Published in final edited form as:

Neuropsychologia. 2015 February ; 68: 31–37. doi:10.1016/j.neuropsychologia.2014.12.021.

Modulation of corticospinal excitability by reward depends on task framing

Eric Mooshagian^{1,2,*}, Aysha Keisler^{1,2}, Trelawny Zimmermann^{1,2}, Janell M. Schweickert³, and Eric M. Wassermann^{1,**}

¹Behavioral Neurology Unit, National Institute of Neurological Disorders and Stroke, National Institutes of Health, 10 CENTER DR, MSC 1440, Bethesda, MD, 20892-1440

²Center for Neuroscience and Regenerative Medicine, Uniformed Services University of Health Sciences, Henry M. Jackson Foundation, Rockville, MD 20852, USA

³The Wake Kendall Group, PLLC, 5247 Wisconsin Avenue NW, Suite 4, Washington DC 20015

Abstract

Findings from previous transcranial magnetic stimulation (TMS) experiments suggest that the primary motor cortex (M1) is sensitive to reward conditions in the environment. However, the nature of this influence on M1 activity is poorly understood. The dopamine neuron response to conditioned stimuli encodes reward probability and outcome uncertainty, or the extent to which the outcome of a situation is known. Reward uncertainty and probability are related: uncertainty is maximal when probability is 0.5 and minimal when probability is 0 or 1 (i.e., certain outcome). Previous TMS-reward studies did not examine these factors independently. Here, we used single-pulse TMS to measure corticospinal excitability in 40 individuals while they performed a simple computer task, making guesses to find or avoid a hidden target. The task stimuli implied three levels of reward probability and two levels of uncertainty. We found that reward probability level interacted with the trial search condition. That is, motor evoked potential (MEP) amplitude, a measure of corticospinal neuron excitability, increased with increasing reward probability when participants were instructed to “find” a target, but not when they were instructed to “avoid” a target. There was no effect of uncertainty on MEPs. Response times varied with the number of choices. A subset of participants also received paired-pulse stimulation to evaluate changes in short-intracortical inhibition (SICI). No effects of SICI were observed. Taken together, the results suggest that the reward-contingent modulation of M1 activity reflects reward probability or a related aspect of utility, not outcome uncertainty, and that this effect is sensitive to the conceptual framing of the task.

^{**}Correspondence to: Eric Wassermann, M.D., 10 CENTER DR, MSC 1440, BETHESDA, MD 20892-1440, wassermann@ninds.nih.gov, Phone +1 301 496 0151.

^{*}Current Address: Department of Anatomy and Neurobiology, Washington University School of Medicine, Saint Louis, MO 63110

The authors have no conflicts of interest to report.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

uncertainty; expectation; reward probability; transcranial magnetic stimulation; corticospinal excitability

Introduction

Estimating the expected value and risk associated with potential action is essential for successful learning and decision-making. The expected value of an action's outcome is the product of the magnitude of the potential reward and the probability of achieving the outcome. When reward magnitude is constant, expected value increases linearly with reward probability. The risk of failure associated with a choice is described by the uncertainty of achieving a particular outcome. Uncertainty also varies with reward probability and is maximal when the probability of reward is 0.5. Single-cell recordings in monkeys show that midbrain dopamine (DA) neurons encode reward probability and uncertainty (Fiorillo, Tobler, & Schultz, 2003). Phasic firing in these neurons increases with both reward probability and magnitude, but the firing rate does not distinguish between these two parameters when the expected reward value is constant. A tonic response, appearing to code uncertainty, peaks when the probability of reward is 0.5 (Fiorillo et al., 2003). In humans, functional magnetic resonance imaging (fMRI) reveals midbrain activations associated with these features (Aron et al., 2004; Dreher, Kohn, & Berman, 2006).

Transcranial magnetic stimulation (TMS) is a useful tool for noninvasive study of motor system physiology in humans. Motor evoked potential (MEP) amplitude reflects the aggregate excitability of primary motor cortex (M1) output cells (Wassermann & Zimmermann, 2012). In a few recent human studies, investigators applied TMS over M1 to measure changes in corticospinal output excitability in response to reward-related events. These studies found increased corticospinal excitability with the desirability of an outcome (Gupta & Aron, 2011) or a momentary reward (Thabit et al., 2011), and increased paired-pulse inhibition with increased expectation of receiving a reward while passively viewing a slot-machine simulation (Kapogiannis, Campion, Grafman, & Wassermann, 2008; Kapogiannis et al., 2011). The neural basis of these effects remains unknown, but they almost certainly reflect neural signaling about outcomes and values (Kapogiannis et al., 2011).

