional 25 ms, leading to ~50 ms latencies generally observed for this loop (Omrani et al., 2016). Consequently, the LLR has no room for latency improvements beyond transmission delays. In contrast, reaction times are often the result of integration and accumulation of sensorimotor and cognitive evidence in associative cortices, the time course of which varies depending on urgency, utility, and value of information (Fernandez-Ruiz et al., 2011; Reddi and Carpenter, 2000; Steverson et al., 2019; Thorpe and Fabre-Thorpe, 2001; Wong et al., 2017).

What is the basis of reward-driven gain increases within the short timescale of the LLR? Functionally, at least two distinct processes underlie the LLR (Lourenço et al., 2006; Matthews, 2006). An earlier, spinal process exhibits "automatic" gain scaling similar to the SLR, and a slower, supraspinal process exhibits voluntary, but not automatic gain scaling (Kimura, 2006; Lourenço et al., 2006; Pruszynski et al., 2011b). The absence of reward modulation of the SLR in our results argues against a contribution of spinal processes to the observed reward modulation of the LLR and suggests instead that the increase in feedback gain during the LLR is

implemented by supraspinal processes. In line with this, we observed that the modulation of feedback gains by reward is larger in the R3 epoch, which is dominated by supra-spinal processes, compared to the R2 epoch, which is governed mainly (though not exclusively) by spinal mechanisms (Figure 2f). We propose that reward-driven increases in feedback gain are implemented by supraspinal processes, and the observed increase in R2 feedback gain represents the early build-up of that process (Pruszynski et al., 2011b).

The supraspinal contribution to the LLR relies on many brain regions, and their temporal recruitment varies (Lee and Tatton, 1975; Omrani et al., 2016; Tatton et al., 1975). Specifically, the R2 epoch response relies mainly on the primary sensorimotor cortex (Matthews, 1991; Omrani et al., 2016; Pruszynski et al., 2011a), while additional regions such as the premotor and posterior parietal cortex (Omrani et al., 2016), supplementary motor area (Spieser et al., 2013), cerebellum (Kurtzer et al., 2013; Strick, 1983), reticular formation (Ravichandran et al., 2013; Shemmell et al., 2009) and red nucleus (Herter et al., 2015) are thought to contribute to the R3 epoch response. There is also a large amount of work showing that the LLR response can be modulated by cognitive processes such as probabilistic outcomes (Beckley et al., 1991), evidence accumulation for decision-making (Selen et al., 2012) and verbal instructions (Hammond, 1956), indicative of pre-modulation by prefrontal cortex (Hartstra et al., 2011, 2012; Noppeney et al., 2010; Philiastides et al., 2011). Our observation of reward modulation of LLR response is in line with these reports of cognitive pre-modulation.

Generally, our results combined with previous work (Carroll et al., 2019; Codol et al., 2020a; Douglas and Parry, 1983; Stillings et al., 1968) show that feedback responses that are thought to directly involve associative cortices are modulated by reward (Figure 6b). In contrast, feedback responses that do not directly recruit associative cortices show no such modulation by reward. Together with previous work showing reduction in peripheral noise with reward (Codol et al., 2020a; Manohar et al., 2019), these results enable us to further complete the picture on how rewarding information triggers improvements in motor performance at the behavioural level. Outstanding questions remain on how reward leads to motor improvement, such as whether noise reduction may also occur centrally (Goard and Dan, 2009; Manohar et al., 2015; Pinto et al., 2013), or whether the cortico-cerebellar feedback loop is also involved in reward-driven improvements (Becker and Person, 2019; Codol et al., 2020a; Miall et al., 2007).

Methods

Dataset and analysis code availability

All behavioural data and analysis code will be made freely available online on the Open Science Framework website at https://osf.io/7t8yj/ upon publication.

Participants

19 participants took part in this study and were remunerated CA$12 or 1 research credit per hour, plus performance-based remuneration. Most participants participated in all five tasks, namely target switch, SLR-LLR, proprioception-cued reaction times, cursor and target jump. The full participation roster is detailed in Table 1. In total, five sessions across five participants had to be excluded either for failing to perform at the required level for the task or as per task instructions. Participants made on average 1.62, 2.14, 1.58, 1.67 and 1.88 Canadian cents per rewarded trial on the target switch, SLR-LLR, cursor jump, target jump and reaction time task, respectively, and earned on average in total $6.98, $7.19, $4.92, $5.20 and $4.06 from performance, respectively. The tasks were divided into three sessions of 90 minutes each and performed two tasks per session maximum. The task order was counterbalanced across participants. All participants signed a consent form prior to the first session. Recruitment and data collection were done in accordance with the requirements of the research ethics board at Western University.

