

RESEARCH ARTICLE | *Control of Movement*

Vigor of reaching movements: reward discounts the cost of effort

 Erik M. Summerside,¹  Reza Shadmehr,² and Alaa A. Ahmed¹

¹*Department of Integrative Physiology, University of Colorado, Boulder, Colorado; and* ²*Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland*

Submitted 8 December 2017; accepted in final form 7 March 2018

Summerside EM, Shadmehr R, Ahmed AA. Vigor of reaching movements: reward discounts the cost of effort. *J Neurophysiol* 119: 2347–2357, 2018. First published March 14, 2018; doi:10.1152/jn.00872.2017.—Making a movement may be thought of as an economic decision in which one spends effort to acquire reward. Time discounts reward, which predicts that the magnitude of reward should affect movement vigor: we should move faster, spending greater effort, when there is greater reward at stake. Indeed, saccade peak velocities are greater and reaction-times shorter when a target is paired with reward. In this study, we focused on human reaching and asked whether movement kinematics were affected by expectation of reward. Participants made out-and-back reaching movements to one of four quadrants of a 14-cm circle. During various periods of the experiment only one of the four quadrants was paired with reward, and the transition from reward to nonreward status of a quadrant occurred randomly. Our experiment design minimized dependence of reward on accuracy, granting the subjects wide latitude in self-selecting their movement speed, amplitude, and variability. When a quadrant was paired with reward, reaching movements had a shorter reaction time, higher peak velocity, and greater amplitude. Despite this greater vigor, movements toward the rewarded quadrant suffered from less variability: both reaction times and reach kinematics were less variable when there was expectation of reward. Importantly, the effect of reward on vigor was specific to the movement component that preceded the time of reward (outward reach), not the movement component that followed it (return reach). Our results suggest that expectation of reward not only increases vigor of human reaching but also decreases its variability.

NEW & NOTEWORTHY Movements may be thought of as an economic transaction where the vigor of the movement represents the effort that the brain is willing to expend to acquire a rewarding state. We show that in reaching, reward discounts the cost of effort, producing movements with shorter reaction time, higher velocity, greater amplitude, and reduced reaction-time variability. These results complement earlier observations in saccades, suggesting a common principle of economics across modalities of motor control.

effort; reaching movements; reward; variability; vigor

INTRODUCTION

Imagine you are sitting at your desk and the phone rings, but you do not recognize the number. You reach for the phone and answer to find it is an old friend. A few weeks later, the friend

calls again, but this time you recognize the number. Again, you reach for the phone, excited to hear how they have been. Both scenarios require execution of a reaching movement. With the assumption that the physical constraints of reaching (i.e., initial arm configuration and end-point goal) are identical, will the reaching movements be the same?

Early motor control models suggested that kinematics of reaching movements might be described through minimizing costs such as end-point variability (Harris and Wolpert 1998; van Beers et al. 2004) and energy consumption (Alexander 1997), but they commonly relied on simplifications that included fixed movement duration. With the use of this framework, movement kinematics were dictated by minimizing the combined weight of these costs (Burdet et al. 2001; Wang et al. 2016). If we apply these models to our example of answering the phone, they predict invariant kinematics in the two situations.

However, if we assume that the purpose of a movement is to acquire a more rewarding state, and that time discounts the value of reward, then movements carry a cost of time (Shadmehr et al. 2010). In this framework, slower movements diminish reward. As a result, reward justifies expenditure of effort to arrive at the goal earlier. Recent contributions have considered this idea by assigning a utility to each action that combines measures of weighted effort and reward (Berret and Jean 2016; Haith et al. 2012; Niv et al. 2007; Rigoux and Guigon 2012; Shadmehr et al. 2016). As a result, the optimal level of vigor (defined as movement speed as a function of distance) is an interaction between optimization of two competing factors: the desire to get reward sooner, balanced via payment of higher effort. According to these models, when you recognize the phone number and expect a pleasant conversation, you will reach with greater vigor, spending more effort to answer the phone sooner.

Experimental evidence has demonstrated that animals produce faster movements when they expect reward. Nonhuman primates make faster saccadic eye movements toward targets paired with juice compared with those same targets without juice (Takikawa et al. 2002). Similarly, humans make faster saccades when those movements are paired with explicit monetary rewards (Manohar et al. 2015, 2017) and also implicit reward, such as when the movement is directed toward a more informative target (Xu-Wilson et al. 2009). As humans deliberate between two rewarding stimuli, saccade velocity is higher when the eyes gaze at the preferred stimulus (Reppert et al. 2015). Furthermore, there is evidence that people who exhibit

Address for reprint requests and other correspondence: E. M. Summerside, Neuromechanics Laboratory, Dept. of Integrative Physiology, Univ. of Colorado Boulder, 354 UCB, Boulder, CO 80309-0354 (e-mail: erik.summerside@colorado.edu).

high temporal discounting in decision making also make more vigorous saccades, suggesting that even in the absence of explicit reward, the cost of time is higher in people who move more vigorously (Choi et al. 2014). Taken together, these experiments demonstrate that in the saccadic system, reward modulates vigor of movements.

The effect of reward on arm movements is less understood. In nonhuman animals, two reports indicated that reward (juice/food) encouraged faster movements (Mosberger et al. 2016; Opris et al. 2011), whereas one report did not make this observation (Pasquereau et al. 2007). In humans, one report stated that reaching was faster when the goal object had higher emotional valence (Esteves et al. 2016).

In the current study we considered a reaching task to test whether reward discounted effort expenditure. Reward may modulate movement vigor, but increased vigor often coincides with reduced accuracy, which can reduce probability of reward. To address this potential confound, our task minimized dependence of reward on accuracy: rather than reaching to a point, participants reached to one of four quadrants. As a result, they had wide latitude in selecting movement velocity, trajectory, and amplitude. When the quadrant was paired with reward, the participants responded by increasing vigor: they reached sooner, with higher velocity, shorter duration, and greater amplitude. Interestingly, we also observed that increased vigor coincided with reduced variability, demonstrating that expectation of reward not only increased vigor but also promoted consistency.

MATERIALS AND METHODS

Participants. Right-handed participants ($n = 20$), naive to the experiment (age: 26 ± 4 yr, mean \pm SD; 10 men and 10 women) gave

written informed consent approved by the University of Colorado Institutional Review Board before participating in the experiment.