Missing from these studies, however, is information on whether reward effects in M1 reflect reward probability or uncertainty coding. This is an important omission, since such data could help identify source of the signals driving the M1 excitability changes. We designed a paradigm, which delivered a fixed reward with varying probability, allowing us to distinguish between reward probability and uncertainty (Figure 1). In a task where a fixed reward is either delivered or withheld, outcome uncertainty is a function of reward probability and maximal when reward probability is least certain ($p_{\text{reward}} = 0.5$) and minimal when reward probability is most certain ($p_{\text{reward}} = 1$ or 0). Our prediction was that if M1 excitability is affected by reward probability, then increasing reward probability should produce a change in MEP amplitude; whereas, if M1 excitability reflects primarily outcome

uncertainty, varying reward probability should produce an inverted U-shaped response, as probability varies from zero to unity with a maximum effect at .5.

Materials and Methods

Participants

We studied 40 healthy, right-handed volunteers (21 women, 19 men) (aged 21–41), all of whom were screened and examined by a neurologist. Exclusion criteria were neuroactive medication use, history of central nervous system disorders, or neurological abnormalities. Participants gave written informed consent and the CNS Institutional Review Board of the National Institutes of Health approved the study.

Task and Stimuli

Figure 2 depicts the stimuli and experimental design. We devised a task where the combination of a trial instruction and a visual stimulus informed the participant about the outcome uncertainty and probability of reward for that trial. The objective of the task was to find or avoid a target stimulus. Successful responses resulted in a prize of 25 cents (\$0.25 USD).

The stimuli were constructed of drawings of circles and squares. The initial display on each trial was one of the following arrangements: a single square outline (“empty square”) in the center of the screen, a single shaded square in the center of the screen, or two shaded squares, one on either side of the center. The target stimulus, a circle, would be hidden by a shaded square. Prior to beginning the experiment, participants were told that the circle was never present behind the single empty square, always present behind the single shaded square, and always present behind one of the two shaded squares. On each trial, participants were instructed either to *Find* or to *Avoid* the location of the target stimulus (search condition). Therefore, the reward contingencies and probabilities were as follows: for the find search condition, a single empty square implied a p_{reward} of 0, a single shaded square implied p_{reward} of 1, and two shaded squares implied p_{reward} of 0.5; for the avoid search condition the probabilities were reversed so that a single empty square implied a p_{reward} of 1, a single shaded square implied p_{reward} of 0, and two shaded squares implied p_{reward} was 0.5. The number of stimuli presented indicated the outcome uncertainty regardless of search condition: a single square, filled or not, implied no outcome uncertainty, while two squares implied maximal uncertainty. Therefore, the combination of the instruction and the target stimulus on each trial provided the participant with complete information about the outcome uncertainty and the probability of receiving a reward (Figure 2A).

To control for possible asymmetrical effects of the task stimuli themselves, half of the participants were instructed to find or avoid the circle as just described; the other half were instructed to find or avoid the square with no circle “behind” it (task instruction). Thus, the target stimulus was the square without a circle behind it. In this condition, the reward probabilities were: for the find search condition, a single empty square implied a p_{reward} of 1, a single shaded square implied p_{reward} of 0, and two shaded squares implied p_{reward} of 0.5; for the avoid search condition the probabilities were reversed so that a single empty square

implied a p_{reward} of 0, a single shaded square implied p_{reward} of 1, and two shaded squares implied p_{reward} was 0.5. In other words, the reward probabilities associated with each stimulus configuration were reversed between the two participant groups (Figure 2A), but the task was otherwise the same.

The sequence and timing of trial events is depicted in Figure 2B. Participants sat in front of a computer monitor with their left index and middle fingers over two response buttons of a response box. The task stimuli were presented centered on the display. Each trial proceeded as follows: At the start of each trial, the trial-specific search instruction to find or avoid the target stimulus was presented. Next, the initial stimuli appeared on the screen. TMS was delivered to the left M1 250 ms after stimulus onset. After 750 ms, participants were prompted to indicate their selection using their left hand on a button box; they were required to make a response for all trials regardless of the expected outcome. In the two-square condition, participants pressed the button corresponding to the location of their choice. On single stimulus trials, regardless of whether the presented square was filled or empty, participants were required to press the button under the middle finger even though the trial outcome was certain in order to equate the motor response requirement across trials. Button presses were made with the left hand (see Electromyography and Transcranial Magnetic Stimulation sections, below).

500-ms after the response, the selection was fed back to the participant by the appearance of a red outline around the selected square (depicted as the dashed gray line in Figure 2B). Finally, the location of the circle, if present, was revealed, accompanied by a message indicating whether the response was correct and whether a prize was awarded. Correct responses resulted in a prize of 25 cents; no money was received or lost for incorrect guesses. Trials were initiated every 8 s and the order of trial types was randomized.

The groups instructed to find or avoid the “square without a circle” (*No Circle*) and the circle (*Circle*) performed 120 and 240 trials, respectively. The additional trials for the *Circle* group were due to the addition of 120 paired-pulse stimulation trials (see below). Trial type was evenly distributed across SEARCH CONDITION (Avoid, Find) and REWARD PROBABILITY (0.00, 0.50, 1.00). The experiment was controlled using the software package Presentation® (www.neurobs.com).

