Influence of Reward on Corticospinal Excitability during Movement Preparation

Pierre-Alexandre Klein, Etienne Olivier, and Julie Duque
Institute of Neuroscience, Université catholique de Louvain, B-1200 Brussels, Belgium

Current models of decision making postulate that action selection entails a competition within motor-related areas. According to this view, during action selection, motor activity should integrate cognitive information (e.g., reward) that drives our decisions. We tested this hypothesis in humans by measuring motor-evoked potentials (MEPs) in a left finger muscle during motor preparation in a hand selection task, in which subjects performed left or right key presses according to an imperative signal. This signal was either obvious or ambiguous, but subjects were always asked to react as fast as possible. When the signal was really indistinct, any key press was regarded as correct, so subjects could respond "at random" in those trials. A score based on reaction times was provided after each correct response, and subjects were told they would receive a monetary reward proportional to their final score. Importantly, the scores were either equitable for both hands or favored implicitly left responses (reward$_{neutral}$ and reward$_{biased}$ blocks, respectively). We found that subjects selected their left hand more often in the reward$_{biased}$ than in the reward$_{neutral}$ condition, particularly after ambiguous signals. Moreover, left MEPs were larger, as soon as the signal appeared, in the reward$_{biased}$ than in the reward$_{neutral}$ conditions. During the course of motor preparation, this effect became strongest following ambiguous signals, a condition in which subjects' choices relied strongly on the reward. These results indicate that motor activity is shaped by a cognitive variable that drives our choices, possibly in the context of a competition taking place within motor-related areas.

Introduction

Traditional approaches have viewed decision making as emerging from a serial process involving independent and successive perceptive, cognitive, and motor operations (Posner, 1978). In the context of action selection, this approach has emphasized that decisions are made at an effector-independent level, with the motor system being involved only after this process, to execute the selected actions. However, more recent computational and neurophysiological studies view decision making as a bounded-accumulation system linking perception, cognition, and action in an integrative parallel process (Link and Heath, 1975; Brown and Heathcote, 2005; Pearson et al., 2009). Such models assume the dynamic activation of multiple action plans in motor-related areas, set by ongoing contextual cues, with a neural activity related to the response preparation reflecting a competition between these simultaneously activated motor plans (Hasbroucq et al., 1997; Cisek, 2006). Selection would occur when the activation associated with one action reaches a certain threshold (Domenech and Dreher, 2010; Kim and Basso, 2010).

A corollary of this idea is that the buildup of activity of potential action representations during response selection should be regulated by the cognitive cues that drive our decisions, possibly through top-down influences originating from the prefrontal cortex (Ridderinkhof et al., 2004). Reward is an important cognitive variable; we typically tend to choose actions that are most rewarding in a given context (Sutton and Barto, 1981). Based on a parallel process, this preference should be reflected in the pattern of motor activations during action selection so that the representation of the most rewarding action reaches threshold first, as shown in human (Selen et al., 2012) and nonhuman primates (Roesch and Olson, 2004; Pastor-Bernier and Cisek, 2011).

Information about reward could tune motor activations by regulating the rate of the activation rise (called the "slope shift" hypothesis; see Fig. 1A, right) or by modulating the level of activity of action representations at the onset of the selection process to adjust the distance to threshold ("starting-point shift" hypothesis; see Fig. 1A, left) (Mazurek et al., 2003; Gold and Shadlen, 2007).

Here, we measured motor-evoked potentials (MEPs) in a left hand muscle while subjects performed a hand selection task requiring to execute left or right key presses according to an imperative signal. The latter was sometimes ambiguous. Hence, subjects had to respond at "random" on a proportion of trials. A monetary reward was provided after each response and was either equal for both hands or higher for left- than right-hand responses (reward$_{neutral}$ and reward$_{biased}$ conditions, respectively). Behaviorally speaking, we expected that subjects would choose their left hand more often in the reward$_{biased}$ compared to the reward$_{neutral}$ condition, especially in ambiguous trials. Moreover, based on the

A

Figure 1. *A*, Schematic representation of how reward may regulate the buildup of motor activity during action selection. One possibility is that it adjusts the initial level of activity of potential action representations at the start of the competition process (starting-point shift hypothesis, early effect, left). Another possibility is that reward modulates the mean rate of the accumulation during the selection process (slope shift hypothesis, late effect, right). *B*, Schematic representation of our results. CS excitability changes are shown separately for ambiguous (left) and obvious signals (right). Reward influenced CS excitability at the onset of the imperative signal (early effect). As such, it induced a starting-point shift. In addition, reward induced a slope shift (late effect greater than early effect) following ambiguous but not obvious signals.

starting-point shift hypothesis, we predicted that left MEPs would be larger in the reward$_{biased}$ compared to the reward$_{neutral}$ condition as soon as the imperative signal appears (early reward effect; see Fig. 1*A*, left). The slope shift hypothesis leads to a different prediction: the reward effect on MEPs should grow during the decision period (late effect greater than early effect; see Fig. 1*A*, right).

Materials and Methods

Participants

A total of 29 subjects participated in one of two main experiments [behavioral experiment, $n = 8$; 4 women; mean age, 25.5 ± 1.23 years old; transcranial magnetic stimulation (TMS) experiment, $n = 11$; 8 women; mean age, 25.4 ± 1.62 years old] or in a control experiment ($n = 10$; 4 women; mean age, 25.2 ± 1.45 years old). None of the participants had any neurological disorder or history of psychiatric illness or drug or alcohol abuse, or were on any drug treatments that could influence performance. All the subjects were right handed according to the condensed version of the Edinburgh Handedness Inventory (Oldfield, 1971) and were financially compensated for their participation. Importantly, they were naive to the purpose of the study. The protocol was approved by the Ethics Committee of the Université catholique de Louvain (Belgium) and all subjects gave written informed consent for their participation.

Hand selection task

In the two main experiments (behavioral and TMS), subjects performed a choice reaction time task, which was implemented by means of Matlab 6.5 (MathWorks) and the Cogent 2000 toolbox (Functional Imaging Laboratory, Laboratory of Neurobiology, and Institute of Cognitive Neuroscience, Wellcome Department of Imaging Neuroscience, Lon-

don, UK). Subjects had to make a binary choice, as quickly as possible, after the onset of an imperative signal, a small colored circle, which was briefly presented at the center of a computer screen, positioned ∼60 cm in front of them. Subjects were instructed to respond with the left index finger following the display of a blue circle and with the right index finger following a red circle presentation. Importantly, the circles could be filled in with different tints of blue and red. Each tint was obtained, respectively, by increasing the saturation of the blue (B) channel or the red (R) channel in the red–green–blue (RGB) model (Fig. 2*A*). As such, the circle colors ranged from "obviously blue" (RGB, [127, 127, 154]), urging for a left-hand response, to "obviously red" (RGB, [154, 127, 127]), urging for a right-hand response, with, in the middle range, a set of more ambiguous grayish tints that were less (or not) saturated in blue or red (gray RGB, [127, 127, 127]). We used a total of 19 different linearly determined RGB color shades (steps of three units in the R or B channel) that were grouped into five categories: "obvious" blue (B channel saturation: 139, 142, 145, 148, 151, and 154), "weak" blue (B channel saturation, 133 and 136), "indistinct" color (B and/or R channel saturation, 127 and 130), "weak" red (R channel saturation, 133 and 136), and "obvious" red (R channel saturation, 139, 142, 145, 148, 151, and 154; Fig. 2*A*). These categories were defined based on pilot experiments in which we determined colors that were easily discernible (obvious category, ∼100% success) or more difficult to discriminate (weak category, ∼70% success). The indistinct category included gray color tints that subjects could not discriminate. The subjects were not told about the different categories and were asked to always make a choice as quickly as possible even if they were not sure about the color of the imperative signal. Note that the indistinct and weak color categories were pooled together for the MEP analysis in the TMS experiment (see below).

The participants sat in front of the computer screen with both forearms resting on a pillow in a semiflexed position and with hands placed palms down on a keyboard. The keyboard was turned upside down so that subjects could press on the required buttons with the left and right index fingers (keys F12 and F5, respectively; Fig. 2*A*). Between each trial, subjects were asked to relax their index fingers on two small yellow rubber pads, which were positioned on the external side of the two target buttons. Hence, each key press required subjects to perform a brisk flexion and abduction movement of the left or right index finger. Note that a strong emphasis was put on the execution of strictly unilateral movements. This aspect of behavior was controlled by continuously looking at the electromyography (EMG) of the left and right first dorsal interosseous (FDI) muscles (agonist in index finger flexion and abduction) during the experiments.

Each trial started with the presentation of a warning signal, a fixation cross (+), displayed at the center of the screen for 500 ms (Fig. 2*B*). This signal indicated the beginning of a trial and was followed by a 500 ms delay period (blank screen). Then, the imperative signal, i.e., the colored circle, appeared and remained on the screen until one of the response buttons was pressed; in the absence of any key press, it disappeared after a delay corresponding to twice the individual median reaction time (RT) value (measured during the training blocks; see below). RTs were computed by means of a homemade hardware (PSB). In brief, the PSB is a system based on a microcontroller (μC; MSP430F249; Texas Instruments) receiving video graphics array (VGA) and keyboard events: a timer starts on specific VGA events (imperative signal) and stops on keyboard events (finger response). The μC sends the pressed key code and the timer value (128 μs resolution) to the main computer through a USB interface, providing RT measurements with very high temporal resolution. Finally, after the offset of the imperative signal, a visual feedback is presented on the screen for 1000 ms.