Task. Participants were seated in a chair that limited trunk movement, and they held the handle of a robotic arm with their right hand (Shoulder-Elbow Robot; Interactive Motion Technologies). Using the handle, they controlled the location of a cursor that was projected on an LCD monitor mounted in front of them at eye level (Fig. 1A). The task was begun by placing the cursor (diameter = 0.6 cm) in the center of a home circle (diameter = 0.9 cm). After the cursor was maintained in the home circle for 150 ms, the visual feedback of the home circle was extinguished and the computer simultaneously delivered an audiovisual cue to begin the trial. The auditory component of the cue was a short beep (50 ms at 110 Hz followed by 50 ms at 220 Hz), and the visual component was the illumination of a large red ring (radius = 14 cm) that was displayed with its center at the home circle. The ring included a marker that indicated the quadrant that served as the goal of the movement. The marker was placed in one of four possible locations (45° , 135° , 225° , or 315° from right horizontal) to specify the intended quadrant (Fig. 1, B and C). The sole criterion for success was that the cursor crossed the ring within a 100° arc centered on the marker. As the reach began, visual feedback of the cursor was blanked. Once the invisible cursor crossed the outer ring within the quadrant, the outer ring changed color from red to gray, indicating that the trial was completed and that the invisible cursor should be brought back to center. We refer to the location where the invisible cursor crossed the ring as the crossing point. There was no time limit to complete the trial, and no instructions were provided regarding a desired reach velocity. The cursor remained invisible until the return aspect of the movement when it entered a region within 9 cm of the center of the home circle. At this point, the cursor and home circle were again made visible and a new trial could begin.

If the quadrant was associated with reward and the invisible cursor crossed within the 100° reward region centered on the marker, the subjects experienced a pleasing sound (50 ms at 880 Hz followed by 50 ms at 3,520 Hz) and a visual animation of the ring at the moment

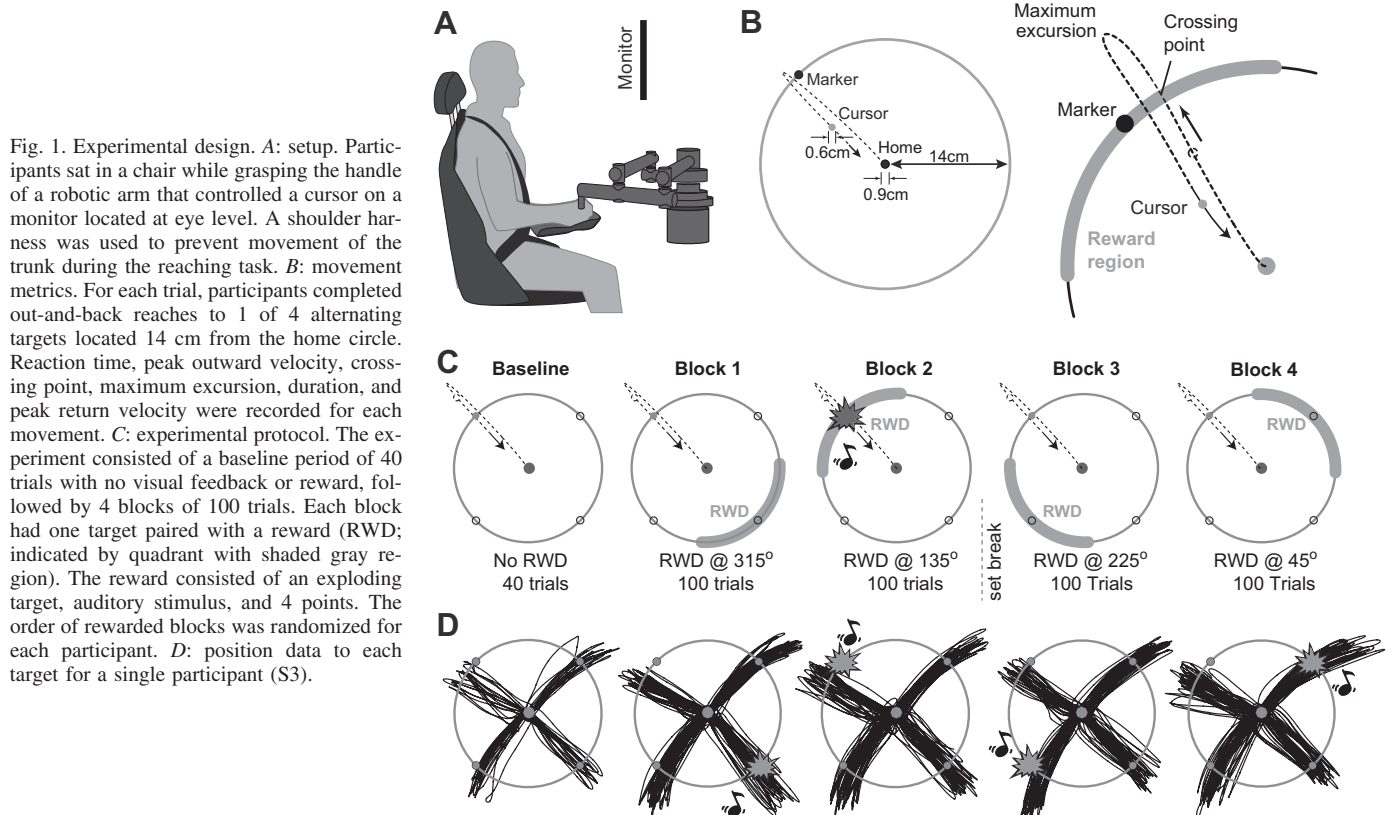


Fig. 1. Experimental design. **A:** setup. Participants sat in a chair while grasping the handle of a robotic arm that controlled a cursor on a monitor located at eye level. A shoulder harness was used to prevent movement of the trunk during the reaching task. **B:** movement metrics. For each trial, participants completed out-and-back reaches to 1 of 4 alternating targets located 14 cm from the home circle. Reaction time, peak outward velocity, crossing point, maximum excursion, duration, and peak return velocity were recorded for each movement. **C:** experimental protocol. The experiment consisted of a baseline period of 40 trials with no visual feedback or reward, followed by 4 blocks of 100 trials. Each block had one target paired with a reward (RWD; indicated by quadrant with shaded gray region). The reward consisted of an exploding target, auditory stimulus, and 4 points. The order of rewarded blocks was randomized for each participant. **D:** position data to each target for a single participant (S3).

the invisible cursor passed the outer ring. The visual animation paired with reward consisted of the entire outer ring rapidly flashing yellow and then disappearing completely (transition from red to yellow to extinguished = 50 ms). The cursor remained invisible throughout the initial aspects of the return movement, after the reward was delivered, to guarantee that the visual qualities of reward were not obstructed by the visual feedback of the cursor. At completion of the trial, participants also received 4 points. The cumulative points were displayed in the upper right corner of the monitor. Participants were not informed of the number of trials they would be performing, only that the experiment would take roughly 1 h. Furthermore, each participant was informed that the compensation for participating in the study session (\$15) was fixed and not contingent on the amount of points received from rewarded trials or any other measure of task performance.