Electromyography

Surface electrodes were applied to the skin over the first dorsal interosseus (FDI) muscle of the right hand. Muscle relaxation was ensured by monitoring the electromyogram continuously on a loudspeaker and instructing participants to maintain silence. The signal was amplified (1000x) and filtered (band-pass 90 Hz; 1 kHz; Coulbourn Instruments, Whitehall, PA), digitized at 2 kHz (Micro 1401, Cambridge Electronics Design, Cambridge, UK) and displayed and stored on a computer for offline analysis.

Transcranial Magnetic Stimulation

TMS was delivered from Magstim 200® stimulators connected via a Bistim® module, through a round coil centered near the vertex, in the optimal position and orientation for

producing a MEP in the right FDI muscle. The round coil was used for both single and paired-pulse stimulation and it has been shown not to affect the measurement of corticospinal excitability compared to figure-of-eight coils (Badawy, Tarletti, Mula, Varrasi, & Cantello, 2011).

Resting motor threshold (RMT) was defined as the lowest stimulator output required to produce a motor evoked potential (MEP) $> 50 \mu\text{V}$ on at least 5 out of 10 consecutive trials. The interval between pulses during thresholding trials varied randomly by 0 – 15% around a mean of 10 s. During baseline stimulation, participants were instructed to sit at rest with their eyes open. The average RMT was ~39% of maximum stimulator output. In both the *Circle* and *No Circle* instruction groups, corticospinal excitability was measured by applying single TMS pulses. MEPs were recorded from the FDI of the right hand after stimulation of the left M1. In the *Circle* group, we also measured short-interval intracortical inhibition (SICI), a measure of intracortical GABA_A-mediated inhibition (Kujirai et al., 1993; Ziemann, Bruns, & Paulus, 1996) with paired-pulse stimulation. The intensity of the conditioning pulse was adjusted to 65% of RMT, and the test pulse intensity was adjusted to produce a 1 mV MEP. The paired-pulses were applied with a 2-ms interstimulus interval and randomly alternated with trials where the test pulse was given alone. SICI was measured as the ratio of the mean MEP amplitude from the paired pulses, divided by the mean amplitude from the test pulses alone, for each experimental condition. A ratio < 1 indicated inhibition by the conditioning pulse and a ratio > 1 indicated facilitation.

Data Processing and Analysis

Trials with pre-stimulus EMG activity in the FDI (signal mean ± 3 SD during the 50 ms preceding the pulse) were discarded. Peak-to-peak MEP amplitude was measured for each trial.

Response time (RT) was measured as the time in milliseconds (ms) between the cue-to-respond to the button press. Trials in which participants responded prematurely (< 150 ms) or late (> 800 ms) were excluded from the analyses.

A preliminary analysis was conducted to determine if there was any effect of the instructed target: A repeated-measures ANOVA of the MEP values, with the between-subjects factor INSTRUCTED TARGET (circle, no circle), and within-subjects factors SEARCH CONDITION (avoid, find), and REWARD PROBABILITY (0, 0.5, 1.0) showed that there was no main effect or interaction with INSTRUCTED TARGET. Because we did not have any hypotheses about INSTRUCTED TARGET and it did not interact with any of our variables of interest, we collapsed the data across instructed targets for our subsequent analyses. Separate repeated-measures ANOVA models were fitted to the MEP and RT data with SEARCH CONDITION, REWARD PROBABILITY and their interactions, as within-subjects factors. We set the criterion for significance at $\alpha = .05$. We applied the Greenhouse–Geisser correction to adjust the degrees of freedom when the sphericity criterion was not met. Post-hoc Bonferroni-corrected t-tests were used to examine significant interactions. Data are reported as the mean $\pm$ standard error of the mean (SEM).

Results

In order to avoid effects on the MEP associated with preparation of the response movement, we recorded MEPs from the non-responding (right) hand, while button presses were made with the left hand, and the motor response occurred at least 750 ms after the TMS pulse occurred (Pascual-Leone et al., 1992). Nevertheless, we looked for an association between RT and MEP amplitude, but found none ($r = .007$, $p = .611$).

There was a borderline main effect of search condition on MEP amplitude, where MEP amplitude tended to be greater when participants were asked to find ($M = 1.22$, $SE = .07$) rather than avoid ($M = 1.17$, $SE = .06$) the target stimulus (SEARCH CONDITION, $F(1,39) = 3.586$, $p = .066$). Surprisingly, there was no main effect of reward probability, but reward probability did interact with search condition (SEARCH CONDITION $\times$ REWARD PROBABILITY, $F(1.731, 67.509) = 5.31$, $p = .01$ (Figure 3). In the Find condition, MEP amplitude increased with increasing reward probability ($F(2,78) = 5.087$, $p = .008$), such that MEP amplitude was lowest for $p_{\text{reward}} = 0.0$ ($M = 1.18$, $SE = .07$), intermediate for $p_{\text{reward}} = 0.5$ ($M = 1.21$, $SE = .07$), and highest for $p_{\text{reward}} = 1.0$ ($M = 1.27$, $SE = .07$). The $p_{\text{reward}} = 0.5$ condition was not significantly greater than the $p_{\text{reward}} = 0$ condition ($p > .05$), but the $p_{\text{reward}} = 1$ condition was significantly greater than both the $p_{\text{reward}} = 0$ [$t(39) = -2.79$, $p = .008$] and $p_{\text{reward}} = 0.5$ [$t(39) = -2.49$, $p < .05$] conditions. On the other hand, in the Avoid condition, MEP amplitude did not vary significantly across reward probability, with similar MEP amplitudes for $p_{\text{reward}} = 0.0$ ($M = 1.19$, $SE = .07$), $p_{\text{reward}} = 0.5$ ($M = 1.18$, $SE = .07$), and $p_{\text{reward}} = 1.0$ ($M = 1.15$, $SE = .06$) ($F(2,78) = .577$, $p > .05$). Post-hoc comparisons between the Find and Avoid conditions showed that MEP amplitude was significantly greater in the Find ($M = 1.27$, $SE = .07$) compared to the Avoid ($M = 1.15$, $SE = .06$) conditions only at $p_{\text{reward}} = 1.0$ ($t(39) = -2.790$, $p < .01$) (Figure 3).