	Task							
Participant	Target Switch	Cursor Jump	SLR-LLR	Reaction Time	Target Jump	**Included**	**Excluded**	**Grand total**
1	1	1	1	1	1	**5**		**5**
2	1	1	1	1	1	**5**		**5**
3	1	1	1	1	1	**5**		**5**
4	1	1	1	1	1	**5**		**5**
5	1	1	1	1	1	**5**		**5**
6	1	1	1	1	1	**5**		**5**
7	1		1	1		**3**		**3**
8	1	0	1	1	1	**4**	**1**	**5**
9	1	1	1	1	1	**5**		**5**
10	1	1	1	1	1	**5**		**5**
11	1	1	1	1	1	**5**		**5**
12	1	1	0	1	1	**4**	**1**	**5**
13	0		1	1		**2**	**1**	**3**
14	1	1	1	1	1	**5**		**5**
15	1	1	1	1	1	**5**		**5**
16	1	1	1	0	1	**4**	**1**	**5**
17		1	0		1	**2**	**1**	**3**
18		1	1			**2**		**2**
Included	**15**	**15**	**16**	**15**	**15**	**76**		
Excluded	**1**	**1**	**2**	**1**			**5**	
Grand Total	**16**	**16**	**18**	**16**	**15**			**81**

Table 1: *Participant inclusion and exclusion for each experimental task. A "1" signifies that the participant data were included for the specific task and a "0" signifies that the participant performed the task but that the data collected were excluded from further analysis. An empty entry signifies that the participant did not take part in the task.*

Apparatus

A BKIN Technologies (Kingston, ON) endpoint KINARM robot was used for the cursor jump and target jump tasks, and an exoskeleton robot from the same company was used for all the other tasks. In both cases the participant was seated in front of a horizontally placed mirror that blocked vision of the participant's arm and reflected a screen above so that visual stimuli appeared in the same plane as the arm. Electromyographic activity of brachioradialis, triceps lateralis, pectoralis major, posterior deltoid, and biceps brachii was recorded using wired surface electrodes (Bagnoli, Delsys, Natick, MA). EMG and kinematics data were recorded at 1000 Hz.

For studies using the exoskeleton KINARM, the participant's arm was rested on a boom that supported the limb against gravity and allowed for movement in a horizontal plane intersecting the centre of the participant's shoulder joint. Pilot tests using an accelerometer fixed on the distal KINARM boom showed that logged perturbation timestamps corresponding to the onset of commanded robot torque preceded the acceleration of the distal end of the robot linkage by 4 ms. Perturbation timestamps were adjusted accordingly for the analysis of experimental data.

For studies using the endpoint KINARM, the robotic handle allowed for movement in a horizontal plane about 10-15 cm below the centre of the shoulder joint. In the experimental tasks tested here the start position of the robot handle was specified to be about 30 cm from the chest midline. In addition to the same five muscles that we recorded during the tasks described above using the exoskeleton KINARM, in the tasks involving the endpoint robot we also recorded the activity of the intermediate deltoid muscle. Reaction time measurements were aligned to the onset/offset of visual stimuli as measured by a using a photodiode attached to the display screen (see *Target jump and cursor jump* description in "Experimental design" section below for details).

Experimental design

General points

In all tasks using mechanical perturbations, perturbation magnitudes were added to the background load. For instance, if a background torque load of -2 Nm was applied and a -4 Nm perturbation was specified, then during the perturbation the robot produced a -6 Nm torque.

In all tasks, the start position and the target(s) were the same colour, which was either pink or cyan blue depending on whether the trial was rewarded or non-rewarded. Target colour assignment to reward conditions was counterbalanced across participants.

SLR-LLR

The location of the tip of the participant's right index finger was indicated by a 3 mm diameter white cursor. At the beginning of each trial, a 3 mm diameter start position appeared, along with a reward sign below the target showing "000" or "$$$" to indicate a non-rewarded or a rewarded trial, respectively. The start position was located so that the participant's external shoulder angle was 45 degrees relative to the left-right axis, and the external elbow angle was 90 degrees relative to the upper arm. When participants moved the cursor inside the start position the cursor disappeared. It reappeared if the participant exited the start position before the perturbation onset. After the cursor remained inside the start position for 150-200 ms, a background torque (+2 Nm) ramped up linearly in 500 ms at the shoulder and elbow to activate the extensor muscles. Then, following another 150-200 ms delay a 10 cm target appeared either at +20 or -20 degrees from the start position (rotated about the elbow joint). Following target appearance and after a 600 ms delay, the robot applied a ±2 Nm perturbation torque at the elbow and shoulder joints (Figure 2a-c). This combination of load on the shoulder and elbow was chosen to create pure elbow motion, as the robot torque applied at the shoulder counteracted the interaction torque arising at the shoulder due to elbow rotation (Maeda et al., 2018, 2020). Because the time interval between the onset of the visual target and the onset of the perturbation was fixed, we tested for anticipatory triceps EMG activity between the manipulation and control groups in a 20 ms window immediately before the perturbation onset. We observed no difference, both for the SLR (no reward: W=83, r=0.61, p=0.43; with reward: W=88, r=0.65, p=0.30) and the LLR (no reward: W=70, r=0.51, p=0.92; with reward: W=82, r=0.60, p=0.46). Following the mechanical perturbation, participants were instructed to move the cursor as fast as possible to the target and stay inside it until the end of the trial. Each trial ended 800 ms after perturbation onset, at which point the target turned dark blue, the reward sign was extinguished, and the final value of the monetary return was displayed in its place. For non-rewarded trials, this was always "0 ¢" and for rewarded trials, this was calculated as the proportion of time spent in the target from the perturbation onset to the trial's end:

$$return = ge^{-\tau p}$$

$$p = 1 - min\left(\frac{x - x_0}{x_f - x_0}, 1\right)$$

where x is the time (ms) spent in the target, $x_0 = 500$ is the minimum amount of time (ms) to receive a return, $x_f = 800$ is the total duration (ms) of the trial, $g = 15$ is the maximum return (¢), and τ is a free parameter adjusted based on pilot data to reduce the discrepancy between easier and harder conditions. In this study, we used τ = 1.428 and τ = 2.600 for an inward and outward perturbation with an outward target, respectively, and τ = 2.766 and τ = 1.351 for an inward and outward perturbation with an inward target, respectively.

The task consisted of 336 trials and was divided into three equal blocks with a free-duration break time between each block. Each block consisted of 112 trials, equally divided between inward and outward perturbation torques, inward and outward target positions, and rewarded and non-rewarded trials. The trial schedule was organised in epochs of 16 trials containing two of each combination of conditions, and the trial order was randomized within each epoch.

For EMG analysis, inward and outward perturbations were used as the manipulation and control conditions, respectively, to observe the SLR on extensor muscles. To observe the LLR on the extensor muscles, inward perturbations were used when combined with an outward and inward target for the manipulation and control condition, respectively.

Reaction times

The proprioception-cued reaction time task used the same task as the SLR-LLR task, with several alterations. First, the background loads were applied to the elbow only, and inverted to -2 Nm to pre-activate the flexor muscles. Only inward targets were presented, meaning the task consisted of elbow flexion movements only. The starting position was located such that the external shoulder angle was 5 degrees relative to the left-right axis, and the external elbow angle was 90 degrees relative to the upper limb. Finally, the perturbation was applied only at the shoulder instead of both shoulder and elbow joints, and the perturbation magnitude was reduced to 0.5 Nm, meaning that the perturbation led to no significant elbow motion. Participants were informed to initiate a movement to the target as fast as possible following the shoulder perturbation. Monetary returns were calculated using the same formula as for the SLR-LLR task but with $\tau = 1.278$

For the choice reaction time task, the methods employed are described in Codol et al. (2020).

Target jump and cursor jump

The position of the endpoint KINARM handle held by participants was indicated by a 5 mm diameter white cursor. At the beginning of each trial, an 8 mm diameter start position was displayed, and below it a reward sign showing “000” or “$$$”to indicate a non-rewarded or a rewarded trial, respectively. At the same time, a 10 mm diameter target appeared 25 cm straight ahead from the start position (Figure 3f). When participants moved the cursor inside the start position a series of 4 tones were played on a speaker located beside the display screen, and participants were instructed to start the movement on the third tone and finish on the fourth tone. They were informed that their monetary returns were proportional to how close to a tone interval a reaching movement duration was. Movement duration was defined as the time interval between exiting the start position and when the cursor was inside the target and its tangential velocity dropped below 2 cm/sec. Once the cursor was inside the target and its tangential velocity dropped below 2 cm/sec, the target turned dark blue and the value of the monetary return for that trial was displayed 3 cm straight ahead of the target. In the target jump task, when the movement crossed past 13 cm from the start position, the target jumped

immediately 3 cm either right or left from its original position or stayed at its original position (no-jump), with all three possibilities occurring with equal frequency. In the cursor jump task, the cursor position rather than the target jumped right or left or did not jump.

Monetary reward was given by an equation based on deviation from the desired movement time, which corresponded to the duration of an interval between two acoustic tones:

$$p = \left|\frac{MT - MT^*}{MT_{dev}}\right|$$

$$return = ge^{-\tau p} + b$$

where MT is the trial's movement time (ms), $MT^* = 600$ is the desired movement time (ms), $MT_{dev} = 300$ is the maximum deviation from the desired movement time (ms) to be rewarded, and g, b, and τ are free parameters. For right and left jump trials, we used $g = 0.136$, $b = -0.067$, and $\tau = 0.856$. For no-jump trials, we used $g = 0.128$, $b = -0.067$, and $\tau = 0.502$.