We assumed that participants planned reaching movements with the goal of wanting to maximize the chance of successfully completing the task and even in the absence of penalty would reach toward the center of the cued quadrant (Trommershäuser et al. 2003). Previous work had demonstrated that, on average, healthy people exhibited regular errors of up to $9^\circ \pm 3^\circ$ while holding a robotic arm and aiming to targets at distances of 10 cm (Smith and Shadmehr 2005). Based on this finding, the size of our rewarded region was more than five times the expected error of reaching such that even in the presence of a persistent error, nearly all attempts should fall within the intended zone. Therefore, an important factor in our experiment design was an attempt to remove accuracy as one of the constraints typically associated with reward.

On arrival to the laboratory, all participants were seated and allowed ~40 trials to familiarize themselves with the robotic manipulandum. All familiarization trials occurred in the absence of reward and with full visual feedback of the cursor during both outward and return components of the movement. On conclusion of the familiarization phase, the experimental protocol consisted of a further 440 reaching trials. At the beginning of the protocol, the participants were informed they would no longer receive visual feedback of the cursor during their reach for the remainder of the experiment. They were also instructed that some trials would now be paired with a reward and that as long as they reached toward the indicated quadrant, they would receive the full reward. Importantly, participants were not told that a direction would be consistently rewarded in a block, nor were they made aware of the underlying block structure.

The first 40 experimental trials occurred in the absence of reward (baseline, Fig. 1C). Following baseline, reward was introduced in one of the four directions (*blocks 1–4*, Fig. 1C). A reach was rewarded if it was within a 100° zone centered on the marker and the direction was paired with reward. There was no feedback of any kind regarding accuracy of the movement: the only feedback was reward, and its only criterion was whether the reach was within a 100° zone centered on the marker. The location of the reward zone was constant within each block of 100 consecutive trials (25 toward the rewarded location) and then changed to a new location for the next 100 trials. There was a short 30-s break between *blocks 2* and *3*. The order of rewarded quadrants was randomized for each participant. For 16 participants, trial-by-trial marker presentation within each block was randomized, meaning that in *blocks 1–4*, there was, on average, a 25% chance that the next trial would be in the rewarded quadrant, even if the previous trial was also rewarded. The remaining four participants had a pseudorandomized presentation of trials such that no rewards were presented consecutively. Participants never received instruction regarding the location of future reward trials, how reward location was distributed across blocks, or when a new block with a new rewarded quadrant began.

Data analysis. Handle position and velocity were recorded at 200 Hz. Reaction time was quantified as the time from the audiovisual start stimulus to movement onset. Movement onset was established via radial acceleration (0.0001 m/s^2) and radial velocity thresholds (0.05 m/s). Distance of the crossing point referred to its distance

relative to the marker, which was reported as the signed difference in degrees, measured from the right horizontal, between the center of the quadrant and where the hand crossed the outer ring. Maximum excursion was calculated as the maximum Euclidean distance between the start marker and the cursor, measured over the course of the entire trial. Peak outward velocity was calculated as the maximum instantaneous radial velocity measured between movement onset and instant of maximum excursion. Movement duration was calculated as the elapsed time between when the cursor crossed a position threshold of 0.3 cm and the crossing point. Peak return velocity was calculated as the maximum instantaneous radial velocity measured after the instant of maximum excursion.

Trials were removed from analysis if reaction times were $>700 \text{ ms}$. Across all participants this accounted for exclusion of 0.43% of trials (43/8,000 trials). In addition, we found that in only 0.03% of trials (2/8,000 trials), the absolute value of the crossing-point distance of the reaching movement was more than 50° , signifying it fell outside the potential reward zone. Therefore, the large size of the reward region allowed for more than 99% of the trials to be potentially rewarding. Errant movements (absolute crossing-point distance of more than 50°) were excluded from analysis.

Experimental design and statistical analysis. The location of the reward zone was reassigned after every 100 trials. In each period of 100 trials, there were 25 movements toward each quadrant. To determine the effect of reward on the current movement, we measured how reaching was altered in the block when that movement was rewarded compared with blocks when that same action was not rewarded. Peak outward velocity, reaction time, maximum excursion, duration, crossing point, and peak return velocity were compared between the rewarded period (100 trials) and nonrewarded periods (300 trials) for each participant. To measure the effects of reward on movement variability, we calculated the variance of peak velocity, reaction time, maximum excursion, duration, and crossing point for each quadrant when that quadrant was rewarded and compared it with the mean variance across the remaining three blocks when that same quadrant was not rewarded. We measured the effect of reward using a two-way repeated-measures analysis of variance (ANOVA) based on block number (discrete), whether the target was rewarded (binary), and a reward \times block interaction. Differences in movements toward each quadrant were compared using a two-way repeated-measures ANOVA based on quadrant location (numbered counterclockwise beginning with the upper right quadrant), reward, and a reward \times quadrant interaction. We used two-sided paired *t*-tests to compare movements toward rewarded quadrants and movements to nonrewarded quadrants in the trials immediately before and after a rewarded trial. Effects of repeating movements to the same quadrant were probed using a repeated-measures ANOVA based on whether the quadrant of the current movement was the same as the previous trial as well as whether the current quadrant was rewarded.

All statistical thresholds were conducted at a significance level of 0.05. All uncorrected *P* values reaching statistical significance were corrected for multiple comparisons using the Holm-Bonferroni method. ANOVAs and paired *t*-tests were corrected for a total of five comparisons, established on the basis of the number of measured behavioral responses (peak velocity, reaction time, crossing point, maximum excursion, and duration). Post hoc comparisons on the effect of blocks and quadrants were corrected for a total of six comparisons. Descriptive statistics are reported as means \pm SE.