The feedback consisted of a numerical score value: it was positive and displayed in green if the response was correct; it was negative and depicted in red following an erroneous response. Again, all responses were regarded as correct following indistinct color signals. In the case of a correct response, the positive score was always inversely proportional to the RT [k/RT^2 (in ms) with $k = 15 \times 10^5$]. Hence, the faster the response was (i.e., the smaller the RT), the higher the score. In contrast, following an incorrect response (or in the absence of response), subjects received a negative score that was proportional to the clearness of the circle color
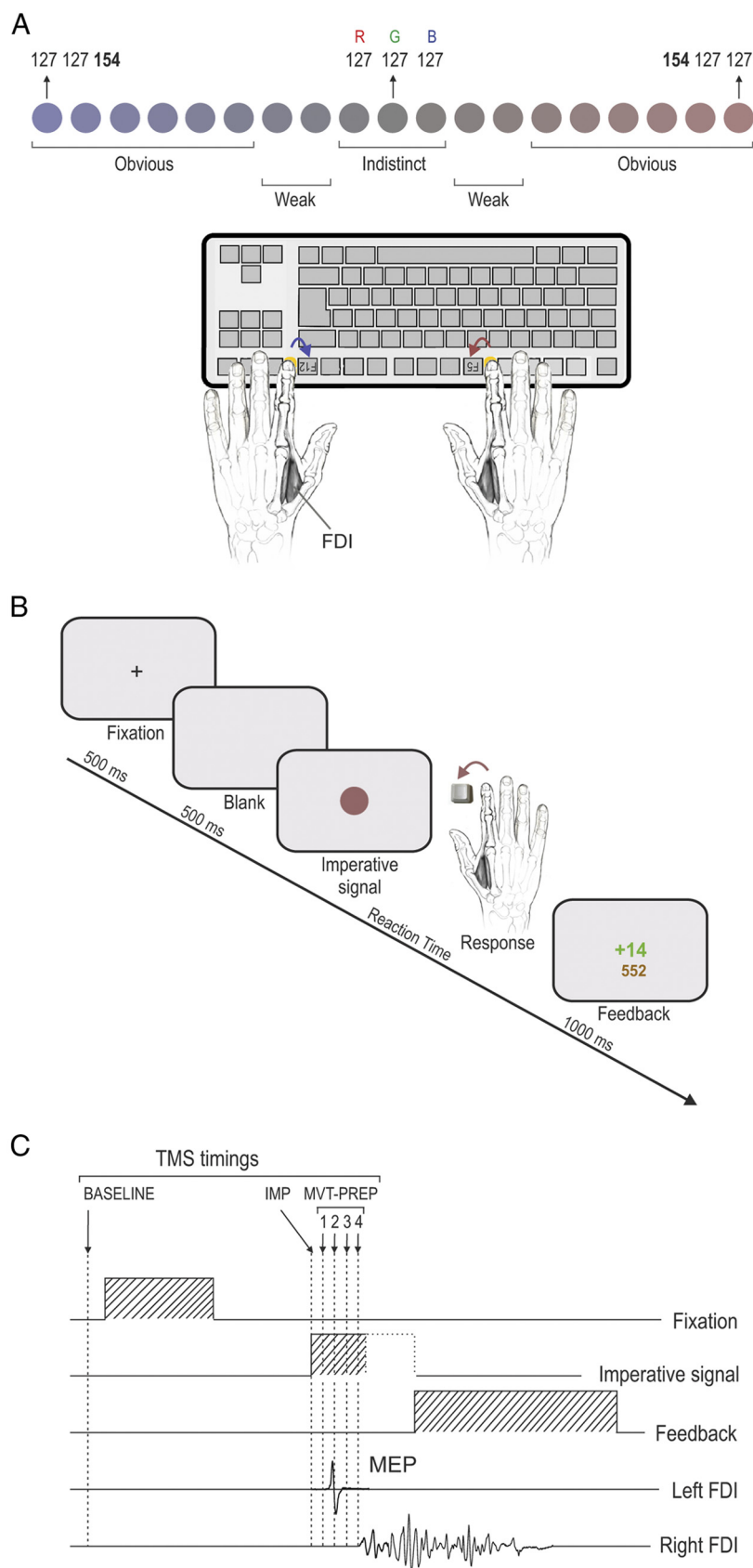
Figure 2. ***A***, Imperative signal consisted of a colored circle displayed in the center of the screen. A range of bluish or reddish saturations filled the circle and indicated a left- or right-hand response, respectively. This color range was subdivided into different categories (indistinct, weak, and obvious), depending on our behavioral results. Subjects had to make an index abduction (left FDI muscle) to selectively and unilaterally press buttons on an keyboard placed upside down. ***B***, Time course of a trial. A visual feedback, computed as a function of the reaction time (see Materials and Methods), was displayed once the response was executed.

[-1.6 times the unit of saturation above 127; maximum withdrawal, -43 (-1.6×27) for the most saturated color ($154 = 127 + 27$); Fig. 2*A*]. Hence, the easier it was to discriminate the color the larger the penalty, regardless of the RT. Finally, in addition to this within-trial score, the feedback screen also displayed the total amount of points accumulated since the beginning of the block. Subjects were told that they would get a financial bonus proportional to their final score. Trials were separated by a blank screen lasting for a variable interval of 1500 to 2000 ms.

Importantly, the within-trial scores were used to shape the relative benefit of left- and right-hand responses. Indeed, in some trials the scores were doubled ($2 \times k/\mathrm{RT}^2$). The proportion of doubled scores was manipulated differently for the left- and right-hand responses to produce two distinct experimental conditions. In a first condition, called the "reward$_{neutral}$" condition, the score was doubled in half of the trials, equivalently for the left- and right-hand responses. Hence, in this condition, our scoring procedure introduced some variability in the feedback value following each response, but it did not modify the relative benefit of left- and right-hand responses. The only element that could potentially induce a discrepancy between the gain of left- and right-hand responses in this condition was an actual intermanual difference in RTs. Given that RTs are often shorter for the dominant hand and that all subjects were right handed in the present study, we predicted that right-hand responses could be associated with slightly larger scores than left-hand responses. As a consequence, in this condition, we expected subjects to show a small preference, if any, for right-hand responses to accumulate a larger financial bonus.

In contrast, in a second condition, called the "reward$_{biased}$" condition, the proportion of trials in which the score was doubled was unequal for left- and right-hand responses. More precisely, this proportion was much larger for left-hand responses (four of five trials) than for right-hand responses (one of five trials). This means that in the reward$_{biased}$ condition, subjects had a greater chance of getting a large score when responding with their left hand than when choosing to respond with their right hand. We predicted that this asymmetrical scoring procedure would lead to a larger inclination for left-hand responses in the reward$_{biased}$ condition compared to the reward$_{neutral}$ condition, especially in the indistinct and weak color range. Moreover, we expected that our manipulation would induce a shortening of left-hand RTs as a result of processes boosting activation of most "valuable" response representations during the competition process. Note that to

←

C, Sequence of events and TMS stimulation timings. A single TMS pulse was applied over the right primary motor cortex (M1) at six possible timings (TMS$_{BASELINE}$, TMS$_{IMP}$, TMS$_{MVT-PREP1}$, TMS$_{MVT-PREP2}$, TMS$_{MVT-PREP3}$, and TMS$_{MVT-PREP4}$).
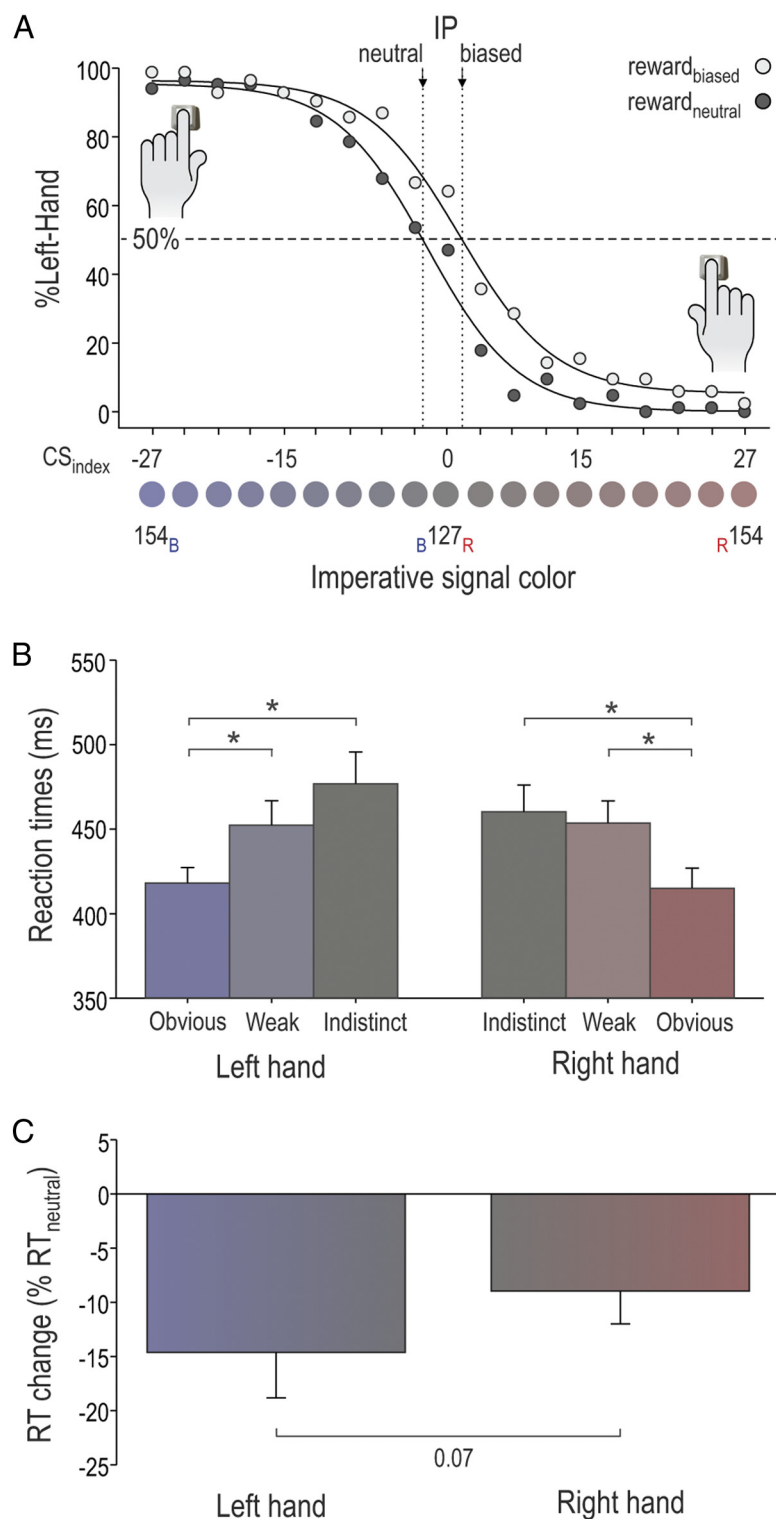
Figure 3. Illustration of inflection points (IPs) and RTs in the behavioral experiment. **A**, Percentage of left-hand responses for each of the 19 color saturation levels in the reward_neutral condition (full dots) and the reward_biased condition (open dots) of the behavioral experiment. Note the shift between the IP_neutral and the IP_biased, illustrating the behavioral effect of the reward bias. **B**, RTs, pooled across reward conditions, for both hands in each color category. **C**, Index reflecting RT change between the two conditions [(RT_biased − RT_neutral)/ RT_neutral] for the left- and right-hand responses. Note the more pronounced shortening of RTs in the left hand than in the right hand. *$p < 0.05$.

avoid "guessing" behaviors with an initiation of responses before the imperative signal, the maximum gain per trial was fixed at 60 points (corresponding to a minimum ceiling RT of 158 ms and 224 ms in the k/RT^2 and $2k/\mathrm{RT}^2$ trials, respectively).

One might wonder why we did not implement the left–right bias by simply controlling the value of the constant (k) in the computation of the feedback score (k/RT^2); we could have used a larger k value for left- than right-hand responses. However, we thought this strategy would be too explicit. The risk was that subjects, understanding the trick, would thoroughly choose their left hand and ignore their right one, especially in the indistinct "random" trials. Instead, we believe that our procedure allowed us to introduce a reward bias while maintaining a reasonable competition between the alternatives. Consistently, if all subjects noticed that the scores were not always coherent, none of them became aware of the actual manipulation. Most of them reported being surprised to respond faster (as they got higher scores) with their left than their right hand. It is of note that the computations of negative scores obtained following incorrect responses were similar in the reward_neutral and reward_biased blocks. Finally, the choice to favor the left rather than the right hand in the reward_biased condition was made based on a pilot behavioral experiment in which we tested both versions of the task and found a much more accentuated biasing effect when it favored the left compared to the right hand. This is likely because right-handed subjects already show a right-hand preference in the reward_neutral condition, making it difficult to further enhance this preference.