These results argue against an M1 response to uncertainty, but we further confirmed this by testing the data directly for an effect of uncertainty. As previously mentioned, outcome uncertainty is a function of reward probability and is maximal when reward probability is least certain ($p_{\text{reward}} = 0.5$) and minimal when reward probability is most certain ($p_{\text{reward}} = 1$ or 0; Figure 1). The data were sorted so that the trials were designated as either certain or uncertain and submitted to repeated-measures ANOVA. As anticipated, sorting the trials in this way did not make a difference. There was no main effect of, or interaction with, uncertainty (all $p > .05$).

Half of the participants also received paired-pulse TMS trials that were interleaved with the single-pulse trials. To address the influence of expected reward probability on SICI, the mean paired-pulse ratio was calculated for each condition for each participant and submitted to another repeated-measures ANOVA. However, this did not yield any significant main effects or interactions (all $p > .05$).

RTs varied significantly across reward probability (main effect of REWARD PROBABILITY, $F(2, 78) = 19.040$, $p < .001$; Figure 4). Bonferroni-corrected pairwise comparisons showed that RTs were slower for $p_{\text{reward}} = .5$ ($M = 472$, $SE = 14$) compared to $p_{\text{reward}} = 0$ ($M = 445$, $SE = 15$), $p = .006$, or $p_{\text{reward}} = 1$ ($M = 424$, $SE = 14$), $p < .017$. There

was no interaction with search condition (SEARCH CONDITION $\times$ REWARD PROBABILITY, $p > .05$).

Discussion

Our primary aim in this study was to test whether the modulation of M1 corticospinal excitability in response to reward-related stimuli better reflects an influence of reward probability or outcome uncertainty. The most striking finding, however, was that the modulation of corticospinal excitability in anticipation of different reward outcomes depended on the larger behavioral context (search condition) i.e., whether participants had been instructed to *find* or *avoid* the target stimulus. We showed that it was only when this larger context was accounted for that an effect of reward probability emerged. These results provide new evidence that changes in corticospinal excitability reflect outcome probability, rather than uncertainty about outcome.

Other TMS studies (Gupta & Aron, 2011; Kapogiannis et al., 2008; 2011) have shown that excitability of the corticospinal system changes in response to predicted outcomes in extrinsically rewarded tasks. However, this is the first TMS experiment attempting to dissociate the influences of probability and uncertainty. We presented participants with stimuli that predicted the likelihood of a fixed reward, and applied TMS over M1 after participants were presented with the predictive stimulus, but before (750 ms) they indicated a selection. Our paradigm was designed to manipulate the extrinsic reward associated with each condition and we expected modulation of corticospinal excitability to be dominated by the influence of reward probability. In addition to the varying reward probability, however, participants were asked to accomplish the task by alternately *finding* or *avoiding* the target stimulus. We found an effect of reward probability only when the framing of the task goal (search condition) was taken into account. The dependence of the reward probability effect on the search condition is a surprising result and suggests that some effect of the search condition cue, instructing how to interact with the stimulus set, is also an important modulator of corticospinal excitability. A complete account of the results must consider the intrinsic motivational contingencies of the task in addition to the extrinsic monetary reward. The effect of search condition on intrinsic motivational state can be interpreted in the framework of approach-avoidance behavior (Carver & Scheier, 1998). In approach motivation, behavior is instigated by a positive or desirable event. In the *Find* search condition, success was framed as acquiring the target stimulus. In avoidance motivation, on the other hand, behavior is instigated by a negative or undesirable event. In the *Avoid* search condition, success was framed as evading the target stimulus. Here, we might expect the avoid stance to result in inhibition of corticospinal excitability.