RESULTS

Participants ($n = 20$) made a self-paced out-and-back reaching movement without visual feedback toward a marker that was positioned at 14 cm in 1 of 4 quadrants (Fig. 1B). In each block of trials, only one of the quadrants (Fig. 1C) was associated with reward (a pleasing sound and animation, as well as 4 points). Figure 1D illustrates reach trajectories for a

single participant in various blocks. At the moment that the unseen cursor crossed the 14-cm ring, the mean absolute distance (for each participant) of the crossing point from the marker was $9.3^\circ \pm 1^\circ$. The sole criterion for success was that on the outward component of the movement, the unseen cursor crossed within a 100° arc centered on the marker. As a result, more than 99.9% of the movements across subjects crossed the outer ring within the potential reward zone. We asked whether expectation of reward altered movement preparation (reaction time) and movement execution (velocity, extent, and variability).

Effect of reward on reach kinematics. We began our analysis by considering how the subjects reacted to presentation of the marker, which acted as the cue to reach to the quadrant. To quantify the effects of the marker appearing in a rewarded quadrant vs. nonreward quadrants, we computed the reaction-time distribution in each condition and then computed a within-subject difference measure (Fig. 2A). This difference measure was calculated for each individual as the probability density of reaction time for all rewarded trials (bins = 5 ms) minus the probability density of all nonrewarded trials, with the difference measure then averaged across participants. It appeared that reward shifted the mode of the distribution earlier and also reduced the variance of reaction-time distribution.

To quantify the within-subject change in the distribution of reaction times, we constructed a delta plot (Ridderinkhof et al. 2005), as shown in Fig. 2D. For each participant and each condition (reward and nonreward trials), we ordered the reaction times from shortest to longest and divided them into 20% quantiles. We computed the mean of each quantile and then measured the within-subject change in the quantile mean due to condition (reward minus nonreward trials). We found that for all quantiles, the change was negative, suggesting that reward reduced reaction times in all ranges of responses. Furthermore, the negative slope indicated that the reaction-time distribution for reward was steeper (less variable) than the nonrewarded distribution, implying a reduced variance. In summary, reward appeared to have two effects on the reaction-time distribution: it shifted the mode of the distribution earlier, and it reduced the variance of the distribution.

We next considered the effects of reward on the kinematics of the reach. We computed radial position and velocity of the hand as a function of time (Fig. 2, B and C) and found that in the rewarded condition, the subjects reached farther (Fig. 2B, right, peak of red curve vs. blue curve) and faster (Fig. 2C, right).

To better characterize the effects of reward, we computed for each participant the change in various parameters of movement when a quadrant was paired with reward compared with when the same quadrant was not paired with reward. In the presence of reward, mean of the reaction times decreased by $5.21 \pm 0.79\%$ (15.20 ± 0.10 ms, $P < 0.001$; Fig. 3A), variance of the reaction times decreased by an average of $24.0 \pm 6.32\%$ ($P = 0.006$; Fig. 3B), outward peak velocity increased by $1.87 \pm 0.88\%$ (0.78 ± 0.01 cm/s, $P = 0.044$; Fig. 3C), maximum excursion increased by $4.14 \pm 0.57\%$ (0.73 ± 0.01 cm, $P < 0.001$; Fig. 3D), and movement duration decreased by $4.56 \pm 1.05\%$ (26.50 ± 8.3 ms, $P = 0.002$; Fig. 3E, where duration refers to time to the crossing point). (All P values reflect corrections for multiple comparisons using the Holm-Bonferroni method). In contrast, we found no effect of reward on mean crossing point ($P = 0.613$). That is, the hand crossed