Behavioral experiment
Experiment goals. The behavioral experiment aimed at comparing the subjects' performance in the reward_neutral and reward_biased conditions. As mentioned above, we expected that making the left-hand responses more beneficial in the reward_biased blocks would boost processes promoting the selection of the left-hand response. To test this hypothesis, we measured two behavioral parameters: (1) the percentage of left-hand responses (%Left-Hand) and (2) the left- and right-hand RTs, according to the spectrum of blue and red color tints (indistinct, weak, or obvious).

As expected, the %Left-Hand data followed a sigmoid pattern, switching from maximum values (~100%) following the most bluish signals to minimum values (~0%) following the most reddish signals (Fig. 3). To evaluate precisely the %Left-Hand according to the spectrum of blue and red color tints, we computed a color saturation index by subtracting the baseline saturation value (127) from the saturation of the relevant channel (blue or red). Negative values were arbitrarily assigned to saturation units above baseline on the blue channel (−1 to −27, indicating increasingly blue tints; Fig. 3A, left), whereas positive values were assigned to saturation units above baseline on the red channel (+1 to +27, indicating increasingly red tints; Fig. 3A, right). Then, a logistic regression that minimized the difference between the actual %Left-Hand data and the best prediction allowed us to fit separate curves for the reward_neutral and reward_biased conditions using the following equation:

Table 1. TMS$_{MVT-PREP}$ timings and mean RTs in the TMS experiment

Subject	Reward$_{neutral}$					Reward$_{biased}$				
	TMS$_{MVT-PREP1}$	TMS$_{MVT-PREP2}$	TMS$_{MVT-PREP3}$	TMS$_{MVT-PREP4}$	RT	TMS$_{MVT-PREP1}$	TMS$_{MVT-PREP2}$	TMS$_{MVT-PREP3}$	TMS$_{MVT-PREP4}$	RT
1	49	98	147	196	445	44	88	132	176	400
2	48	96	144	192	436	48	96	144	192	436
3	53	106	159	212	482	49	98	147	196	445
4	44	88	132	176	400	40	80	120	160	364
5	56	112	168	224	509	63	126	189	252	573
6	57	114	171	228	518	54	108	162	216	491
7	50	100	150	200	455	50	100	150	200	455
8	46	92	138	184	418	42	84	126	168	382
9	53	106	159	212	482	49	98	147	196	445
10	47	94	141	188	427	43	86	129	172	391
11	47	94	141	188	427	35	70	105	140	318
Mean	50	100	150	200	455	47	94	141	188	427
SE	1.3	2.5	3.8	5.1	11.6	2.3	4.5	6.8	9.1	20.6

The four TMS timings, TMS$_{MVT-PREP1}$, TMS$_{MVT-PREP2}$, TMS$_{MVT-PREP3}$, and TMS$_{MVT-PREP4}$, were determined on an individual basis (0.17, 0.33, 0.50, and 0.66 × 66% of RT) to assess corticospinal excitability during the preparation period following the imperative signal. Left- and right-hand RTs were pooled together to determine the TMS$_{MVT-PREP}$ timings.

$$f(x) = y_0 + \frac{a}{1 + e^{-\left(\frac{x - x_o}{b}\right)}}$$

To compare the %Left-Hand in the reward$_{neutral}$ and reward$_{biased}$ conditions, we computed the inflection point (IP; x_0 parameter in the equation above) of the sigmoid curves obtained for each subject in each condition (IP$_{neutral}$ and IP$_{biased}$, respectively). As shown in Figure 3A, this IP reflects the color saturation index at which the response switched from being mostly performed with the left hand (%Left-Hand, >50%; left) to mostly performed with the right hand (%Left-Hand, <50%; right). Hence, an IP value of 0 means that the point of switch (50%) occurs at the exact midpoint of the color spectrum; an IP of <0 means that the switch occurs on the blue side of the color spectrum, reflecting a preference for right-hand responses, whereas an IP of >0 reflects a preference for left-hand responses (switch on the right side of the color spectrum). As such, here we expected IP$_{biased}$ values to be larger than IP$_{neutral}$ values, indicating a larger preference for left-hand responses in the reward$_{biased}$ than in the reward$_{neutral}$ condition.

Blocks and sessions. The behavioral experiment involved a total of 10 blocks of 80 trials. First, subjects performed three blocks in the reward$_{neutral}$ condition so as to assess their baseline behavior in a nonbiased reward context. Then, they performed seven blocks in the reward$_{biased}$ condition; the three last blocks were used to assess performance in the reward$_{biased}$ condition. For the analysis, we compared the %Left-Hand and the RTs in the three last reward$_{biased}$ blocks with respect to the three first reward$_{neutral}$ blocks. Of course, as already mentioned, subjects were never informed about the different types of blocks they were engaged in. Each block lasted ~7–8 min. A 5 min break was made every other block.

Statistical analyses. The IP$_{neutral}$ and IP$_{biased}$ of the %Left-Hand data were calculated for each subject and then compared between the reward$_{neutral}$ and reward$_{biased}$ conditions by means of a paired t test. A three-way repeated-measures ANOVA (ANOVA$_{RM}$) was then conducted on RTs with condition (neutral, biased), hand (left, right), and color saturation (indistinct, weak, and obvious) as factors. After that, we computed an index reflecting the impact of the reward bias on the RTs (RT$_{biased}$ − RT$_{neutral}$/RT$_{neutral}$), which we analyzed using a two-way ANOVA$_{RM}$ with factors hand (left, right) and color saturation (indistinct, weak, and obvious). *Post hoc* comparisons were conducted using Fisher's least significant difference (LSD) procedure.

TMS experiment
Experiment goals. An important finding in the behavioral experiment is that subjects showed an augmented preference for left-hand responses in the reward$_{biased}$ compared to the reward$_{neutral}$ blocks. This effect was manifest both when considering the RTs and the %Left-Hand data (see behavioral results).

The aim of the TMS experiment was to investigate whether this increased left-hand preference in the reward$_{biased}$ condition could be ac-

counted for by a specific change in the activation of its motor representation during response preparation. To test this idea, we applied single pulse TMS over the right M1 to measure the corticospinal (CS) excitability of the left FDI muscle in the same task as the one used in the behavioral experiment.

TMS procedure. A figure-of-eight coil (wing external diameter, 70 mm) connected to a Magstim 200 magnetic stimulator was placed tangentially on the scalp; the handle was oriented toward the back of the head and laterally at a 45° angle away from the midline, approximately perpendicular to the central sulcus. We identified the optimal spot for eliciting MEPs in the left FDI muscle, and this location was marked on an electroencephalography cap fitted on the participant's head to provide a reference point throughout the experimental session. The resting motor threshold (rMT) was defined as the minimal TMS intensity required to evoke MEPs of ~50 μV peak to peak in the targeted muscle in 5 of 10 consecutive trials. Across participants, the rMT corresponded to 43 ± 2.4% ($n = 11$) of the maximum stimulator output. The intensity of TMS for the experimental sessions was always set at 120% of the individual rMT.

To assess CS excitability of the left FDI muscle during response preparation, we applied TMS at six different timings (Fig. 2C), although only one single pulse TMS was delivered in each trial. First, to establish a baseline of the CS excitability, TMS pulses were applied during the intertrial interval. More specifically, this timing of stimulation, referred to as TMS$_{BASELINE}$, fell 200 to 600 ms before the onset of the fixation cross. Second, TMS pulses could also occur at the onset of the imperative signal (TMS$_{IMP}$) or at one out of four timings during movement preparation (MVT-PREP), that is, between the onset of the imperative signal and the motor response. These four timings, referred to as TMS$_{MVT-PREP1}$, TMS$_{MVT-PREP2}$, TMS$_{MVT-PREP3}$, and TMS$_{MVT-PREP4}$, were determined on an individual basis and corresponded, respectively, to 0.17, 0.33, 0.50, and 0.67 × 66% of the individual median RT (Table 1). This RT was measured at the beginning of each session in the last no-TMS training block (see Blocks and sessions, below) and corresponded to the time elapsed between the onset of the imperative signal and the detection of the button press. The value of 66% of the RT was chosen based on pilot experiments, which showed that this value corresponds approximately to the onset of the FDI muscle EMG activity preceding the button press. This was confirmed in four subjects of the current experiment in whom we compared the time of EMG onset with the RT recorded from the keyboard; we found that the EMG onset occurred on average at 65.9 ± 4.1% of the RT. Finally, the timing of the last TMS pulse (0.67 × 66%) was also chosen based on pilot experiments. It corresponded to the latest time at which we can elicit MEPs without taking the risk of having too many TMS pulses falling after the EMG onset given the RT variability. This was an important point because assessing CS excitability changes during response preparation requires including only the trials in which the TMS falls before EMG onset (Chen and Hallett, 1999); trials in which the TMS pulse falls after EMG onset need to be removed from the data set. A previous study using a similar task also showed that most selection-related CS excitability

changes occur before that time (Michelet et al., 2010). Hence, we believe that the four $TMS_{MVT-PREP}$ timings (1–4) provide us with a valuable probe of CS excitability changes during response preparation (Rossini et al., 1988; Reynolds and Ashby, 1999; Leocani et al., 2000) with only a marginal amount of data loss due to variations in the EMG onset timing. For the analysis of the MEP data, we focused on three measures: TMS_{IMP}, which consisted of MEPs elicited when TMS was applied at the imperative signal onset; $TMS_{IMP+100}$, which consisted of MEPs elicited with TMS at 0.33 × 66% of the mean RT (corresponding to ∼100 ms after the imperative signal; Table 1); and $TMS_{MVT-175}$, in which we pooled together all MEPs that were elicited by TMS falling from 175 to 0 ms preceding the expected time of EMG onset [66% of the trial RT (time to key press)]. Following this procedure, a minimum of 10 MEPs remained to estimate CS excitability at $TMS_{MVT-175}$ in each condition. On average, these MEPs were elicited at a comparable delay from the estimated time of EMG onset in the two reward conditions ($reward_{neutral}$, −118 ± 36 ms; $reward_{biased}$, −119 ± 35 ms; $t_{(10)} =$ 0.10, $p > 0.91$).

Together, these measures allowed us to assess MEPs at different time points during movement preparation: at the onset of the imperative signal (TMS_{IMP}), ∼100 ms later ($TMS_{IMP+100}$), and ∼100 ms before movement onset ($TMS_{MVT-175}$). Notably, we choose not to include the $TMS_{MVT-PREP1}$ timing (0.17 × 66% of the mean RT, i.e., ∼50 ms after TMS_{IMP}) in our final analysis because the amplitude of MEPs elicited at this time point was similar to that of MEPs elicited at the signal onset (all $F < 2.64$; all $p > 0.136$).

Finally, in four subjects, we also assessed CS excitability outside the task by applying 20 TMS pulses between the blocks ($TMS_{BASELINE-OUT}$). This control was performed to check for the occurrence of a general CS excitability change in the $reward_{biased}$ condition that would extend to baseline MEPs.