If M1 corticospinal output were sensitive to reward probability, one might predict MEP amplitude to approximate a linear function of probability, ranging from the $p_{\text{reward}} = 0$ condition, through $p_{\text{reward}} = .5$ condition, with a maximal effect occurring in the $p_{\text{reward}} = 1$ condition. If M1 were sensitive to outcome uncertainty, the MEP response might approximate a U-shaped function, with the maximal modulation occurring in the $p_{\text{reward}} = .5$ condition, and minimal effects when the outcome is certain in the $p_{\text{reward}} = 0$ or $p_{\text{reward}} = 1$ conditions. In the *Find* search condition, corticospinal excitability increased, while in the

Avoid condition, corticospinal excitability decreased, though non-significantly, with reward probability. In other words, the linear function of probability is positively weighted under the *Find* condition, and negatively weighted under the *Avoid* condition. In the $p_{\text{reward}} = 0$ condition, regardless of the intrinsic motivation generated by the search instruction, there was no chance of success; participants were forced to commit what was, in effect, an error. In both search conditions, the weight is multiplied by the zero probability of being correct resulting in the same reward probability effect in each case. This is consistent with the lack of difference in MEP amplitude for the *Find* and *Avoid* search conditions in the $p_{\text{reward}} = 0$ condition. On the other hand, in the $p_{\text{reward}} = 1$ condition, participants were ensured of answering correctly. The reward probability is multiplied by a positive weight for the *Find* search condition and by a negative weight for the *Avoid* search condition and should take on their most extreme values. This is consistent with the largest MEP amplitude being observed in the *Find* search condition and the smallest MEP amplitude being observed in the *Avoid* search condition as well as the significant difference in MEP amplitude between those two conditions. Neither the probability nor the uncertainty analyses supported uncertainty as a significant determinant of excitability.

The observation of changing corticospinal excitability in response to varying expectation of reward outcome is consistent with the involvement of M1 in motor learning. This involvement has been demonstrated in physiological studies using TMS (Pascual-Leone, Grafman, & Hallett, 1994), EEG (Zhuang et al., 1998; 1997) and PET (Grafton et al., 1992; Honda et al., 1998). Inhibitory theta-burst stimulation over M1 abolishes probabilistic sequence learning while the same stimulation to supplementary motor or dorsolateral prefrontal cortex does not (Rosenthal, Roche-Kelly, Husain, & Kennard, 2009; Wilkinson, Teo, Obeso, Rothwell, & Jahanshahi, 2010). Plastic changes in M1 take place under the control of DA projections: DA D2 agonists enhance practice dependent plasticity (Meintzschel & Ziemann, 2006) and recent evidence from rats shows that DA depletion in M1 impairs acquisition of motor skill, but not execution of already learned skills (Molina-Luna et al., 2009). In addition to the much discussed basal ganglia circuits, integration of reward signals with behavior might occur in the motor cortex, another site where DA signaling influences behavioral control (Wickens, Reynolds, & Hyland, 2003). More generally, our finding is consistent with the view of M1 as more than the output stage of the executive stream, but also a contributor to higher motor processes (Graziano, 2006). Current models of action selection posit the parallel processing of potential actions before a final response is reached (Cisek, 2007), allowing responses to be selected as far downstream as M1. For instance, M1 neuron population activity is influenced by partial movement information (Bastian, Schöner, & Riehle, 2003) and cortical excitability is enhanced by response competition (Michelet, Duncan, & Cisek, 2010), implicating M1 in decision making (Selen, Shadlen, & Wolpert, 2012). Our results show that reward probability also modulates M1 activity prior to response selection, adding to the evidence that cognitive factors can exert influence even at late stages of processing.

In our experiment, we had participants respond with the hand not activated by TMS, in order to avoid contamination of the EMG response by motor preparation; correlational analysis between the mean RT and MEP amplitude for each participant confirmed that there was no

relationship between the two. This rules out the possibility that the changes in MEP amplitude that we observed simply reflected preparation of different motor responses. Additionally, on single stimulus trials (i.e., the $p_{\text{reward}} = 0$ and the $p_{\text{reward}} = 1$ conditions) participants always pressed the left response button. Yet, corticospinal excitability still varied depending on the reward probability. In fact, it is difficult to see how the interaction between search condition and reward probability observed for MEP amplitude could be due to preparation of particular motor responses. If anything, one might expect that a more pronounced effect of reward probability on the EMG response would be observed if the EMG was recorded from the acting hand.

In contrast to the physiological findings, the RT data reflect the number of choices, with the slowest RT in both groups occurring in the $p_{\text{reward}} = .5$ condition, consistent with the literature on choice RT (Hick, 1952) as well as with another recent TMS study of reward (Suzuki et al., 2014). An inverse relationship between reward probability and RT has been observed for saccadic reaction times in monkeys (Milstein & Dorris, 2007; 2011) and nose-pokes in rats (Lauwereyns & Wisniewski, 2006). Our primary interest in this experiment was to capture the influence of reward on corticospinal excitability and we did not want that measure contaminated by changes in excitability related to movement preparation. We therefore asked participants to withhold their response until they received a go cue, 1 s after the stimulus appeared and 750 ms after the TMS pulse. Our task also included $p_{\text{reward}} = 0$ and $p_{\text{reward}} = 1$ conditions for which the response was identical and did not require a decision. A task that required participants to respond closer in time to stimulus presentation, used intermediate reward probabilities, or required a choice response in all conditions, might have had a different RT outcome.