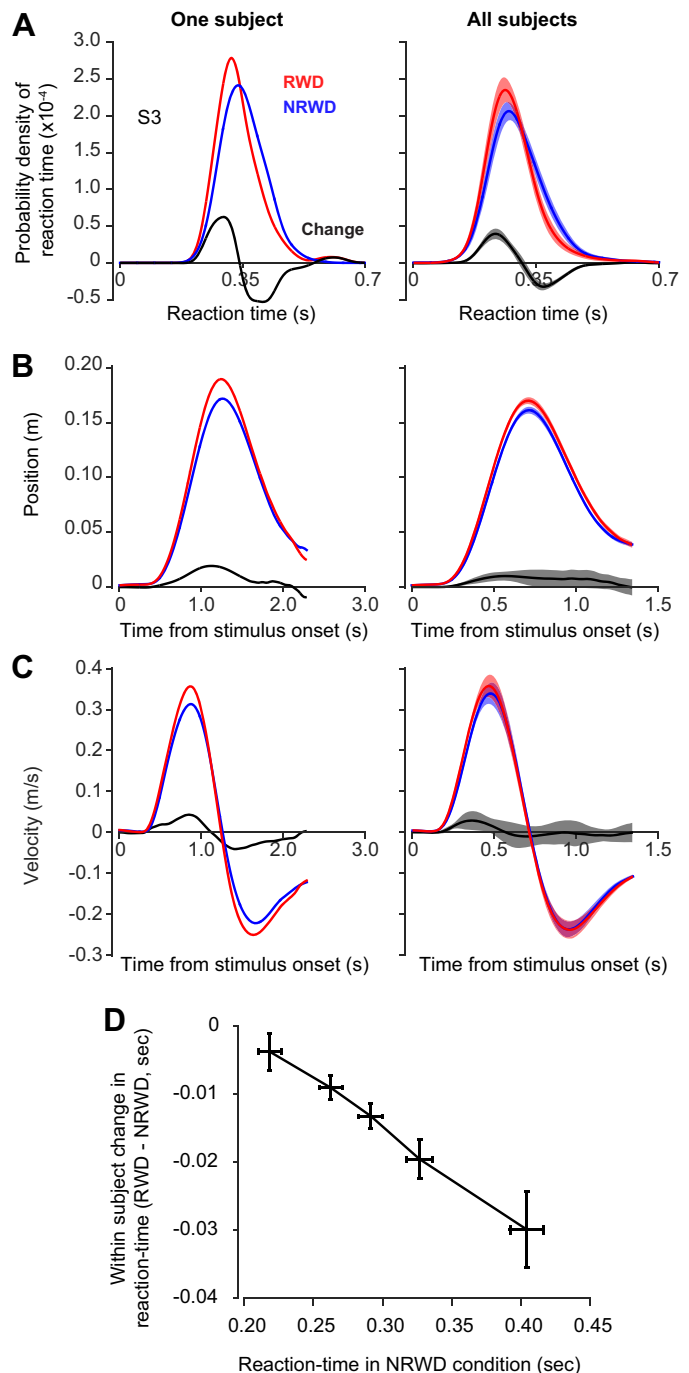


Fig. 2. Movement characteristics. A: probability distribution of reaction time was estimated for each subject in each condition using a nonparametric approach (bin size = 5 ms). The change in reaction time is a within-subject measure. Mode of the reaction time appeared to shift earlier, and the variance appeared to decrease. B: radial position trajectory and the change in radial position as a function of time. The hand appeared to reach farther in the rewarded condition. C: radial velocity and the change in radial velocity as a function of time. The hand appeared to reach faster in the rewarded condition. Because of the range of movement durations selected across participants, group averages are displayed up to the point of the shortest individual curve. Shaded regions are \pm SE. D: delta plot of reaction time across 20% quantiles. For each subject, reaction times in each condition were rank ordered and sorted into 20% quantiles. Values along the x-axis represent mean reaction time for no reward (NRWD) at each quantile. Values along the y-axis represent the change in the mean from reward (RWD) to NRWD condition. Negative values indicate that reward decreased reaction times, and the negative slope suggests that reward reduced the variance of reaction times. Error bars are \pm SE.

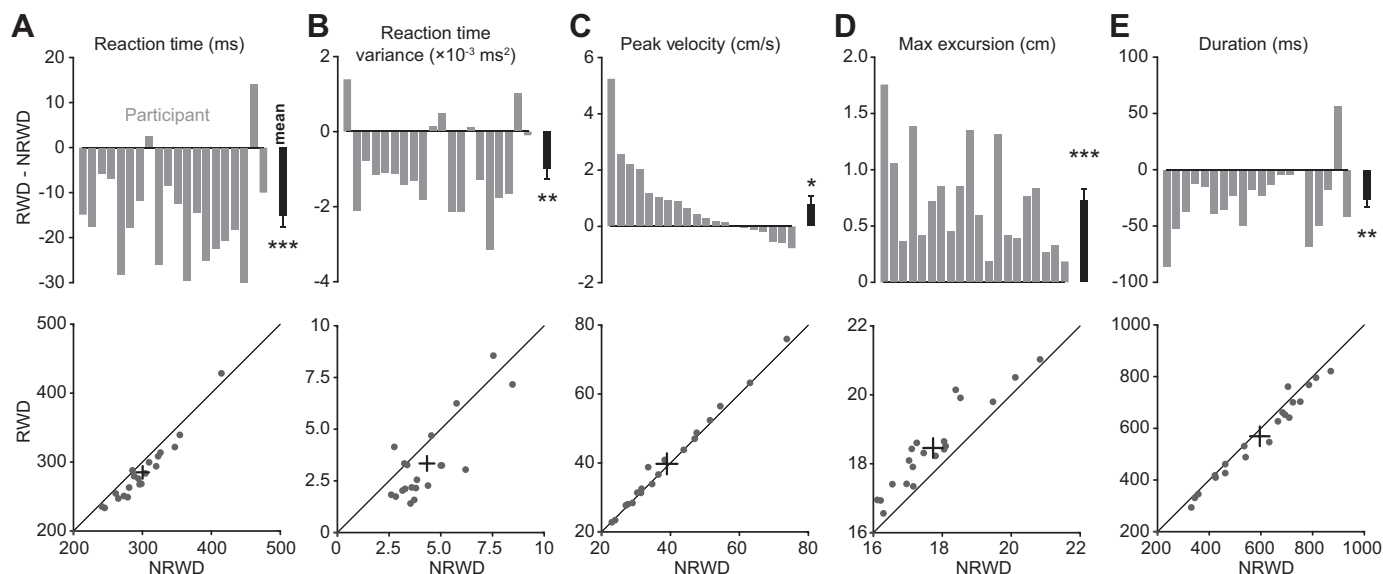


Fig. 3. Within-subject measures on the effects of reward. Reward-dependent changes in mean of the reaction times (A), variance of the reaction times (B), peak velocity of the outward movement (C), maximum excursion (D), and duration of the outward movement (E) are shown. In bar plots, gray bars represent within-subject change and black bars represent group means \pm SE. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. Participants were ranked according to their relative change in peak velocity (C). Differences represent reward minus nonreward. In scatter plots, black dots represent individual participants and black lines represent group means \pm SE.

the outer ring at a location (with respect to the marker) that was, on average, unchanged with reward. In contrast, reward decreased crossing-point variance by an average of $10.10 \pm 4.18\%$. However, this effect was lost when corrected for multiple comparisons (uncorrected $P = 0.024$, corrected $P = 0.096$). We observed no effect of reward on the variance of peak velocity, maximum excursion, or duration (all $P > 0.05$).

Effect of reward across blocks. Previous work had noted that if subjects repeatedly made saccadic eye movements toward the same stimulus, the movements tended to become slower (Chen-Harris et al. 2008; Xu-Wilson et al. 2009). In this study, we observed the opposite tendency: as the experiment progressed, participants increased the speed of their reaching movements [RM-ANOVA, main effect of block, peak outward velocity: $F(3,57) = 8.748$, $P < 0.001$; Fig. 4A]. Similarly, progression of the experiment coincided with a reduction in the mean of the reaction times [$F(3,57) = 10.500$, $P < 0.001$; Fig. 4B] as well as the variance of the reaction times [$F(3,57) = 4.692$, $P = 0.005$; Fig. 4F]. As the experiment progressed, duration of the reaching movements decreased [$F(3,57) = 9.478$, $P < 0.001$; Fig. 4D]. There were no changes across blocks for maximum excursion [$F(3,57) = 1.873$, $P = 0.144$; Fig. 4C] and no changes across blocks for mean crossing-point distance [$F(3,57) = 0.662$, $P = 0.579$; Fig. 4E] or the variance in crossing-point distance [$F(3,57) = 0.356$, $P = 0.785$].