Blocks and sessions. The TMS experiment extended over two sessions (one for each experimental condition) performed on different days; the order of the $reward_{neutral}$ and $reward_{biased}$ sessions was counterbalanced between subjects. Each session involved eight blocks of 84 trials (∼7 min each). The first three blocks were performed in the absence of TMS and served to train the subjects in a specific feedback context ($reward_{neutral}$ or $reward_{biased}$); the median RT was then extracted from performance in the third block to establish the individual $TMS_{MVT-PREP}$ timings (1–4) within that session. Then, in the main phase of the experiment, subjects performed five TMS blocks during which we elicited MEPs at the different timings of interest. This procedure allowed us to assess the amplitude of left FDI muscle MEPs when the muscle was either selected (left-hand response) or nonselected (right-hand response), following an obvious, weak or indistinct imperative signal, in the $reward_{neutral}$ and $reward_{biased}$ conditions. Twenty MEPs were elicited to assess CS excitability in each condition. There was a 5 min break every other block or whenever the subjects felt the need to rest.

EMG recordings. EMG activity was recorded from surface electrodes (Neuroline; Medicotest) placed over the left and right FDI muscles. EMG data were collected for 2600 ms on each trial, starting at least 200 ms before the TMS pulse. The EMG signals were amplified and bandpass filtered on-line (10–500 Hz; NeuroLog; Digitimer), and digitized at 2000 Hz for off-line analysis. The EMG signals were used to measure peak-to-peak amplitudes of the left FDI muscle MEPs. Importantly, a careful inspection of each single trial allowed us to identify, and then to exclude from the analysis, trials in which the TMS pulse fell after the EMG onset or trials with any background EMG activity larger than 100 μV in the 200 ms window preceding the TMS pulse. This technique has proved useful in the past to prevent contamination of the MEP measurements by significant fluctuations in background EMG (Duque et al., 2005; Sartori et al., 2011; Cavallo et al., 2012). Finally, trials in which subjects pushed the wrong button were also removed from the data set. After trimming the data for errors, background EMG activity, and outliers, a minimum of 10 MEPs remained to assess CS excitability in each condition.

Statistical analyses. To analyze behavior in the TMS experiment, we pooled together all trials, regardless of the TMS timings. We chose to do so to increase the number of observations. We recognize that the TMS stimulation is likely to impact behavior (Rossini, 1988; Hallett, 2007).

However, we assumed that this TMS effect would be similar in the $reward_{neutral}$ and $reward_{biased}$ conditions, and thus that this procedure would not preclude us from making valid comparisons between the two experimental conditions. To run analyses with a sufficient amount of MEPs in each data set, the indistinct and weak color categories were pooled together (into an "ambiguous" category). All analyses were similar to those performed in the behavioral experiment. Briefly, the $IP_{neutral}$ and IP_{biased} of the %Left-Hand data were compared using a paired t test. RTs were analyzed using a three-way $ANOVA_{RM}$ with condition (neutral, biased), hand (left, right), and color saturation (ambiguous, obvious) as factors. Finally, as for the behavioral experiment, an index of the bias effect on RTs ($RT_{biased} - RT_{neutral}/RT_{neutral}$) was computed and analyzed using a two-way $ANOVA_{RM}$ with factors hand (left, right) and color saturation (ambiguous, obvious).

Left FDI muscle MEPs elicited during the preparation period (TMS_{IMP}, $TMS_{IMP+100}$, and $TMS_{MVT-175}$) were always expressed with respect to MEPs (as a percentage) elicited at $TMS_{BASELINE}$ within the same session (either $reward_{neutral}$ or $reward_{biased}$). Hence, MEPs at $TMS_{BASELINE}$ were used as reference for the subsequent TMS timings within each session, cancelling out any possible block effect in MEPs measured during movement preparation (at TMS_{IMP}, $TMS_{IMP+100}$, or $TMS_{MVT-175}$). As these MEPs (%$TMS_{BASELINE}$) were not normally distributed (Kolmogorov–Smirnov test; failed, $p < 0.01$), a logarithmic transformation was applied before the statistical tests to obtain a normal distribution (Kolmogorov–Smirnov test; passed, $p > 0.20$). A four-way $ANOVA_{RM}$ was run on these log-transformed MEPs with the factors epoch (TMS_{IMP}, $TMS_{IMP+100}$, and $TMS_{MVT-175}$), condition (neutral, biased), hand (left, right), and color saturation (ambiguous, obvious). Then, we assessed the effect of the left bias by expressing each log-transformed MEP measure (TMS_{IMP}, $TMS_{IMP+100}$, and $TMS_{MVT-175}$) in the $reward_{biased}$ condition with respect to that in the $reward_{neutral}$ condition [%Reward-Effect ($^{log}MEP_{biased} - ^{log}MEP_{neutral}/^{log}MEP_{neutral}$)]. We analyzed this %Reward-Effect by means of an $ANOVA_{RM}$ with epoch (TMS_{IMP}, $TMS_{IMP+100}$, and $TMS_{MVT-175}$), hand (left, right) and color saturation (ambiguous, obvious) as factors. This analysis allowed us to compare the strength of the %Reward-Effect at the different epochs and thus to specifically test our two main hypotheses (see Introduction; Fig. 1A).

Finally, in an additional analysis, we expressed MEPs elicited at $TMS_{IMP+100}$ and $TMS_{MVT-175}$ with respect to MEPs elicited at TMS_{IMP}. This was done to eliminate the starting-point effect (at TMS_{IMP}; see Results) in our MEP measures taken at the subsequent time points ($TMS_{IMP+100}$ and $TMS_{MVT-175}$). These values were used (after having been log-transformed) to compute a %Reward-Effect (see above) specific to the preparation period (%Reward-Effect$_{PREP}$), which was analyzed using a three-way $ANOVA_{RM}$ with factors epoch ($TMS_{IMP+100}$ and $TMS_{MVT-175}$), hand (left, right), and color saturation (ambiguous, obvious). All *post hoc* comparisons were conducted using Fisher's LSD procedure. All of the data are expressed as mean ± SE. In addition, figures display nontransformed data to facilitate visualization of our data.

Control experiment
Left MEP measures in the TMS experiment revealed specific differences in CS excitability changes during movement preparation in the $reward_{biased}$ and $reward_{neutral}$ conditions (see Results). One possibility is that these effects were due to different reward-based, top-down influences over the left response motor representation in the two conditions, consistent with the idea that this cognitive variable may bias motor competition during response preparation. However, another explanation is that these differences evidenced in the main TMS experiment were related to the fact that, as a consequence of the reward bias, subjects prepared their left hand more in the $reward_{biased}$ compared to the $reward_{neutral}$ condition to generate faster responses and to receive the largest reward. In that case, contrary to the first hypothesis, the changes in motor excitability found in the $reward_{biased}$ condition would only be an indirect consequence of reward, through enhanced preparatory processes favoring left-hand responses. A control experiment was done to test the viability of the latter hypothesis. In particular, we aimed at assess-

ing whether the occurrence of a preparation bias could explain our data in the main TMS experiment.

To do so, we tested 10 additional subjects in two experimental conditions (performed on the same day in a counterbalanced order) that differed according to the proportion of trials that required left- or right-hand responses. In one condition (major$_{left}$), the imperative signal (weak and obvious color conditions) called for left-hand responses on 75% of trials, whereas in the other condition (major$_{right}$), the imperative signal only called for left-hand responses on 25% of trials (75% right-hand responses). Subjects were aware of these different proportions in the two conditions and were asked to favor preparation of the most likely response, to be as fast as possible. Hence, we assumed that preparation of left-hand responses would be larger in the major$_{left}$ condition compared to the major$_{right}$ condition, as reflected in the RT data (see Results). Importantly, reward was always constant and similar across conditions in this experiment (correct, +10; error, −5), to avoid any potential direct influence of reward on CS excitability. All other aspects of the tasks were the same as in the main experiment, including the required response to bluish and reddish circles (left and right, respectively), the range of color saturations (indistinct, weak, and bright), as well as the TMS measures (TMS$_{BASELINE}$, TMS$_{IMP}$, TMS$_{IMP+100}$, and TMS$_{MVT-175}$). All analyses were similar, including the log-transformation of MEPs before statistical tests, as in the main experiment.

Results

Behavioral experiment

Figure 3A shows the %Left-Hand responses according to the color of the imperative signal in the reward$_{neutral}$ (filled dots) and reward$_{biased}$ (open dots) conditions as well as the two fitted sigmoid curves. In both conditions, the nonlinear regressions explained ~99% of the observed data variance. The IP$_{neutral}$ and IP$_{biased}$ equaled −2.40 ± 1.28 and +1.10 ± 1.78, respectively. Importantly, these two values were significantly different ($t_{(7)}$ = 2.4, $p < 0.047$), suggesting that individuals switched from preferring their right hand in the reward$_{neutral}$ condition (IP$_{neutral}$, <0) to preferring their left hand in the reward$_{biased}$ condition (IP$_{biased}$, >0). This effect was even more obvious when considering the indistinct color condition alone; whereas subjects only choose their left hand in 42.1% of trials in the reward$_{neutral}$ condition, they used it in 58.2% of trials in the reward$_{biased}$ condition ($t_{(10)}$ = 3.36, $p < 0.012$).

The mean RTs for left-index-finger responses were 488 ± 24 ms ($n = 8$) and 411 ± 14 ms in the reward$_{neutral}$ (pretraining) and reward$_{biased}$ (posttraining) conditions, respectively; the mean RTs for right-index-finger responses were 466 ± 23 ms and 420 ± 11 ms, respectively. The ANOVA$_{RM}$ revealed a main effect of condition ($F_{(1,7)}$ = 12.53, $p < 0.009$) and of color saturation ($F_{(2,14)}$ = 14.47, $p < 0.001$) on RTs. Given the sequence of blocks (reward$_{biased}$ blocks always came after reward$_{neutral}$ blocks), RTs were generally smaller in the reward$_{biased}$ (415 ± 14 ms) than in the reward$_{neutral}$ condition (477 ± 23 ms). In addition, key presses following obvious circles (417 ± 17 ms) were faster than those following weak (453 ± 21 ms, $p < 0.001$) or indistinct circles (469 ± 25 ms, $p < 0.001$), regardless of the responding hand (Fig. 3B).

More importantly, a condition by hand interaction ($F_{(7,11)}$ = 5.09, $p < 0.059$) was nearly significant. Consistently, *post hoc* tests revealed that RTs tended to be slower in the left than in the right hand in the pretraining reward$_{neutral}$ condition ($p < 0.060$), but became similar following training in the reward$_{biased}$ condition ($p > 0.373$). This suggests that RTs improved more from pretraining to posttraining in the left compared to the right hand. Similarly, ANOVA$_{RM}$ revealed a nearly significant effect of the factor hand ($F_{(1,7)}$ = 4.71, $p < 0.067$), regardless of the color

saturation ($F_{(2,14)}$ = 0.46, $p > 0.642$; Fig. 3C), on the index computed to assess RT change [(RT$_{biased}$ − RT$_{neutral}$)/RT$_{neutral}$].