In this study, we found significant effects only on the response to single TMS pulses. Some of us (Kapogiannis et al., 2008), on the other hand, previously reported increased paired-pulse inhibition when reward probability changed to 0.5 from 0 and did not report effects on the response to single pulses. However, reevaluation of those data revealed a trend in the single-pulse data for increased excitability in the maximum, compared to the minimum, outcome uncertainty condition (unpublished observation in data reported Kapogiannis et al., 2008). While the raw paired-pulse data showed a trend toward greater inhibition when the probability of reward increased, larger MEPs in response to the test pulse alone also contributed to the apparent inhibition when the data were expressed as the ratio of conditioned to unconditioned responses. The co-occurrence of larger unconditioned and smaller conditioned MEPs seems incongruous, but a change in neuronal excitability, which increased the sensitivity to both the test *and* conditioning pulses, might produce greater excitatory *and* inhibitory effects.

Several challenges remain. In the present experiment, we used only one reward amount, so expected reward value/utility and probability covaried. Therefore, we could not distinguish between expected reward value and reward probability. However, single-unit recording experiments show that individual DA neurons increase their firing rate with reward probability and magnitude, suggesting that both factors contribute to value (Tobler, Fiorillo, & Schultz, 2005). We also measured the cortical response at only one time point (250 ms after stimulus presentation) and it is possible that the modulation of excitability is different

closer to movement onset or after the response. Single-unit recordings from DA neurons (Fiorillo et al., 2003), suggest that outcome uncertainty is coded by sustained firing, which occurs after the phasic probability response. It is also possible that TMS applied after the delivery or omission of reward would be sensitive to reward prediction errors (Bayer & Glimcher, 2005; Morris, Arkadir, Nevet, Vaadia, & Bergman, 2004; Schultz, Dayan, & Montague, 1997). The present study, however, was not designed to address those issues. Additionally, only positive rewards were tested, i.e., participants did not lose money for incorrect guesses, so whether the expectation of punishment or loss is reflected in M1 excitability remains an open question.

In summary, we demonstrated that the reward-contingent response of M1 corticospinal excitability to TMS reflects reward probability and that task-framing can modulate this response. These findings have implications for motor learning and decision-making, both of which rely, in part, upon M1 (Honda et al., 1998; Karni et al., 1995; Michelet et al., 2010) and dopamine-mediated reward (D'Ardenne, McClure, Nystrom, & Cohen, 2008; Molina-Luna et al., 2009; Schultz, 2007; Schultz et al., 2008). This study also adds to the evidence that TMS provides a useful means of monitoring the activity of the human reward system during behavior.

Acknowledgments

Support for this work included intramural funding from the Clinical Neuroscience Program of the National Institute of Neurological Disorders and Stroke (EW) and funding from the Center for Neuroscience and Regenerative Medicine (AK, EM, EW, and TZ), via the Henry M. Jackson Foundation for the Advancement of Military Medicine. The authors would like to thank Mr. Eric Emmons for his assistance with pilot data collection.

References

- Aron AR, Shohamy D, Clark J, Myers C, Gluck MA, Poldrack RA. Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *Journal of Neurophysiology*. 2004; 92(2):1144–1152.10.1152/jn.01209.2003 [PubMed: 15014103]
- Badawy RAB, Tarletti R, Mula M, Varrasi C, Cantello R. The routine circular coil is reliable in paired-TMS studies. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*. 2011; 122(4):784–788.10.1016/j.clinph.2010.10.027 [PubMed: 21071268]
- Bastian A, Schöner G, Riehle A. Preshaping and continuous evolution of motor cortical representations during movement preparation. *The European Journal of Neuroscience*. 2003; 18(7): 2047–2058. [PubMed: 14622238]
- Bayer HM, Glimcher PW. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*. 2005; 47(1):129–141.10.1016/j.neuron.2005.05.020 [PubMed: 15996553]
- Carver, CS.; Scheier, MF. *On the self-regulation of behavior*. New York: Cambridge University Press; 1998.
- Cisek P. Cortical mechanisms of action selection: the affordance competition hypothesis. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*. 2007; 362(1485):1585–1599.10.1098/rstb.2007.2054
- D'Ardenne K, McClure SM, Nystrom LE, Cohen JD. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science (New York, NY)*. 2008; 319(5867):1264–1267.10.1126/science.1150605
- Dreher JC, Kohn P, Berman KF. Neural coding of distinct statistical properties of reward information in humans. *Cerebral Cortex (New York, NY: 1991)*. 2006; 16(4):561–573.10.1093/cercor/bhj004