Importantly, the effect of reward on all movement parameters was consistent throughout the duration of the experiment [RM-ANOVA, block \times reward interaction, peak outward velocity: $F(3,57) = 0.509$, $P = 0.678$; reaction time: $F(3,57) = 1.344$, $P = 0.269$; maximum excursion: $F(3,57) = 1.484$, $P = 0.229$; duration: $F(3,57) = 0.515$, $P = 0.674$; crossing point: $F(3,57) = 0.602$, $P = 0.616$]. In summary, with the progression of the experiment, reach velocities tended to increase and reaction times tended to decrease. However, within-subject effects of reward remained consistent, influencing peak outward velocity [main effect of reward on peak velocity: $F(1,19) = 6.273$,

$P = 0.044$; Fig. 4A], mean reaction time [$F(1,19) = 38.47$, $P < 0.001$; Fig. 4B], maximum excursion [$F(1,19) = 51.77$, $P < 0.001$; Fig. 4C], movement duration [$F(1,19) = 15.95$, $P = 0.002$; Fig. 4D], and variance of reaction time [$F(1,19) = 14.5$, $P = 0.010$; Fig. 4F].

Effect of reward across quadrants. Movements to each of the four quadrants required a unique combination of elbow and shoulder torques. This difference in joint torque combinations introduced the possibility that the amount of effort required for reaching was dependent on quadrant location (Schweighofer et al. 2015) and that reward may have affected movements differently at each location. Indeed, movement characteristics differed depending on which quadrant was cued. There was a main effect of quadrant on peak velocity [$F(3,57) = 8.68$, $P < 0.001$], maximum excursion [$F(3,57) = 9.43$, $P < 0.001$], and crossing point [$F(3,57) = 42.12$, $P < 0.001$]. Post hoc comparisons indicated that peak velocity was slowest in *quadrant 1* (Q1: 36.70 ± 2.84 cm/s, Q2: 40.64 ± 3.35 cm/s, Q3: 38.24 ± 2.67 cm/s, Q4: 41.07 ± 3.49 cm/s, $P_{1,2} = 0.002$, $P_{1,3} = 0.020$, $P_{1,4} = 0.002$, $P_{2,3} = 0.046$, $P_{2,4} = 0.606$, $P_{3,4} = 0.049$). Maximum excursion was shortest for movements toward *quadrant 2* (Q1: 18.15 ± 0.32 cm, Q2: 17.02 ± 0.30 cm, Q3: 18.21 ± 0.30 cm, Q4: 18.29 ± 0.37 cm, $P_{1,2} < 0.001$, $P_{1,3} = 0.773$, $P_{1,4} = 0.640$, $P_{2,3} < 0.001$, $P_{2,4} = 0.003$, $P_{3,4} = 0.768$). Both crossing-point distance mean and variance were affected by quadrant location. Mean crossing-point distance was most positive (counterclockwise from quadrant center) in *quadrant 3* and most negative in *quadrant 4* (Q1: $7.66^\circ \pm 1.46^\circ$, Q2: $2.27^\circ \pm 1.66^\circ$, Q3: $14.63^\circ \pm 1.28^\circ$, Q4: $-4.37^\circ \pm 0.94^\circ$, $P_{1,2} = 0.038$, $P_{1,3} < 0.001$, $P_{1,4} < 0.001$, $P_{2,3} < 0.001$, $P_{2,4} = 0.001$, $P_{3,4} < 0.001$). Variance in crossing point was greater in *quadrant 1* compared with *quadrants 3* and *4*, with all other pairs being indistinguishable (Q1: $33.56^\circ \pm 3.24^\circ$, Q2: $26.39^\circ \pm 3.00^\circ$, Q3: $24.11^\circ \pm 2.20^\circ$, Q4: $22.65^\circ \pm 2.48^\circ$, $P_{1,2} = 0.099$, $P_{1,3} = 0.022$, $P_{1,4} = 0.022$, $P_{2,3} < 0.532$, $P_{2,4} < 0.213$, $P_{3,4} < 0.565$).