Together, these findings support the view that the gain associated with each response is evaluated to build up a prediction of the reward one can expect from future responses; this information is used to direct our choices toward the most satisfying actions.
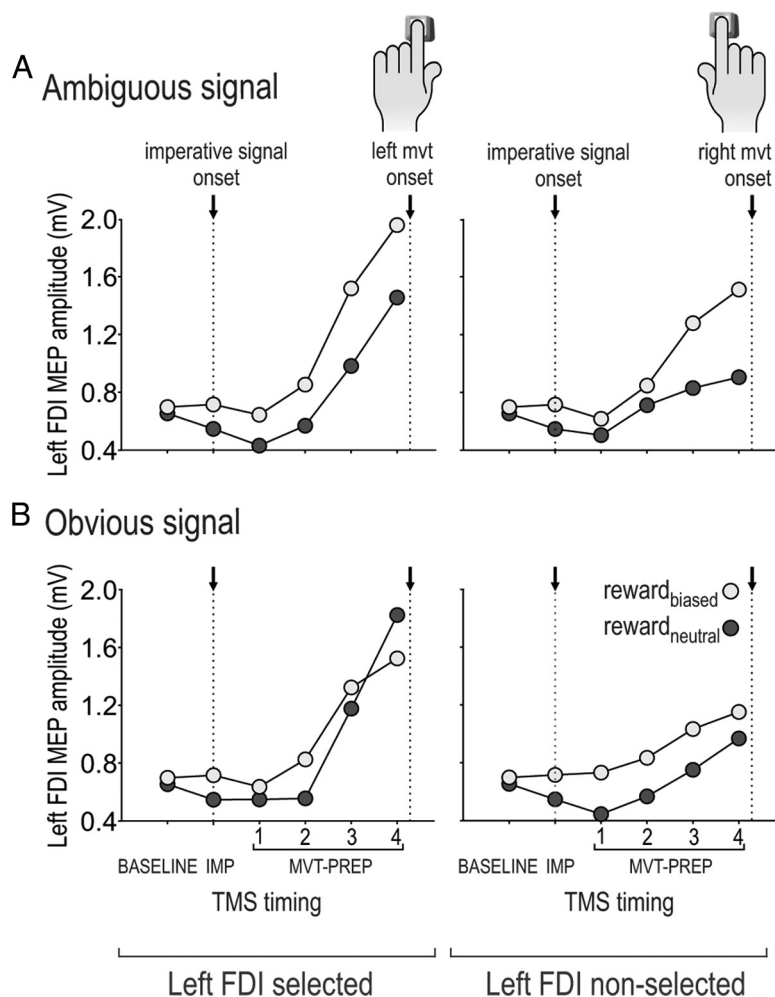
TMS experiment

The IP$_{neutral}$ and IP$_{biased}$ equaled +0.33 ± 0.82 and +1.79 ± 0.58, respectively. These two values were significantly different ($t_{(10)}$ = 3.2, $p < 0.012$), confirming an impact of the left bias on the %Left-Hand.

The mean RTs for left-index-finger responses were 488 ± 18 ms ($n = 11$) and 460 ± 17 ms in the reward$_{neutral}$ and reward$_{biased}$ conditions, respectively; the mean RTs for right-index-finger responses were 450 ± 11 and 453 ± 14 ms, respectively. As in the behavioral experiment, the effect of color saturation on RTs was significant ($F_{(1,10)}$ = 4.91, $p < 0.019$), as well as the condition by hand interaction ($F_{(1,10)}$ = 5.67, $p < 0.038$). Left RTs ($p < 0.013$), but not right RTs ($p > 0.717$), were found to be shorter in the reward$_{biased}$ than in the reward$_{neutral}$ condition, an effect also reflected in the RT change due to the reward bias (main effect of hand, $F_{(1,10)}$ = 6.21, $p < 0.032$).

Factorial ANOVAs were used to compare directly the performance in the behavioral and TMS experiments. These tests revealed a significant difference in IPs between the two experiments (main effect of experiment, $F_{(1,34)}$ = 4.88, $p < 0.035$). Subjects in the behavioral experiment used, in general, less their left hand (mean IP, −0.65 ± 1.53) than the subjects recruited in the TMS experiment (mean IP, 1.06 ± 0.70). This difference could be related to the recruitment of different subjects in the two experiments. Another possibility is that MEPs, which were always elicited in the left hand, tended to increase the selection of that hand in the TMS experiment. However, more importantly, we found a significant main effect of factor condition ($F_{(1,34)}$ = 5.32, $p < 0.027$) and no significant experiment by condition interaction ($F_{(1,34)}$ = 0.95, $p > 0.336$). This means that despite the difference highlighted above, the reward manipulation influenced hand preference in a similar way in the two experiments. Concerning RTs, a significant experiment by condition interaction was found ($F_{(1,204)}$ = 13.03, $p < 0.001$). Given the design, RTs in the behavioral experiment were globally much longer in the neutral (pretraining) than in the biased (posttraining) condition (see numbers above; $p < 0.001$). Such a global effect of condition was not found for the TMS experiment ($p > 0.179$). This is due to the fact that in that case, the biased and neutral conditions were both tested after training. In addition, as reflected in the numbers (see previous sections), RTs in the TMS experiment (posttraining sessions) were generally slower than RTs in the posttraining session of the behavioral experiment (both hands, all $p < 0.014$). Again, this is possibly due to the fact that subjects were different in the two experiments; alternatively, TMS may have delayed RTs. Consistent with this idea, RTs in the TMS experiment were found to be slower when TMS was applied at the latest time points (at TMS$_{MVT-PREP3}$ and TMS$_{MVT-PREP4}$, 486 ± 19 ms) compared to when it was applied at the earliest time points (at TMS$_{BASELINE}$ and TMS$_{IMP}$, 460 ± 19 ms; $t_{(10)}$ = −2.63, $p < 0.026$). Importantly, despite the differences mentioned above, both experiments revealed a specific speeding of RTs for left-hand responses compared to right-hand responses; this confirmed the effectiveness of the task manipulation.

Figure 4. Illustration of MEP results (TMS experiment) in a representative subject. The amplitude of raw left FDI muscle MEPs elicited at all TMS timings (TMS$_{BASELINE}$, TMS$_{IMP}$, and TMS$_{MVT-PREP1}$–TMS$_{MVT-PREP4}$) preceding left (left hand selected) and right (left hand not selected) key presses. Note the systematically larger MEPs in the reward$_{biased}$ compared to the reward$_{neutral}$ condition. ***A***, Following ambiguous imperative signals. ***B***, Following obvious imperative signals.
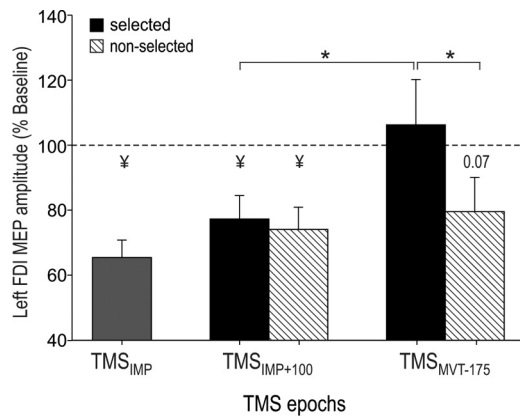
not modulated in a distinctive way with respect to MEPs elicited outside the blocks, at TMS$_{BASELINE-OUT}$. Hence, the larger reward associated with left-hand responses in the reward$_{biased}$ condition did not impact on baseline motor excitability of the left FDI muscle. Rather, it specifically boosted activity of left FDI muscle at a time when it was competing for selection.

Figure 4AB displays raw MEP amplitudes (in millivolts) in a representative subject; left FDI muscle MEPs are shown for the TMS$_{BASELINE}$, TMS$_{IMP}$, and each of the TMS$_{MVT-PREP}$ timings (1–4) preceding left (left FDI muscle selected) and right (left FDI muscle nonselected) key presses for this single subject. Note the systematically larger MEPs at TMS$_{IMP}$ in the reward$_{biased}$ compared to the reward$_{neutral}$ condition (in this individual subject, $t_{(45)} = 3.61$, $p < 0.001$). As such, in this subject, MEPs were suppressed in the reward$_{neutral}$ condition ($t_{(45)} = 3.52$, $p < 0.001$, when compared to baseline in this individual subject) but not in the reward$_{biased}$ condition (although it was to some degree at the group level, $p < 0.0001$; see below). In addition, the effect of the reward bias increased at the latest TMS time point (TMS$_{MVT-PREP4}$) in the ambiguous color condition, but not in the obvious situation ($t_{(70)} = 2.70, p < 0.008$ and $t_{(42)} = 0.47, p > 0.630$, respectively; Fig. 4$A$). This pattern of CS excitability changes was confirmed by the analyses performed at the group level.

For the group analysis of CS excitability changes during the preparation period, MEPs elicited at TMS$_{IMP}$, TMS$_{IMP+100}$, and TMS$_{MVT-175}$ were all expressed with respect to MEPs at TMS$_{BASELINE}$ (and then log-transformed for statistical analyses). Consistent with previous studies, we found that MEPs were strongly suppressed at the onset of the imperative signal (TMS$_{IMP}$; $t_{(10)} = -5.23, p < 0.001$), reaching, on average, $66 \pm 5.7\%$ of the baseline value (Fig. 5). In the past, we have referred to this form of inhibition as "impulse control," reflecting the idea that it prevents the premature initiation of potential responses (Duque and Ivry, 2009; Duque et al., 2012). Then, a four-way ANOVA$_{RM}$ revealed a significant epoch by hand interaction on left FDI muscle MEPs elicited during the preparation period ($F_{(2,20)} = 10.02, p < 0.001$; Fig. 5). The amplitude of left MEPs progressively increased when the imperative signal indicated a left-hand response, but remained largely suppressed when the imperative signal indicated a right key press. As a result, MEPs at TMS$_{MVT-175}$ were significantly larger in the former (left selected) compared to the latter (left nonselected) condition ($p < 0.0001$).