- Fiorillo CD, Tobler PN, Schultz W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science (New York, NY)*. 2003; 299(5614):1898–1902.10.1126/science.1077349
- Grafton ST, Mazziotta JC, Presty S, Friston KJ, Frackowiak RS, Phelps ME. Functional anatomy of human procedural learning determined with regional cerebral blood flow and PET. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*. 1992; 12(7):2542–2548. [PubMed: 1613546]
- Graziano M. THE ORGANIZATION OF BEHAVIORAL REPERTOIRE IN MOTOR CORTEX. *Annual Review of Neuroscience*. 2006; 29(1):105–134.10.1146/neuro.2006.29.issue-1
- Gupta N, Aron AR. Urges for food and money spill over into motor system excitability before action is taken. *The European Journal of Neuroscience*. 2011; 33(1):183–188.10.1111/j.1460-9568.2010.07510.x [PubMed: 21091805]
- Hick WE. On the rate of gain of information. *Quarterly Journal of Experimental Psychology*. 1952; 4(1):11–26.10.1080/17470215208416600
- Honda M, Deiber MP, Ibáñez V, Pascual-Leone A, Zhuang P, Hallett M. Dynamic cortical involvement in implicit and explicit motor sequence learning. A PET study. *Brain: a Journal of Neurology*. 1998; 121(Pt 11):2159–2173. [PubMed: 9827775]
- Kapogiannis D, Campion P, Grafman J, Wassermann EM. Reward-related activity in the human motor cortex. *The European Journal of Neuroscience*. 2008; 27(7):1836–1842.10.1111/j.1460-9568.2008.06147.x [PubMed: 18371077]
- Kapogiannis D, Mooshagian E, Campion P, Grafman J, Zimmermann TJ, Ladit KC, Wassermann EM. Reward processing abnormalities in Parkinson's disease. *Movement Disorders: Official Journal of the Movement Disorder Society*. 2011; 26(8):1451–1457.10.1002/mds.23701 [PubMed: 21538525]
- Karni A, Meyer G, Jezzard P, Adams MM, Turner R, Ungerleider LG. Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature*. 1995; 377(6545):155–158.10.1038/377155a0 [PubMed: 7675082]
- Kujirai T, Caramia MD, Rothwell JC, Day BL, Thompson PD, Ferbert A, et al. Corticocortical inhibition in human motor cortex. *The Journal of Physiology*. 1993; 471:501–519. [PubMed: 8120818]
- Lauwereyns J, Wisniewski RG. A reaction-time paradigm to measure reward-oriented bias in rats. *Journal of Experimental Psychology Animal Behavior Processes*. 2006; 32(4):467–473.10.1037/0097-7403.32.4.467 [PubMed: 17044749]
- Meintzschel F, Ziemann U. Modification of practice-dependent plasticity in human motor cortex by neuromodulators. *Cerebral Cortex (New York, NY: 1991)*. 2006; 16(8):1106–1115.10.1093/cercor/bhj052
- Michelet T, Duncan GH, Cisek P. Response competition in the primary motor cortex: corticospinal excitability reflects response replacement during simple decisions. *Journal of Neurophysiology*. 2010; 104(1):119–127.10.1152/jn.00819.2009 [PubMed: 20445034]
- Milstein DM, Dorris MC. The influence of expected value on saccadic preparation. *Journal of Neuroscience*. 2007; 27(18):4810–4818.10.1523/JNEUROSCI.0577-07.2007 [PubMed: 17475788]
- Milstein DM, Dorris MC. The Relationship between Saccadic Choice and Reaction Times with Manipulations of Target Value. *Frontiers in Neuroscience*. 2011; 5:122.10.3389/fnins.2011.00122 [PubMed: 22028681]
- Molina-Luna K, Pekanovic A, Röhrich S, Hertler B, Schubring-Giese M, Rioult-Pedotti MS, Luft AR. Dopamine in motor cortex is necessary for skill learning and synaptic plasticity. *PLoS ONE*. 2009; 4(9):e7082.10.1371/journal.pone.0007082 [PubMed: 19759902]
- Morris G, Arkadir D, Nevet A, Vaadia E, Bergman H. Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron*. 2004; 43(1):133–143.10.1016/j.neuron.2004.06.012 [PubMed: 15233923]
- Pascual-Leone A, Grafman J, Hallett M. Modulation of cortical motor output maps during development of implicit and explicit knowledge. *Science (New York, NY)*. 1994; 263(5151):1287–1289.