Both the target and cursor jumps consisted of 312 trials and were divided into two blocks with a free-duration break time between each block. Each block consisted of 156 trials, equally divided between rewarded and non-rewarded trials, and a third of leftward trials, rightward trials, and no-jump trials. The trial schedule was organised in epochs of 12 trials containing two of each combination of conditions, and the trial order was randomized within each epoch.

For EMG analysis, leftward target jumps and rightward cursor jumps were used as the manipulation condition because they both led to contraction of the pectoralis muscle to adjust the movement. Rightward target jumps and leftward cursor jumps were used as control conditions. No-jump conditions were not used for EMG analysis. All EMG signals were aligned to a photodiode signal recording the appearance of a white, 8 mm diameter target at the same time as the jump occurrence. The photodiode target was positioned on the screen horizontally 35 cm to the left from the starting position and vertically at the same position as the cursor jumping position (in the cursor jump task) or as the target position (target jump task). This target was covered by the photodiode sensor and was therefore not visible to participants.

Target switch

The position of participants' right index fingertip was indicated by a 3 mm diameter white cursor. At the beginning of each trial, a 1 cm start position appeared, and a reward sign showing "000" or "$$$" was displayed to indicate a non-rewarded or a rewarded trial, respectively. The start position was located so that the external shoulder angle was 45 degrees relative to the left-right axis, and the external elbow angle was 90 degrees relative to the upper arm. When participants moved the cursor inside the start position the cursor disappeared. it reappeared if the cursor exited the start position before the target(s) appeared. Once inside the start position, the robot applied -2 Nm background torques which were ramped up linearly in 500 ms at the shoulder and elbow to activate the flexor muscles. Then, following a delay of 600 ms to record baseline EMG activity, a 2 cm target appeared at +20 degrees (inward) from the start position (rotated about the elbow joint). In half of trials, a second, identical target also

appeared at -20 degrees (outward) from the start position. Once the target(s) appeared, participants initiated their reaching movement. Participants were instructed to always reach for the inward target, even if the second, outwards target also appeared. During the movement itself, the cursor position was not visible to prevent any visual feedback from contributing to performance. In half of trials, a -4 Nm perturbation occurred at the elbow and shoulder joints when the movement crossed an invisible 2 cm threshold distance from the centre of the start position. Note that this perturbation led to an outward perturbation because the torque is negative**.** Participants were instructed to push through the perturbation to reach and stop inside the inward target if only the inward target was available, leading to increased brachioradialis activity (control condition); if two targets were available, they were instructed to relax their arm and let the perturbation push them toward the outward target and stop there (manipulation condition). Therefore, this task design resulted in a co-occurrence of the target switch and divergence of EMG activity at the brachioradialis level, enabling us to assess the feedback response that underlies goal target switch. In trials with no perturbations, participants were instructed to carry on the reach toward the inward target, regardless of whether one or two targets were displayed in that trial.

Cursor feedback was restored when the hand was less than 3 cm away from the centre of the target. When the (correct) target was reached, all the targets turned dark blue, the reward sign was extinguished, and the final monetary return for the trial appeared where the reward sign was located before. For non-rewarded trials, this was always “0 ¢” and for rewarded trials, the return was higher for shorter movement times. Movement times were defined as the time interval (ms) between exiting the start position and the trial finishing. The start position was considered as exited when the radial distance from its centre was greater than 2 cm, at which point the perturbation would then occur if the trial included one. A trial was considered finished when the cursor was inside the (correct) target and its tangential velocity dropped below 10 cm/sec. The exact return equation was:

$$p = \frac{MT}{MT_{norm}}$$

$$return = g * max(ae^{-\tau p} + b\ , 0)$$

where MT is the trial’s movement time (ms), $MT_{norm} = 600$ is a normalising scalar, $g = 25$ is the maximum return (dollar cents), and $a, b = -0.267$, and τ are free parameters. The parameter a was set at 0.587 and 0.749 for two-target trials with and without a perturbation, respectively, and 0.985 and 0.739 for one-target trials with and without a perturbation, respectively. The parameter τ was set at 0.331 and 1.382 for two-target trials with and without a perturbation, respectively, and 1.346 and 1.343 for one-target trials with and without a perturbation, respectively.

The task consisted of 432 trials and was divided into 3 blocks with a break time between each block. Each block consisted of 144 trials, equally divided between one- and two-targets trials, perturbation and no-perturbation trials, and rewarded and non-rewarded trials. The trial schedule was organised in epochs of 16 trials containing two of each combination of conditions, and the trial order was randomized within each epoch.