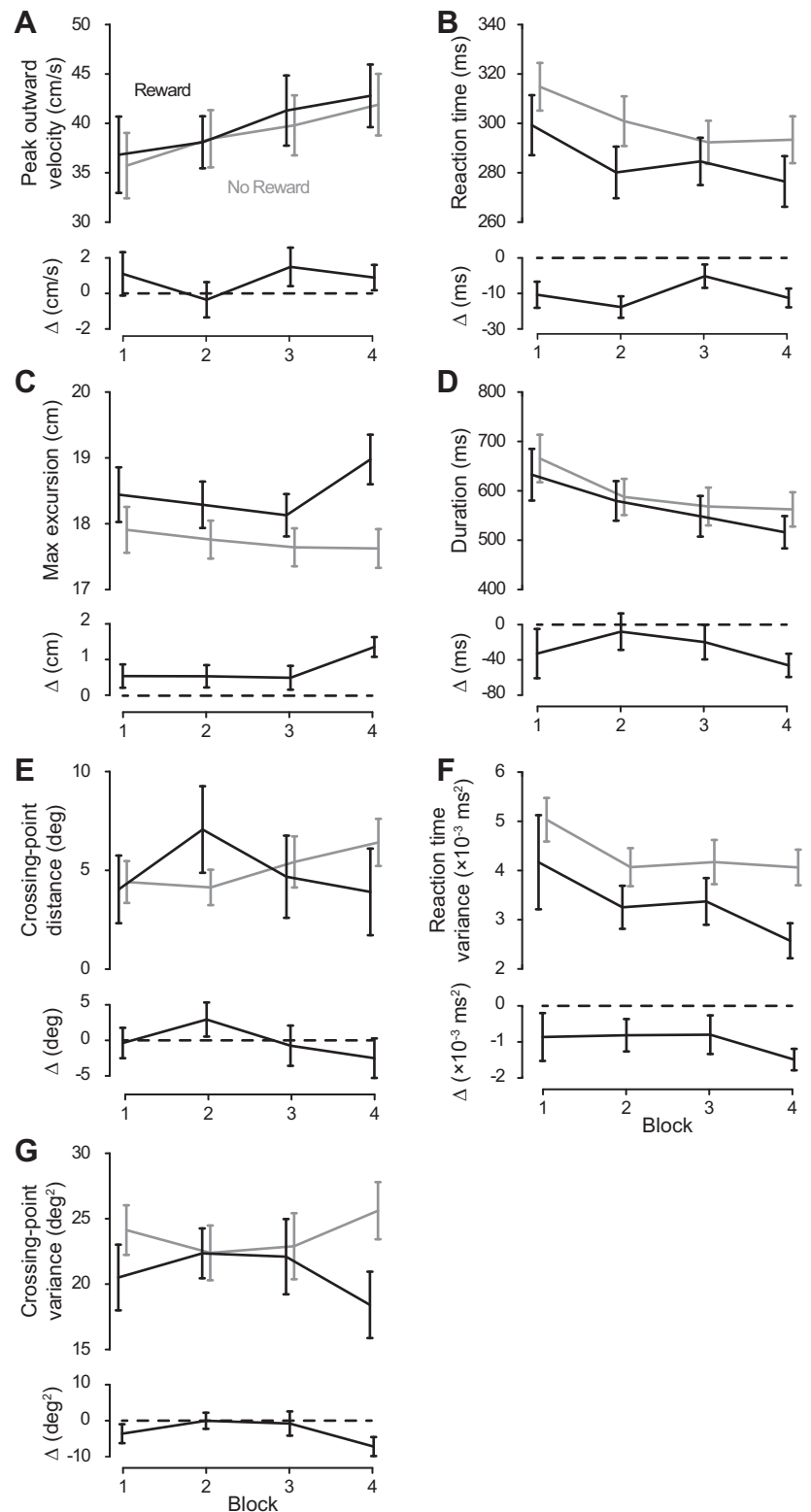


Fig. 4. Effect of block number on reaching movements. Effects of block on peak outward velocity (A), reaction time (B), maximum excursion (C), duration (D), crossing-point distance (E), reaction-time variance (F), and crossing-point variance (G) are shown. Black and gray lines represent rewarded and nonrewarded trials, respectively. Bars are means \pm SE. Graphs have been slightly offset horizontally to improve contrast. Results from post hoc comparisons regarding the main effect of block are based on averages combining both rewarded and nonrewarded movements. * $P < 0.05$; ** $P < 0.01$. Subset graphs represent within-subject difference (Δ) calculated as reward minus nonreward at each block.

There was no effect of quadrant on reaction-time mean, reaction-time variance, or duration.

Although it was evident that the location of the quadrant affected a few of the movement kinematics, we found no interaction effects between reward and quadrant in any of the measured metrics [peak velocity: $F(3,37) = 1.01$, $P = 0.394$; reaction time: $F(3,37) = 0.23$, $P = 0.878$; maximum excursion:

$F(3,37) = 0.77$, $P = 0.514$; duration: $F(3,37) = 0.13$, $P = 0.942$; crossing point: $F(3,37) = 1.77$, $P = 0.163$]. In summary, the location of the quadrant influenced movement vigor, but the effect of reward was quadrant independent.

Effect of temporal proximity to a rewarding movement. If expectation of reward affected movement vigor, what was the temporal window of these effects? Did increased vigor due to

reward on one trial influence vigor of the subsequent movements? To explore these questions, we compared movements to the rewarded quadrant with the movements that were made immediately before and after, toward other (nonrewarded) quadrants (Fig. 5). We found that compared with the rewarded trial, the immediately preceding nonrewarded trial had reduced outward peak velocity (2-sided paired *t*-test, reward trial compared with previous trial, $P = 0.041$), increased reaction time ($P < 0.001$), reduced excursion ($P < 0.001$), and increased duration ($P = 0.037$). Similarly, the nonrewarded trial immediately following the rewarded trial exhibited reduced peak outward velocity (2-sided paired *t*-test, reward trial compared with subsequent trial, $P = 0.006$), increased reaction time ($P < 0.001$), reduced excursion ($P < 0.001$), and increased duration ($P < 0.001$). The average crossing point was unchanged between reward and surrounding nonrewarded trials (all $P > 0.05$).

We next considered the effects of reward on movement variance and found that reaction-time variance was lower in the rewarded trial compared with the preceding nonrewarded trial ($P = 0.030$). However, this same comparison in variance for crossing point resulted in indistinguishable differences ($P = 0.126$). Trials immediately following reward exhibited increased variance in both reaction time ($P = 0.019$) and crossing point ($P = 0.024$). Therefore, increased vigor and reduced variability were specific to the rewarding target and were not shared with temporally nearby movements to nonrewarding quadrants.

Effect of spatial proximity to a rewarding movement. We tested whether spatial proximity to the rewarded quadrant influenced the vigor of the movements toward adjacent and opposite nonrewarded quadrants. We measured kinematics of nonrewarded movements (reaction time, peak velocity, maximum excursion, and crossing point) when a quadrant was adjacent to reward and compared it with the kinematics when that same quadrant was opposite the reward quadrant. We found no difference between any measures for movements adjacent and opposite reward (all $P > 0.05$).

The large 100° arc for each quadrant meant large deviations from the center would still result in a successfully completed trial. When a rewarded quadrant was adjacent to a cued nonrewarded quadrant, it was possible that the rewarded quadrant could act as a distractor (or attractor) and influence the crossing point for the nonrewarded movement either toward or away from the direction of the rewarded quadrant. To test for the presence of a reward influenced bias, we measured crossing-point distance for each target when the clockwise quadrant was rewarded and compared it with the crossing-point distance when the counterclockwise quadrant was rewarded. For example, when testing for the effect of reward proximity in *quadrant 1*, we averaged crossing-point distance in nonrewarded movements to *quadrant 1* when *quadrant 3* was rewarded and compared it with crossing-point distance in nonrewarded movements to *quadrant 1* when *quadrant 4* was rewarded. Looking at movements to each quadrant independently, we found that there were no differences in crossing-point distances in *quadrant 1* [reward in *quadrant 3* (R_3) = 7.51 ± 1.43 , reward in *quadrant 4* (R_4) = 8.37 ± 1.58 , $P = 0.316$], in *quadrant 2* ($R_3 = 14.57 \pm 1.53$, $R_4 = 14.95 \pm 1.42$, $P = 0.714$), in *quadrant 3* ($R_1 = 2.42 \pm 1.77$, $R_2 = 1.79 \pm$