- Pascual-Leone A, Valls-Solé J, Wassermann EM, Brasil-Neto J, Cohen LG, Hallett M. Effects of focal transcranial magnetic stimulation on simple reaction time to acoustic, visual and somatosensory stimuli. *Brain: a Journal of Neurology*. 1992; 115(Pt 4):1045–1059. [PubMed: 1393501]
- Rosenthal CR, Roche-Kelly EE, Husain M, Kennard C. Response-dependent contributions of human primary motor cortex and angular gyrus to manual and perceptual sequence learning. *Journal of Neuroscience*. 2009; 29(48):15115–15125.10.1523/JNEUROSCI.2603-09.2009 [PubMed: 19955363]
- Schultz W. Behavioral dopamine signals. *Trends in Neurosciences*. 2007; 30(5):203–210.10.1016/j.tins.2007.03.007 [PubMed: 17400301]
- Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science (New York, NY)*. 1997; 275(5306):1593–1599.
- Schultz W, Preusschoff K, Camerer C, Hsu M, Fiorillo CD, Tobler PN, Bossaerts P. Explicit neural signals reflecting reward uncertainty. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*. 2008; 363(1511):3801–3811.10.1098/rstb.2008.0152
- Selen LPJ, Shadlen MN, Wolpert DM. Deliberation in the motor system: reflex gains track evolving evidence leading to a decision. *Journal of Neuroscience*. 2012; 32(7):2276–2286.10.1523/JNEUROSCI.5273-11.2012 [PubMed: 22396403]
- Suzuki M, Kirimoto H, Sugawara K, Oyama M, Yamada S, Yamamoto JI, et al. Motor cortex-evoked activity in reciprocal muscles is modulated by reward probability. *PLoS ONE*. 2014; 9(6):e90773.10.1371/journal.pone.0090773 [PubMed: 24603644]
- Thabit MN, Nakatsuka M, Koganemaru S, Fawi G, Fukuyama H, Mima T. Momentary reward induce changes in excitability of primary motor cortex. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*. 2011; 122(9):1764–1770.10.1016/j.clinph.2011.02.021 [PubMed: 21439903]
- Tobler PN, Fiorillo CD, Schultz W. Adaptive coding of reward value by dopamine neurons. *Science (New York, NY)*. 2005; 307(5715):1642–1645.10.1126/science.1105370
- Wassermann EM, Zimmermann T. Transcranial magnetic brain stimulation: therapeutic promises and scientific gaps. *Pharmacology & Therapeutics*. 2012; 133(1):98–107.10.1016/j.pharmthera.2011.09.003 [PubMed: 21924290]
- Wickens JR, Reynolds JNJ, Hyland BI. Neural mechanisms of reward-related motor learning. *Current Opinion in Neurobiology*. 2003; 13(6):685–690. [PubMed: 14662369]
- Wilkinson L, Teo JT, Obeso I, Rothwell JC, Jahanshahi M. The contribution of primary motor cortex is essential for probabilistic implicit sequence learning: evidence from theta burst magnetic stimulation. *Journal of Cognitive Neuroscience*. 2010; 22(3):427–436.10.1162/jocn.2009.21208 [PubMed: 19301999]
- Zhuang P, Dang N, Waziri A, Gerloff C, Cohen LG, Hallett M, Warzeri A. Implicit and explicit learning in an auditory serial reaction time task. *Acta Neurologica Scandinavica*. 1998; 97(2):131–137. [PubMed: 9517864]
- Zhuang P, Toro C, Grafman J, Manganotti P, Leocani L, Hallett M. Event-related desynchronization (ERD) in the alpha frequency during development of implicit and explicit learning. *Electroencephalography and Clinical Neurophysiology*. 1997; 102(4):374–381. [PubMed: 9146500]
- Ziemann U, Bruns D, Paulus W. Enhancement of human motor cortex inhibition by the dopamine receptor agonist pergolide: evidence from transcranial magnetic stimulation. *Neuroscience Letters*. 1996; 208(3):187–190. [PubMed: 8733301]

Highlights

- Primary motor cortex (M1) excitability responds to reward-related signals.
- We test if M1 excitability response reflects probability or uncertainty.
- M1 excitability changes across reward probability, not uncertainty.
- The direction of change depends on larger task context.

Figure 1.

Reward processing in M1 may reflect either outcome uncertainty or expected reward value. Outcome uncertainty is reflected by an inverted U-shaped function with maximal response under maximal outcome uncertainty (50/50). Expected reward value is the product of the reward probability and reward magnitude and is reflected by a linear increase in response to increased reward probability.

Figure 2.

(A) Reward probability and outcome uncertainty associated with each stimulus for the Find and Avoid search conditions for the *Circle* instruction group. Stimulus was a filled square (left), two filled squares (middle), or an empty square (right). Probability values to the right indicate the probabilities of choosing correctly, under the Avoid and Find search conditions, respectively. Separate groups participated in the *Circle* and *No Circle* instruction groups. For the “No Circle” instruction group, the same stimuli were presented, however the trial contingencies were reversed from what is shown here for the Avoid and Find search conditions. (B) Time course of a trial. For each trial, participants were instructed either to Find or Avoid the designated target stimulus for that experiment. After the trial instruction was given, the initial stimuli appeared on the screen; after 1 s, participants were prompted to indicate their selection on a button box; after a 500-ms delay, the selection was fed back to the participant by the appearance of a red outline around the selected square (dashed gray line in the figure). TMS was delivered to the left M1 250 ms after stimulus onset and 750 ms before the response prompt. Finally, the location of the circle, if present, was revealed, accompanied by a message indicating whether the response was correct and whether a prize was awarded.

Figure 3.
Mean motor-evoked potential (MEP) amplitude (mV), for each level of Reward Probability.
* = significant difference at $p < .05$; Error bars indicate SEM.

Figure 4.

Mean response times (RT) for each level of reward probability. * = significant difference at $p < .05$; Error bars indicate SEM.