EMG signal processing

EMG signals were sampled at 1000 Hz, band-pass filtered between 20 Hz and 250 Hz, and full-wave rectified. Before each task, participants' baseline EMG signal was acquired by asking participants to position their arm such that the cursor remained, motionless, at the start position for 2 seconds (against the background load, if applicable for the task). This was repeated 6 times, after which the task started normally. Following band-pass filtering and full-wave rectification, the EMG signal of each muscle from 250 ms after entering the start position to 250 ms before the end of the 2 seconds window was concatenated across all 6 trials and averaged to obtain a mean baseline EMG scalar. EMG measures during the task were then normalised by each muscle's baseline scalar. Follow-up analyses (latency, feedback gains, co-contraction) were performed subsequently on the filtered, full wave rectified and normalised EMG traces.

Statistical analysis

To determine the time at which EMG signals for different task conditions diverged, we used Receiver operating characteristic (ROC) analysis. We used the same approach as in Weiler et al. (2015), using a 25-75% threshold of area under the curve (AUC) for establishing signal discrimination. Discrimination was done for each participant and each reward condition independently, using all trials available for each contrast without averaging. Once the AUC threshold was crossed, we performed a segmented linear regression on the AUC before it crossed the 25-75% threshold. We minimized the sums-of-squared residuals to find the inflexion point, that is, where the two segments of the segmented linear regression form an angle (see Weiler et al. (2015) and analysis code online for details).

To compute feedback gains, for each feedback response considered we defined a 25 ms window that started at that response's latency found for each participant independently using ROC analysis. We then calculated the integral of average EMG signals in that window using the trapezoid rule (MATLAB's built-in *trapz* function), for the control and manipulation condition and each reward value. For instance, for the target switch task the control condition is defined as trials with a perturbation and only one target (no switch), while the manipulation condition is defined as trials with a perturbation and two targets (switch occurring). We then calculated the absolute difference between those two conditions as a measure of feedback gains. We then calculated the log-ratio of the rewarded to non-rewarded conditions as $\log(rewarded\ gain/non\ rewarded\ gain)$. For the R2 and R3 epochs of the LLR, the 25 ms window was defined as the 1st-to-25th ms and 26th-to-50th ms post-response, respectively. To test for the epoch-reward interaction between R2 and R3, the rewarded gain and non-rewarded gain for each epoch was used directly instead of computing their log-ratio. We used ratios to ensure that changes in feedback gains are normalized within participants to EMG activity in the non-rewarded condition. This is because EMG activity levels can greatly vary over

time, making normalization by subtraction unviable. The log function was then applied to linearize the ratio values.

To test for differences between conditions we used Wilcoxon signed-rank tests. For each test, we reported the test statistic W, the effect size r (Kerby, 2014) and the p-value. To compare LLR feedback gains between the R2 and R3 epochs, a repeated-measure Analysis of Variance was used, as there is no established non-parametric equivalent for repeated measures designs.

Contributions

OC conceptualized and designed the research question; OC, CJF, JMG, and PLG designed the experiments; OC implemented the tasks, and collected and analysed the datasets; OC, CJF, JMG and PLG interpreted the results; OC made the figures and wrote the first draft; OC, CJF, JMG, and PLG edited and approved the final version of the manuscript.

Acknowledgements

This work was supported by the Natural Science and Engineering Council of Canada (RGPIN-2018-05458 to PLG) and the Canadian Institutes of Health Research (PJT-156241 to PLG). We thank Jonathan M. Michaels and J. Andrew Pruszynski for helpful comments and suggestions.

Competing interests

The authors declare no competing financial or non-financial interests.

References

Becker, M.I., and Person, A.L. (2019). Cerebellar Control of Reach Kinematics for Endpoint Precision. Neuron *103*, 335–348.

Beckley, D.J., Bloem, B.R., Remler, M.P., Roos, R.A.C., and Van Dijk, J.G. (1991). Long latency postural responses are functionally modified by cognitive set. Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section *81*, 353–358.

Carroll, T.J., McNamee, D., Ingram, J.N., and Wolpert, D.M. (2019). Rapid Visuomotor Responses Reflect Value-Based Decisions. J. Neurosci. *39*, 3906–3920.

Cashaback, J.G.A., Lao, C.K., Palidis, D.J., Coltman, S.K., McGregor, H.R., and Gribble, P.L. (2019). The gradient of the reinforcement landscape influences sensorimotor learning. PLoS Comput Biol *15*, e1006839.