More importantly, the epoch by condition by color saturation interaction was significant ($F_{(2,20)} = 4.82, p < 0.019$; Fig. 6$A$). A first notable result is the significant upregulation of left MEP amplitudes at TMS$_{IMP}$ in the reward$_{biased}$ condition compared to the reward$_{neutral}$ condition ($p < 0.001$). This means that CS excitability associated with a left response was larger at the onset of the decision process when this response was associated with a

The mean amplitudes of left FDI muscle MEPs applied during the intertrial intervals (at TMS$_{BASELINE}$) of the reward$_{neutral}$ and reward$_{biased}$ blocks were 2.27 ± 0.45 mV ($n = 11$) and 2.24 ± 0.39 mV, respectively. These baseline values were highly comparable in the two block types ($t_{(10)} = 0.11, p > 0.916$). However, because the two conditions were tested on different days, the direct comparison of MEPs elicited at TMS$_{BASELINE}$ might be biased by several uncontrolled aspects that changed between sessions (e.g., position of the electrodes, level of alertness of the subjects, location of the TMS coil, etc.), precluding us from observing any global block effect in the present study. Hence, to test for the occurrence of such a change in CS excitability at a block level, we compared, in four subjects, the MEPs obtained at TMS$_{BASELINE}$ to those elicited at TMS$_{BASELINE-OUT}$ (outside the block) by means of individual two-way ANOVAs with the factors condition (neutral, biased) and epoch (TMS$_{BASELINE}$, TMS$_{BASELINE-OUT}$). In one subject (out of four tested), we found a significant main-effect of TMS-epoch (TMS$_{BASELINE}$ vs TMS$_{BASELINE-OUT}$; $p < 0.0001$); that is, MEPs were larger in the block compared to outside the block, a finding that has been reported previously in the literature (Labruna et al., 2011). More importantly, the condition by epoch interaction was never significant (all four subjects, $F_{(1,2)} < 2.01$; all $p > 0.16$). This means that MEPs elicited at TMS$_{BASELINE}$ were

Figure 5. Illustration of MEP changes when left FDI muscle was selected or not selected in the TMS experiment ($n = 11$). The amplitudes of MEPs, expressed as percentages of baseline, are depicted for the different TMS epochs (see Materials and Methods) when the left FDI muscle was selected (black histograms) or not selected (hatched histograms). MEPs are pooled across reward conditions. In addition, MEPs are pooled across hands at TMS_{IMP} given that, at that time, participants don't know yet which hand they will select. MEPs are shown as raw data but were log-transformed for statistical analyses. *$p < 0.05$ (significantly different); ¥$p < 0.05$ (significantly different from MEPs elicited at $TMS_{BASELINE}$).

larger reward, an effect that meets the prediction of the starting-point shift hypothesis (Fig. 1A, left).

The upregulation of left MEPs in the reward$_{biased}$ condition was maintained at $TMS_{IMP+100}$ ($p < 0.002$) and $TMS_{MVT-175}$ ($p < 0.008$, Fig. 6A). Yet, the magnitude of this upregulation depended on the color saturation of the imperative signal (ambiguous, obvious), especially at $TMS_{MVT-175}$. At this last epoch, MEPs were larger following obvious than ambiguous colors in the reward$_{neutral}$ condition ($p < 0.022$), an effect consistent with models of decision making assuming that the degree of perceptual evidence can influence the accumulation process (Gold and Shadlen, 2007). Interestingly, such an effect was not found in the reward$_{biased}$ condition ($p > 0.061$), probably because the reward bias had a different impact on CS excitability in the two color categories. To investigate this point, we assessed the %Reward-Effect computed for each epoch (TMS_{IMP}, $TMS_{IMP+100}$, and $TMS_{MVT-175}$) in the ambiguous and obvious color conditions (Fig. 6B, gray and black circles, respectively). The ANOVA$_{RM}$ performed on these data revealed a significant interaction between the epoch and color-saturation factors ($F_{(2,20)} = 5.78$, $p < 0.010$). At the early stages of the preparation period, the bias effect averaged $20.32 \pm 8.94\%$ (at TMS_{IMP}) and then $28.97 \pm 8.37\%$ (at $TMS_{IMP+100}$); these values were found to be similar following ambiguous and obvious imperative signals (all $p > 0.47$). Then, at timings closer to movement onset (at $TMS_{MVT-175}$), the bias effect increased for the ambiguous color category ($39.66 \pm 10.94\%$, $p < 0.027$, compared to bias at $TMS_{IMP+100}$), but decreased for the obvious color category ($14.89 \pm 9.54\%$, $p < 0.005$, compared to bias at $TMS_{IMP+100}$). Consequently, at $TMS_{MVT-175}$, the bias effect was significantly larger in the ambiguous than in the obvious condition ($p < 0.001$). The larger %Reward-Effect at $TMS_{MVT-175}$ in the ambiguous condition with respect to the previous timings is consistent with the occurrence of a slope shift in that condition (Fig. 1A, right).

Importantly, the stronger the %Reward-Effect at $TMS_{IMP+100}$ in the ambiguous color condition, the more the subjects switched from using their right hand to using their left hand in that same color condition ($r = 0.78$, $p < 0.006$; Fig. 6C). No such correlation was found at $TMS_{MVT-175}$, possibly because the increased

variability when considering MEP amplitudes closer to movement onset precluded us from identifying precise links between this neurophysiological measure and behavior.

Finally, a last analysis was run in which we expressed MEPs elicited at $TMS_{IMP+100}$ and $TMS_{MVT-175}$ with respect to MEPs elicited at TMS_{IMP}, thereby cancelling out the starting-point effect for the analysis of subsequent time points. These values were then used to compute a %Reward-Effect specific to the preparation period (%Reward-Effect$_{PREP}$ at $TMS_{IMP+100}$ and $TMS_{MVT-175}$; Fig. 6B, triangles). Following this procedure, ANOVA$_{RM}$ revealed a significant color saturation by epoch interaction ($F_{(1,10)} = 6.66$, $p < 0.027$) on the %Reward-Effect$_{PREP}$. Similar to the initial analysis, the %Reward-Effect$_{PREP}$ was found to be significantly larger following ambiguous compared to obvious signals at $TMS_{MVT-175}$ ($p < 0.011$), but not at $TMS_{IMP+100}$ ($p < 0.575$). This last analysis provides strong evidence for the occurrence of a reward-related slope shift on top of the starting-point shift.

Control experiment

In the control experiment ($n = 10$), we tested whether our findings in the TMS experiment could be accounted for by the occurrence of a preparation bias. Figure 7, A–D, shows the RT and MEP data for the major$_{left}$ and major$_{right}$ blocks of this control experiment (see Materials and Methods).

The inspection of RTs confirmed the occurrence of a preparation bias in the two conditions of the control experiment. ANOVA$_{RM}$ revealed a significant main effect of color saturation ($F_{(2,18)} = 14.53$, $p < 0.001$) and a significant hand by condition interaction ($F_{(1,9)} = 25.26$, $p < 0.001$), but no color saturation by condition interaction ($F = 0.96$, $p > 0.401$) on RTs. As expected, left responses were faster in the major$_{left}$ than in the major$_{right}$ blocks ($p < 0.027$; Fig. 7A), whereas the reversed effect was found for right-hand responses ($p < 0.002$; Fig. 7A, inset), confirming that subjects prepared the most frequent response more in this control experiment.

Figure 7, B–D, depicts left FDI muscle MEP results. Similar to the main experiment, all MEPs elicited during movement preparation were expressed with respect to those elicited at baseline (log-transformed for statistical analyses). Figure 7B illustrates the overall change in left MEPs when the left FDI muscle was selected or not selected. These results are similar to those found in the main experiment (Fig. 5); left MEPs were strongly suppressed at the onset of the imperative signal ($TMS_{IMP} = 86 \pm 3.8\%$ of the baseline value; $t_{(9)} = -3.66$, $p < 0.005$). They then increased when the imperative signal indicated a left-hand response but remained largely suppressed when the imperative signal indicated a right key press (significant epoch by hand interaction, $F_{(2,18)} = 4.69$, $p < 0.023$). Figure 7C shows the amplitude of left FDI muscle MEPs (%Baseline) at TMS_{IMP}, $TMS_{IMP+100}$, and $TMS_{MVT-175}$ in the major$_{left}$ and major$_{right}$ conditions for each color saturation (similar to Fig. 6A in the main experiment). Importantly, there was no significant difference in left MEP amplitudes between the two block types (condition, $F_{(1,9)} = 0.12$, $p > 0.741$), either in the ambiguous or the obvious color category (condition by color saturation interaction, $F_{(1,9)} = 2.80$, $p > 0.129$). At TMS_{IMP}, the average left MEPs even tended to be smaller in the condition in which the left hand was most prepared (major$_{left}$; but it was not significant, $p > 0.55$). In fact, this effect would be coherent with the idea that response preparation is associated with impulse control inhibition at the end of a delay period (Duque et al., 2012). Hence, the starting-point shift in the reward$_{biased}$ blocks of the main experiment cannot be accounted for by the larger preparation of left responses in that condition.
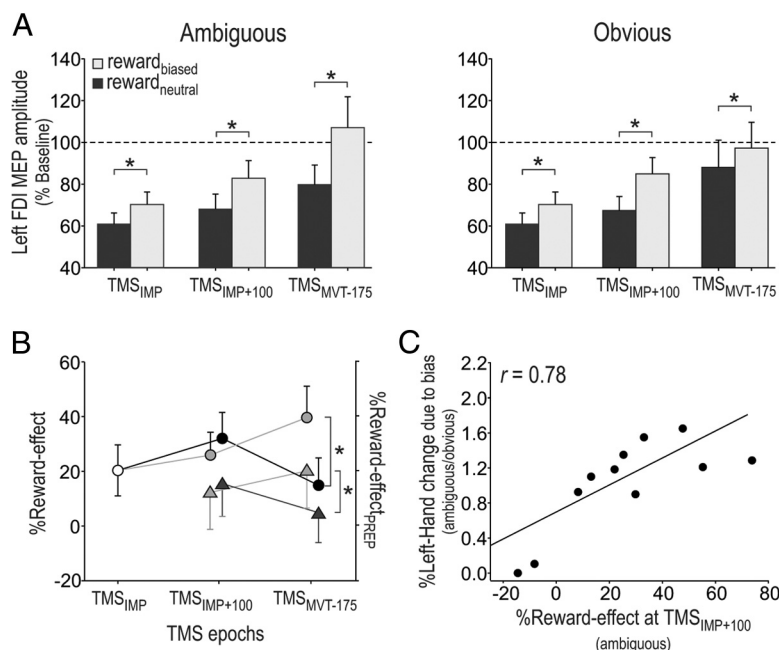
Figure 6. Illustration of the reward effect in the TMS experiment ($n = 11$). MEPs are shown as raw data but were log-transformed for statistical analyses. *A*, The amplitudes of left FDI muscle MEPs, expressed as percentages of baseline following ambiguous (left side) and obvious (right side) imperative signals in the reward$_{neutral}$ (dark histograms) and the reward$_{biased}$ (light histograms) conditions. Note the global upregulation of left FDI muscle MEP amplitudes in the reward$_{biased}$ condition (all $p <$ 0.002), which seems to be present from the starting point of the accumulation process (TMS$_{IMP}$). This upregulation was maintained at TMS$_{IMP+100}$ and TMS$_{MVT-175}$, even though at this late timing, its magnitude was influenced by the color-saturation category. *B*, The %Reward-Effect and the %Reward-Effect$_{PREP}$ (circles and triangles, respectively; see Materials and Methods) analyses revealed that at the early decision period (TMS$_{IMP+100}$), the effect was the same for both color-saturation categories (ambiguous, light gray; obvious, dark gray). Closer to movement onset (at TMS$_{MVT-175}$), the bias effect increased for the ambiguous color category but decreased for the obvious color category. As a result, the %Reward-Effect was significantly larger in the ambiguous than in the obvious condition ($p < 0.001$). *C*, The strength of the %Reward-Effect on motor excitability at TMS$_{IMP+100}$ in the ambiguous trials correlated with the degree to which the %Left-Hand changed due to the bias, also in ambiguous trials, compared with the change observed for obvious trials ($r = 0.78, p < 0.006$). *$p < 0.05$.