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

Chen, X., Rutledge, R.B., Brown, H.R., Dolan, R.J., Bestmann, S., and Galea, J.M. (2018a). Age-dependent Pavlovian biases influence motor decision-making. PLoS Comput Biol *14*, e1006304.

Chen, X., Holland, P., and Galea, J.M. (2018b). The effects of reward and punishment on motor skill learning. Current Opinion in Behavioral Sciences *20*, 83–88.

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

Codol, O., Holland, P.J., Manohar, S.G., and Galea, J.M. (2020a). Reward-Based Improvements in Motor Control Are Driven by Multiple Error-Reducing Mechanisms. J. Neurosci. *40*, 3604–3620.

Codol, O., Galea, J.M., Jalali, R., and Holland, P.J. (2020b). Reward-driven enhancements in motor control are robust to TMS manipulation. Exp Brain Res *238*, 1781–1793.

Crone, C., and Nielsen, J. (1994). Central control of disynaptic reciprocal inhibition in humans. Acta Physiologica Scandinavica *152*, 351–363.

Day, B.L., and Brown, P. (2001). Evidence for subcortical involvement in the visual control of human reaching. Brain *124*, 1832–1840.

De Comité, A., Crevecoeur, F., and Lefèvre, P. (2021). Reward-dependent selection of feedback gains impacts rapid motor decisions.

Desmurget, M., Gaveau, V., Vindras, P., Turner, R.S., Broussolle, E., and Thobois, S. (2004). On-line motor control in patients with Parkinson's disease. Brain *127*, 1755–1773.

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.

Douglas, V.I., and Parry, P.A. (1983). Effects of reward on delayed reaction time task performance of hyperactive children. J Abnorm Child Psychol *11*, 313–326.

Fernandez-Ruiz, J., Wong, W., Armstrong, I.T., and Flanagan, J.R. (2011). Relation between reaction time and reach errors during visuomotor adaptation. Behavioural Brain Research *219*, 8–14.

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

Goard, M., and Dan, Y. (2009). Basal forebrain activation enhances cortical coding of natural scenes. Nature Neuroscience *12*, 1444–1449.

Goodman, R.N., Rietschel, J.C., Roy, A., Jung, B.C., Diaz, J., Macko, R.F., and Forrester, L.W. (2014). Increased reward in ankle robotics training enhances motor control and cortical efficiency in stroke. Journal of Rehabilitation Research and Development *51*, 213–228.

de Graaf, J.B., Frolov, A., Fiocchi, M., Nazarian, B., Anton, J.-L., Pailhous, J., and Bonnard, M. (2009). Preparing for a motor perturbation: Early implication of primary motor and somatosensory cortices. Human Brain Mapping *30*, 575–587.

Guitart-Masip, M., Duzel, E., Dolan, R., and Dayan, P. (2014). Action versus valence in decision making. Trends in Cognitive Sciences *18*, 194–202.

Haith, A.M., Huberdeau, D.M., and Krakauer, J.W. (2015). The Influence of Movement Preparation Time on the Expression of Visuomotor Learning and Savings. Journal of Neuroscience *35*, 5109–5117.

Hammond, P.H. (1956). The influence of prior instruction to the subject on an apparently involuntary neuro-muscular response. J Physiol *132*, 17-18P.

Hartstra, E., Kühn, S., Verguts, T., and Brass, M. (2011). The implementation of verbal instructions: An fMRI study. Human Brain Mapping *32*, 1811–1824.

Hartstra, E., Waszak, F., and Brass, M. (2012). The implementation of verbal instructions: Dissociating motor preparation from the formation of stimulus–response associations. NeuroImage *63*, 1143–1153.

Herter, T.M., Takei, T., Munoz, D.P., and Scott, S.H. (2015). Neurons in red nucleus and primary motor cortex exhibit similar responses to mechanical perturbations applied to the upper-limb during posture. Front. Integr. Neurosci. *0*.

Hübner, R., and Schlösser, J. (2010). Monetary reward increases attentional effort in the flanker task. Psychonomic Bulletin & Review *17*, 821–826.

Izawa, J., and Shadmehr, R. (2011). Learning from Sensory and Reward Prediction Errors during Motor Adaptation. PLoS Comput Biol *7*, e1002012.

Kerby, D.S. (2014). The Simple Difference Formula: An Approach to Teaching Nonparametric Correlation. Comprehensive Psychology *3*, 11.IT.3.1.

Kimura, T. (2006). Transcranial Magnetic Stimulation over Sensorimotor Cortex Disrupts Anticipatory Reflex Gain Modulation for Skilled Action. Journal of Neuroscience *26*, 9272–9281.

Kurtzer, I., Trautman, P., Rasquinha, R.J., Bhanpuri, N.H., Scott, S.H., and Bastian, A.J. (2013). Cerebellar damage diminishes long-latency responses to multijoint perturbations. J Neurophysiol *109*, 2228–2241.

Lee, R.G., and Tatton, W.G. (1975). Motor Responses to Sudden Limb Displacements in Primates with Specific CNS Lesions and in Human Patients with Motor System Disorders. Canadian Journal of Neurological Sciences *2*, 285–293.

Lourenço, G., Iglesias, C., Cavallari, P., Pierrot-Deseilligny, E., and Marchand-Pauvert, V. (2006). Mediation of late excitation from human hand muscles via parallel group II spinal and group I transcortical pathways. The Journal of Physiology *572*, 585–603.

Lowet, A.S., Zheng, Q., Matias, S., Drugowitsch, J., and Uchida, N. (2020). Distributional Reinforcement Learning in the Brain. Trends in Neurosciences *43*, 980–997.

Maeda, R.S., Cluff, T., Gribble, P.L., and Pruszynski, J.A. (2018). Feedforward and Feedback Control Share an Internal Model of the Arm's Dynamics. J. Neurosci. *38*, 10505–10514.

Maeda, R.S., Zdybal, J.M., Gribble, P.L., and Pruszynski, J.A. (2020). Generalizing movement patterns following shoulder fixation. Journal of Neurophysiology *123*, 1193–1205.

Manley, H., Dayan, P., and Diedrichsen, J. (2014). When Money Is Not Enough: Awareness, Success, and Variability in Motor Learning. PLoS ONE *9*, e86580.

Manohar, S.G., Chong, T.T.-J., Apps, M.A.J., Batla, A., Stamelou, M., Jarman, P.R., Bhatia, K.P., and Husain, M. (2015). Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control. Current Biology *25*, 1707–1716.

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

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

Matthews, P.B.C. (1991). The human stretch reflex and the motor cortex. Trends in Neurosciences *14*, 87–91.

Matthews, P.B.C. (2006). Restoring balance to the reflex actions of the muscle spindle: the secondary endings also matter. The Journal of Physiology *572*, 309–310.

McDougle, S.D., Ballard, I.C., Baribault, B., Bishop, S.J., and Collins, A.G.E. (2021). Executive Function Assigns Value to Novel Goal-Congruent Outcomes. Cerebral Cortex bhab205.

Miall, R.C., Christensen, L.O.D., Cain, O., and Stanley, J. (2007). Disruption of State Estimation in the Human Lateral Cerebellum. PLoS Biology *5*, e316.

Nielsen, J., and Kagamihara, Y. (1992). The regulation of disynaptic reciprocal Ia inhibition during co-contraction of antagonistic muscles in man. The Journal of Physiology *456*, 373–391.

Nielsen, J., and Kagamihara, Y. (1993). The regulation of presynaptic inhibition during co-contraction of antagonistic muscles in man. The Journal of Physiology *464*, 575–593.

Noppeney, U., Ostwald, D., and Werner, S. (2010). Perceptual Decisions Formed by Accumulation of Audiovisual Evidence in Prefrontal Cortex. J. Neurosci. *30*, 7434–7446.

Omrani, M., Murnaghan, C.D., Pruszynski, J.A., and Scott, S.H. (2016). Distributed task-specific processing of somatosensory feedback for voluntary motor control. ELife *5*, e13141.

Parvin, D.E., McDougle, S.D., Taylor, J.A., and Ivry, R.B. (2018). Credit Assignment in a Motor Decision Making Task Is Influenced by Agency and Not Sensory Prediction Errors. J. Neurosci. *38*, 4521–4530.

Pekny, S.E., Izawa, J., and Shadmehr, R. (2015). Reward-Dependent Modulation of Movement Variability. Journal of Neuroscience *35*, 4015–4024.

Philiastides, M.G., Auksztulewicz, R., Heekeren, H.R., and Blankenburg, F. (2011). Causal Role of Dorsolateral Prefrontal Cortex in Human Perceptual Decision Making. Curr. Biol. *21*, 980–983.

Pinto, L., Goard, M.J., Estandian, D., Xu, M., Kwan, A.C., Lee, S.-H., Harrison, T.C., Feng, G., and Dan, Y. (2013). Fast modulation of visual perception by basal forebrain cholinergic neurons. Nature Neuroscience *16*, 1857–1863.

Poscente, S., Peters, R., Cashaback, J., and Cluff, T. (2021). Rapid Feedback Responses Parallel the Urgency of Voluntary Reaching Movements. Neuroscience.

Pruszynski, J.A., Kurtzer, I., and Scott, S.H. (2008). Rapid Motor Responses Are Appropriately Tuned to the Metrics of a Visuospatial Task. Journal of Neurophysiology *100*, 224–238.