In addition, as shown in Figure 7, *C* and *D*, we did not find any slope shift between the TMS$_{+100}$ and TMS$_{MVT-175}$ timings in the ambiguous color category. In the ambiguous condition, left MEPs at TMS$_{+100}$ tended to be larger in the major$_{left}$ compared to major$_{right}$ conditions, possibly due to a faster initial CS excitability increase due to enhanced preparation (Fig. 7C). However, again this effect was not significant ($p > 0.22$). Furthermore, it disappeared at TMS$_{MVT-175}$, the timing at which we found the strongest %Reward-Effect in the main experiment. These differences are also evident when observing the %Preparation-Effect (Fig. 7D), homologous to the %Reward-Effect in the main experiment. Hence, it is very unlikely that our results in the main experiment can be accounted for by different levels of preparation in the two block types.

Together, findings in the main and control experiments of the present study indicate that the reward bias influenced CS excitability as soon as the imperative signal was presented. Then, this effect increased when the subjects were rather free to choose their response (following ambiguous imperative signals), but declined when the signal color unequivocally constrained the response (obvious signals). These findings suggest the occurrence of a starting-point shift on top of which a slope shift overlapped when the reward information remained a relevant factor for choosing the response. Eventually, the strength of the reward effect on CS excitability in the ambiguous trials correlated with the degree to which subjects preferred their left hand in that same condition.

Discussion

It has been suggested that reward drives motor decisions by regulating a competition between potential action representations within motor areas (Cisek and Kalaska, 2010; Sul et al., 2011; Klein-Flügge and Bestmann, 2012). A first fundamental step in the verification of this assumption consists in showing that reward can modulate motor activity during the preparation of a movement requiring an action selection (Davranche et al., 2005; Bogacz et al., 2007). Here, we provide evidence in support of this idea by demonstrating significant differences in the pattern of CS excitability changes during action selection in two different monetary reward conditions.

In the reward$_{biased}$ condition, left-hand responses were associated with more advantageous rewards than right-hand responses; in the reward$_{neutral}$ condition, the reward was equivalent for left- and right-hand responses. Behavioral results indicate that reward influenced the subjects' performance, increasing their propensity for producing responses with the left hand in the reward$_{biased}$ condition, compared to the reward$_{neutral}$ condition. In addition, the analysis of RTs reveals a specific gain in the time to prepare left-hand responses in the reward$_{biased}$ condition. Altogether, this attitude guaranteed the biggest financial gain, a finding consistent with the idea that any behavior is associated with the quest for the largest reward (D. Lee et al., 2012).

The effect of reward on CS excitability was assessed by measuring MEPs in a left hand muscle at several time points during movement preparation. Consistent with many previous reports, MEPs were strongly suppressed at the onset of the imperative signal. This effect reflects the recruitment of a mechanism that suppresses preplanned responses to avoid their hasty initiation during delay periods (Sinclair and Hammond, 2009; Duque et al., 2010). Then, following the imperative signal, the amplitude of MEPs progressively increased. This increase was more pronounced when the imperative signal had indicated a left-hand response, compared to when it had indicated a right-hand response (left-hand nonselected), reflecting a selective buildup of activity in the motor cortex controlling the forthcoming response (Duque et al., 2008; Michelet et al., 2010).

Based on bounded-accumulation models, two predictions exist on how reward could regulate the buildup of motor activity during response selection (Fig. 1*A*). One possibility is that it adjusts the initial level of activity of potential action representations at the start of the competition process. This adjustment, called "starting-point shift," regulates the amount of accumulated activation required to reach the selection threshold (Domenech and Dreher, 2010). Another possibility is that reward modulates the mean rate of the accumulation during the selection process, called a "slope shift" (Garrett et al., 2012; Y. Lee et al., 2012). In the present study, we aimed at directly testing these two predictions.

At the beginning of the selection process, the two reward conditions were associated with distinct levels of CS excitability. As such,

left MEPs elicited at the onset of the imperative signal were larger in the condition where the left hand was associated with larger gains (reward$_{biased}$) compared to the reward$_{neutral}$ condition. Hence, the starting point of the accumulation of motor activity was upregulated in the reward$_{biased}$ condition; this effect was not due to a larger preparation of left responses in the reward$_{biased}$ condition (see above, Results, Control experiment). Interestingly, at later stages of the decision period, the reward effect depended on the clearness of the imperative signal. When the choice was based on ambiguous signals, the reward effect increased; that is, CS facilitation in the reward$_{biased}$ condition was augmented when compared to that at the beginning of the selection process. This suggests that, following ambiguous imperative signals, a reward-related slope-shift effect occurred on top of the starting-point shift (Fig. 1B, left). In contrast, when the imperative signal was obvious, calling explicitly for a left- or right-hand response, the effect of the reward bias on CS excitability became less pronounced (Fig. 1B, right); that is, the CS facilitation in the reward$_{biased}$ condition was decreased with respect to that at the beginning of the selection process.

Interestingly, the reward-related change in CS excitability following ambiguous signals correlated with the degree to which subjects switched from preferring their right to their left hand following ambiguous (compared to obvious signals). This indicates a link between our CS excitability observations during action selection and the subjects' behavior in the reward$_{biased}$ condition. Finally, baseline MEPs (elicited during the intertrial interval) were comparable in the two conditions, suggesting that the reward-related changes in CS excitability were specific to the selection period.

The reward bias induced a slope shift following ambiguous, but not following obvious imperative signals. These results suggest that the influence of reward on CS excitability depended on whether this information was important for making a choice. After obvious signals, the choice was entirely constrained by the imperative signal, and hence the reward value of potential actions was irrelevant. Rather, taking this information into account could result in errors and thus in a penalty (negative score), which was particularly severe following obvious signals. In contrast, after ambiguous signals, the penalty was moderate. Moreover, when the signal was indistinct, all responses were regarded as correct. Hence, in that condition, reward information was much more important, because it guided selection toward the most beneficial outcome.

Our results provide direct evidence for the idea that reward modulates motor excitability during motor preparation. Especially relevant to the current issue are two previous works (Kapogiannis et al., 2008; Gupta and Aron, 2011) that revealed reward-related activity in the motor cortex by means of TMS.
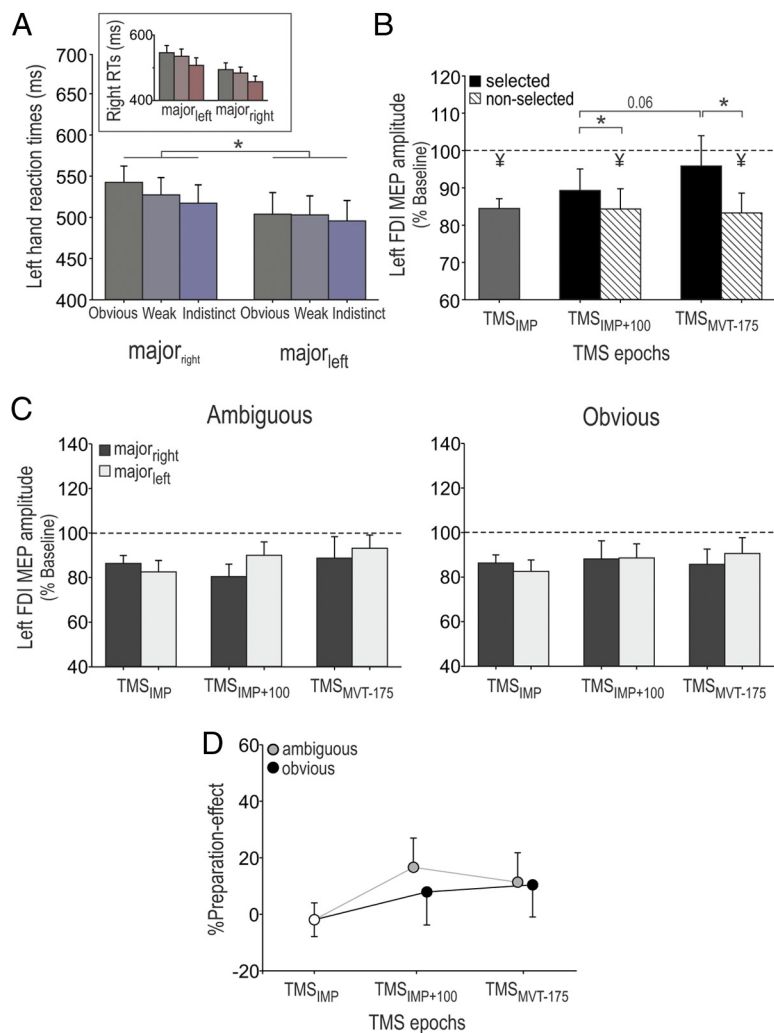
Figure 7. Illustration of the preparation effect in the control experiment (*n* = 10). MEPs are shown as raw data but were log-transformed for statistical analyses. *A*, Left-hand RTs are shown for each color category in each condition (major$_{left}$ and major$_{right}$); right-hand RTs are shown in the inset. Note the systematically faster RTs for the most frequent response in each condition (left in major$_{left}$ and right in the major$_{right}$). *B*, Amplitudes of MEPs, expressed as percentages of baseline, when the left FDI muscle was selected (black histograms) or not selected (hatched histograms) at the different TMS epochs. *C*, Amplitude of left FDI muscle MEPs following ambiguous (left) or obvious (right) signals in the major$_{right}$ (dark histograms) and the major$_{left}$ (light histograms) conditions. *D*, The %Preparation-Effect (MEP$_{major-left}$ − MEP$_{major-right}$/MEP$_{major-right}$) did not increase along the preparation period, whatever the saturation of the imperative signal. *p < 0.05. ¥Significantly different (*p < 0.05) from MEPs elicited at TMS$_{BASELINE}$.

However, in the first study (Kapogiannis et al., 2008), because subjects were at rest and no movement was required, it was not possible to tell whether the observed reward-related changes in CS excitability were specific to action selection. In the second study (Gupta and Aron, 2011), subjects were presented with stimuli about food or money and were required to indicate, after an imperative signal, whether they would be interested in obtaining them after the experiment. The authors showed that CS excitability was largest following presentation of preferred stimuli. However, CS excitability was assessed by applying TMS before the imperative signal and thus before any response (still unspecified at that time) could be selected. Hence, it was again not possible in that study, unlike in the present one, to assess whether reward modulates the buildup of motor activity during response preparation.

In the present study, the reward was always more advantageous for left-hand responses, although implicitly. This aspect of

the task is likely responsible for the fact that left MEPs were already upregulated at the onset of the imperative signal. Future experiments are required to assess the impact of reward on CS excitability when subjects cannot anticipate what the most rewarding response will be. In addition, it is worth noting that in the present design, both the absolute and relative (with respect to the right hand) reward values of left-hand responses were enhanced in the reward_biased condition. We decided to do so to increase our chances of observing an influence of reward in the present study. Our results reveal that reward modulates motor activations during response preparation. Yet, it remains to be determined which aspect of the change in reward (absolute and/or relative value) drove the modulation of CS excitability.