Pruszynski, J.A., Kurtzer, I., Nashed, J.Y., Omrani, M., Brouwer, B., and Scott, S.H. (2011a). Primary motor cortex underlies multi-joint integration for fast feedback control. Nature *478*, 387–390.

Pruszynski, J.A., Kurtzer, I., and Scott, S.H. (2011b). The long-latency reflex is composed of at least two functionally independent processes. Journal of Neurophysiology *106*, 449–459.

Quattrocchi, G., Greenwood, R., Rothwell, J.C., Galea, J.M., and Bestmann, S. (2017). Reward and punishment enhance motor adaptation in stroke. J Neurol Neurosurg Psychiatry *88*, 730.

Ravichandran, V.J., Honeycutt, C.F., Shemmell, J., and Perreault, E.J. (2013). Instruction-dependent modulation of the long-latency stretch reflex is associated with indicators of startle. Exp Brain Res *230*, 59–69.

Reddi, B. a. J., and Carpenter, R.H.S. (2000). The influence of urgency on decision time. Nat Neurosci *3*, 827–830.

Reschechtko, S., and Pruszynski, J.A. (2020). Stretch reflexes. Current Biology *30*, R1025–R1030.

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

Scott, S.H. (2012). The computational and neural basis of voluntary motor control and planning. Trends in Cognitive Sciences *16*, 541–549.

Scott, S.H. (2016). A Functional Taxonomy of Bottom-Up Sensory Feedback Processing for Motor Actions. Trends in Neurosciences *39*, 512–526.

Selen, L.P.J., Shadlen, M.N., and Wolpert, D.M. (2012). Deliberation in the Motor System: Reflex Gains Track Evolving Evidence Leading to a Decision. J. Neurosci. *32*, 2276–2286.

Shemmell, J., An, J.H., and Perreault, E.J. (2009). The Differential Role of Motor Cortex in Stretch Reflex Modulation Induced by Changes in Environmental Mechanics and Verbal Instruction. J. Neurosci. *29*, 13255–13263.

Sherrington, C.S. (1906). The integrative action of the nervous system (New Haven, CT, US: Yale University Press).

Sherrington, C.S. (1913). Reflex Inhibition as a Factor in the Co-Ordination of Movements and Postures. Quarterly Journal of Experimental Physiology *6*, 251–310.

Spieser, L., Aubert, S., and Bonnard, M. (2013). Involvement of SMAp in the intention-related long latency stretch reflex modulation: a TMS study. Neuroscience *246*, 329–341.

Steverson, K., Chung, H.-K., Zimmermann, J., Louie, K., and Glimcher, P. (2019). Sensitivity of reaction time to the magnitude of rewards reveals the cost-structure of time. Sci Rep *9*, 20053.

Stillings, N.A., Allen, G.A., and Estes, W.K. (1968). Reaction time as a function of noncontingent reward magnitude. Psychon Sci *10*, 337–338.

Strick, P.L. (1983). The influence of motor preparation on the response of cerebellar neurons to limb displacements. J Neurosci *3*, 2007–2020.

Summerside, E.M., Shadmehr, R., and Ahmed, A.A. (2018). Vigor of reaching movements: reward discounts the cost of effort. Journal of Neurophysiology *119*, 2347–2357.

Tatton, W.G., Forner, S.D., Gerstein, G.L., Chambers, W.W., and Liu, C.N. (1975). The effect of postcentral cortical lesions on motor responses to sudden upper limb displacements in monkeys. Brain Research *96*, 108–113.

Therrien, A.S., Wolpert, D.M., and Bastian, A.J. (2016). Effective reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. Brain *139*, 101–114.

Therrien, A.S., Wolpert, D.M., and Bastian, A.J. (2018). Increasing Motor Noise Impairs Reinforcement Learning in Healthy Individuals. 14.

Thorpe, S.J., and Fabre-Thorpe, M. (2001). Seeking Categories in the Brain. Science *291*, 260–263.

Weiler, J., Gribble, P.L., and Pruszynski, J.A. (2015). Goal-dependent modulation of the long-latency stretch response at the shoulder, elbow, and wrist. J. Neurophysiol. *114*, 3242–3254.

Weiler, J., Gribble, P.L., and Pruszynski, J.A. (2019). Spinal stretch reflexes support efficient hand control. Nat Neurosci *22*, 529–533.

Wong, A.L., Goldsmith, J., Forrence, A.D., Haith, A.M., and Krakauer, J.W. (2017). Reaction times can reflect habits rather than computations. ELife *6*, e28075.

Zonnino, A., Farrens, A.J., Ress, D., and Sergi, F. (2021). Measurement of stretch-evoked brainstem function using fMRI. Sci Rep *11*, 12544.