On the basis of our findings, interesting hypotheses can be formulated concerning the role of the motor cortex in action selection. In particular, one central question concerns the exact functional contribution of the reward-related motor regulation observed in the present experiment to action selection. One possibility is that this tuning of motor representations in M1 is critical, as it constitutes the information on the basis of which the choices are made; that is, motor decisions would directly emerge from the regulation of M1 activity (Hare et al., 2011). The integration of reward information within competing action plans in the motor cortex could rely on top-down influences originating from regions in the prefrontal (Noonan et al., 2011; Rushworth et al., 2011) and parietal cortices (Andersen and Cui, 2009), as well as in the basal ganglia (Neubert et al., 2010; Frank and Badre, 2012). Another possibility is that the changes in CS excitability observed in the present study are due to selection processes occurring in other areas, such as the premotor cortex (Donner et al., 2009; Gail et al., 2009), which spread to M1 (Fecteau and Munoz, 2005; Enama et al., 2012). In this case, activations in M1 could still reflect the accumulated evidence toward a choice but would not be critical for the selection process itself. Rather, these early adjustments could serve to preactivate competing motor plans, allowing a prompt initiation of selected actions (O'Shea et al., 2007).

In conclusion, our results suggest that motor decisions encompass a competition in motor areas, tuned by contextual cues in the setting of an integrative parallel process (Cisek, 2007; McKinstry et al., 2008; Tosoni et al., 2008). Yet, such an assertion still requires future studies to show that the motor tuning observed during response selection is functionally critical to the elaboration of motor decisions.

References

Andersen RA, Cui H (2009) Intention, action planning, and decision making in parietal-frontal circuits. Neuron 63:568–583. CrossRef Medline

Bogacz R, Usher M, Zhang J, McClelland JL (2007) Extending a biologically inspired model of choice: multi-alternatives, nonlinearity and value-based multidimensional choice. Philos Trans R Soc Lond B Biol Sci 362:1655–1670. CrossRef Medline

Brown S, Heathcote A (2005) A ballistic model of choice response time. Psychol Rev 112:117–128. CrossRef Medline

Cavallo A, Becchio C, Sartori L, Bucchioni G, Castiello U (2012) Grasping with tools: corticospinal excitability reflects observed hand movements. Cereb Cortex 22:710–716. CrossRef Medline

Chen R, Hallett M (1999) The time course of changes in motor cortex excitability associated with voluntary movement. Can J Neurol Sci 26:163–169. Medline

Cisek P (2006) Integrated neural processes for defining potential actions and deciding between them: a computational model. J Neurosci 26:9761–9770. CrossRef Medline

Cisek P (2007) Cortical mechanisms of action selection: the affordance competition hypothesis. Philos Trans R Soc Lond B Biol Sci 362:1585–1599. Medline

Cisek P, Kalaska JF (2010) Neural mechanisms for interacting with a world full of action choices. Annu Rev Neurosci 33:269–298. CrossRef Medline

Davranche K, Burle B, Audiffren M, Hasbroucq T (2005) Information processing during physical exercise: a chronometric and electromyographic study. Exp Brain Res 165:532–540. CrossRef Medline

Domenech P, Dreher JC (2010) Decision threshold modulation in the human brain. J Neurosci 30:14305–14317. CrossRef Medline

Donner TH, Siegel M, Fries P, Engel AK (2009) Buildup of choice-predictive activity in human motor cortex during perceptual decision making. Current Biol 19:1581–1585. CrossRef Medline

Duque J, Ivry RB (2009) Role of corticospinal suppression during motor preparation. Cereb Cortex 19:2013–2024. CrossRef Medline

Duque J, Mazzocchio R, Dambrosia J, Murase N, Olivier E, Cohen LG (2005) Kinematically specific interhemispheric inhibition operating in the process of generation of a voluntary movement. Cereb Cortex 15:588–593. Medline

Duque J, Mazzocchio R, Stefan K, Hummel F, Olivier E, Cohen LG (2008) Memory formation in the motor cortex ipsilateral to a training hand. Cereb Cortex 18:1395–1406. CrossRef Medline

Duque J, Lew D, Mazzocchio R, Olivier E, Ivry RB (2010) Evidence for two concurrent inhibitory mechanisms during response preparation. J Neurosci 30:3793–3802. CrossRef Medline

Duque J, Labruna L, Verset S, Olivier E, Ivry RB (2012) Dissociating the role of prefrontal and premotor cortices in controlling inhibitory mechanisms during motor preparation. J Neurosci 32:806–816. CrossRef Medline

Enama ME, Hu Z, Gordon I, Costner P, Ledgerwood JE, Grady C (2012) Randomization to standard and concise informed consent forms: Development of evidence-based consent practices. Contemp Clin Trials 33:895–902. CrossRef Medline

Fecteau JH, Munoz DP (2005) Correlates of capture of attention and inhibition of return across stages of visual processing. J Cogn Neurosci 17:1714–1727. CrossRef Medline

Frank MJ, Badre D (2012) Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. Cereb Cortex 22:527–536. CrossRef Medline

Gail A, Klaes C, Westendorff S (2009) Implementation of spatial transformation rules for goal-directed reaching via gain modulation in monkey parietal and premotor cortex. J Neurosci 29:9490–9499. CrossRef Medline

Garrett DD, Kovacevic N, McIntosh AR, Grady CL (2012) The modulation of BOLD variability between cognitive states varies by age and processing speed. Cereb Cortex. Advance online publication. doi:10.1093/cercor/bhs055 CrossRef Medline

Gold JI, Shadlen MN (2007) The neural basis of decision making. Annu Rev Neurosci 30:535–574. CrossRef Medline

Gupta N, Aron AR (2011) Urges for food and money spill over into motor system excitability before action is taken. Eur J Neurosci 33:183–188. CrossRef Medline

Hallett M (2007) Transcranial magnetic stimulation: a primer. Neuron 55:187–199. CrossRef Medline

Hare TA, Schultz W, Camerer CF, O'Doherty JP, Rangel A (2011) Transformation of stimulus value signals into motor commands during simple choice. Proc Natl Acad Sci U S A 108:18120–18125. CrossRef Medline

Hasbroucq T, Kaneko H, Akamatsu M, Possamaï CA (1997) Preparatory inhibition of cortico-spinal excitability: a transcranial magnetic stimulation study in man. Brain Res Cogn Brain Res 5:185–192. CrossRef Medline

Kapogiannis D, Campion P, Grafman J, Wassermann EM (2008) Reward-related activity in the human motor cortex. Eur J Neurosci 27:1836–1842. CrossRef Medline

Kim B, Basso MA (2010) A probabilistic strategy for understanding action selection. J Neurosci 30:2340–2355. CrossRef Medline

Klein-Flügge MC, Bestmann S (2012) Time-dependent changes in human corticospinal excitability reveal value-based competition for action during decision processing. J Neurosci 32:8373–8382. CrossRef Medline

Labruna L, Fernández-del-Olmo M, Ivry RB (2011) Comparison of different baseline conditions in evaluating factors that influence motor cortex excitability. Brain Stim 4:152–155. CrossRef Medline

Lee D, Seo H, Jung MW (2012) Neural basis of reinforcement learning and decision making. Annu Rev Neurosci 35:287–308. CrossRef Medline

Lee Y, Anaki D, Grady CL, Moscovitch M (2012) Neural correlates of tem-

poral integration in face recognition: An fMRI study. Neuroimage 61: 1287–1299. CrossRef Medline

Leocani L, Cohen LG, Wassermann EM, Ikoma K, Hallett M (2000) Human corticospinal excitability evaluated with transcranial magnetic stimulation during different reaction time paradigms. Brain 123:1161–1173. CrossRef Medline

Link SW, Heath RA (1975) A sequential theory of psychological discrimination. Psychometrika 40:77–105. CrossRef

Mazurek ME, Roitman JD, Ditterich J, Shadlen MN (2003) A role for neural integrators in perceptual decision making. Cereb Cortex 13:1257–1269. CrossRef

McKinstry C, Dale R, Spivey MJ (2008) Action dynamics reveal parallel competition in decision making. Psychol Sci 19:22–24. CrossRef Medline

Michelet T, Duncan GH, Cisek P (2010) Response competition in the primary motor cortex: corticospinal excitability reflects response replacement during simple decisions. J Neurophysiol 104:119–127. CrossRef Medline

Neubert FX, Mars RB, Buch ER, Olivier E, Rushworth MF (2010) Cortical and subcortical interactions during action reprogramming and their related white matter pathways. Proc Natl Acad Sci U S A 107:13240–13245. CrossRef Medline

Noonan MP, Mars RB, Rushworth MF (2011) Distinct roles of three frontal cortical areas in reward-guided behavior. J Neurosci 31:14399–14412. CrossRef Medline

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113. CrossRef Medline

O'Shea J, Sebastian C, Boorman ED, Johansen-Berg H, Rushworth MF (2007) Functional specificity of human premotor-motor cortical interactions during action selection. Eur J Neurosci 26:2085–2095. CrossRef Medline

Pastor-Bernier A, Cisek P (2011) Neural correlates of biased competition in premotor cortex. J Neurosci 31:7083–7088. CrossRef Medline

Pearson JM, Hayden BY, Raghavachari S, Platt ML (2009) Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task. Curr Biol 19:1532–1537. CrossRef Medline

Posner, M.I. (1978) Chronometric explorations of mind. Hillsdale, NJ: Lawrence Erlbaum.

Reynolds C, Ashby P (1999) Inhibition in the human motor cortex is reduced just before a voluntary contraction. Neurology 53:730–735. CrossRef Medline

Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S (2004) The role of the medial frontal cortex in cognitive control. Science 306:443–447. CrossRef Medline

Roesch MR, Olson CR (2004) Neuronal activity related to reward value and motivation in primate frontal cortex. Science 304:307–310. CrossRef Medline

Rossini PM (1988) The anatomic and physiologic bases of motor-evoked potentials. Neurol Clin 6:751–769. Medline

Rossini PM, Zarola F, Stalberg E, Caramia M (1988) Pre-movement facilitation of motor-evoked potentials in man during transcranial stimulation of the central motor pathways. Brain Res 458:20–30. CrossRef Medline

Rushworth MF, Noonan MP, Boorman ED, Walton ME, Behrens TE (2011) Frontal cortex and reward-guided learning and decision-making. Neuron 70:1054–1069. CrossRef Medline

Sartori L, Cavallo A, Bucchioni G, Castiello U (2011) Corticospinal excitability is specifically modulated by the social dimension of observed actions. Exp Brain Res 211:557–568. CrossRef Medline

Selen LP, Shadlen MN, Wolpert DM (2012) Deliberation in the motor system: reflex gains track evolving evidence leading to a decision. J Neurosci 32:2276–2286. CrossRef Medline

Sinclair C, Hammond GR (2009) Excitatory and inhibitory processes in primary motor cortex during the foreperiod of a warned reaction time task are unrelated to response expectancy. Exp Brain Res 194:103–113. CrossRef Medline

Sul JH, Jo S, Lee D, Jung MW (2011) Role of rodent secondary motor cortex in value-based action selection. Nat Neurosci 14:1202–1208. CrossRef Medline

Sutton RS, Barto AG (1981) Toward a modern theory of adaptive networks: expectation and prediction. Psychol Rev 88:135–170. CrossRef Medline

Tosoni A, Galati G, Romani GL, Corbetta M (2008) Sensory-motor mechanisms in human parietal cortex underlie arbitrary visual decisions. Nat Neurosci 11:1446–1453. CrossRef Medline