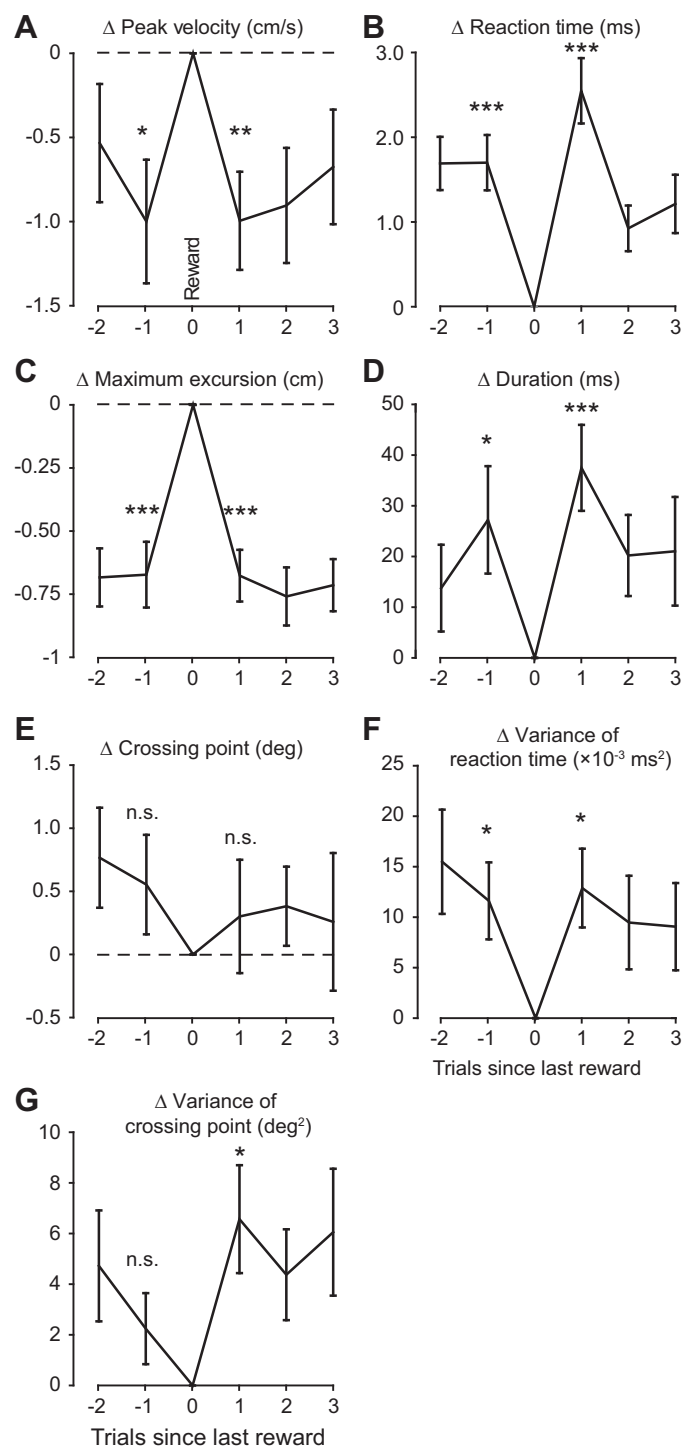


Fig. 5. Trial-to-trial effect of reward. Changes in peak outward velocity (A), reaction time (B), maximum excursion (C), duration (D), crossing-point distance (E), reaction-time variance (F), and crossing-point variance (G) as a result of reward on subsequent and preceding nonrewarded targets are shown. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$ compared with reward; n.s., not significant. All reported values are relative to the rewarded trial. Bars are means \pm SE.

1.87, $P = 0.367$), or *quadrant 4* ($R_1 = -4.70 \pm 0.97$, $R_2 = -4.58 \pm 0.98$, $P = 0.855$).

In summary, we found that the effects of reward were both temporally and spatially specific to the quadrant that was rewarded.

Effects of repeating movements to the same quadrant within a block. Movement history appears to influence arm choice by discounting effort when movements are repeated with the same arm as in preceding trials (Schweighofer et al. 2015). In our current paradigm, there were several instances in which subsequent movements were cued to the same quadrant (~25% of the time). If we assume that reward discounts effort, then our observed effects of reward may be enhanced by movements being repeated. We found that repeating consecutive trials to the same quadrant increased peak velocity [main effect of repetition, peak velocity: $F(1,15) = 18.47$, $P = 0.016$; increased reaction time: $F(1,15) = 15.76$, $P = 0.004$; increased excursion: $F(1,15) = 25.78$, $P < 0.001$; reduced duration: $F(1,15) = 9.692$, $P = 0.014$; and increased reaction-time variance: $F(1,15) = 19.08$, $P = 0.003$], but not crossing-point mean or variance.

Our main question was whether the presence of reward affected these changes. Indeed, we found that the effects of repetition on reaction time, maximum excursion, and duration depended on whether the movements were rewarded or not [reward \times repetition interaction, reaction time: $F(1,15) = 25.00$, $P = 0.001$; maximum excursion: $F(1,15) = 10.49$, $P = 0.010$; duration: $F(1,15) = 13.36$, $P = 0.008$; crossing point: $F(1,15) = 22.05$, $P < 0.001$]. In the presence of reward, repetition further increased maximum excursion (0.93 ± 0.20 cm, $P = 0.001$) and further reduced duration (56 ± 13 ms, $P = 0.002$). Repetition of rewarded trials also increased crossing-point distance ($1.64^\circ \pm 0.44^\circ$, $P = 0.006$). There was no effect of repetition in rewarded trials for peak velocity, reaction time, reaction-time variance, or crossing-point variance (all $P > 0.05$). In the absence of reward, repetition led to longer reaction times (30 ± 6 ms, $P < 0.001$), greater maximum excursion (0.32 ± 0.10 cm, $P = 0.015$), and increased reaction-time variance ($2 \pm 0.5 \times 10^{-3}$ ms², $P = 0.001$). Therefore, repetition of reward led to faster and larger movements.

Effect of reward across segments of a single reaching movement. The auditory and visual cues that indicated success were delivered as the unseen hand crossed the outer circle. However, the movement continued to a self-selected turn-around point, and then the participants brought their hand back to center. Therefore, the trial was composed of two phases of movement (out and back). During rewarded trials the visual target explosion and auditory beep were delivered at the crossing point of the outward movement. That is, acquisition of reward was associated with only the outward phase, not the return phase. Did reward modulate vigor during both movement phases?

We found that the outward peak velocity was, on average, 5.49 ± 0.88 cm/s ($15.59 \pm 2.17\%$) faster than the peak return velocity (2-sided paired t -test, $P < 0.001$). Whereas reward produced an increase in the peak outward velocity of 0.78 ± 0.01 cm/s ($1.87 \pm 0.88\%$, $P = 0.044$; Fig. 3A), the return velocity of the same movement was indistinguishable between rewarded and nonrewarded trials [rewarded: 33.92 ± 2.47 cm/s, nonrewarded: 33.60 ± 2.39 cm/s; ANOVA, main effect of reward: $F(1,19) = 1.273$, $P = 0.273$]. In summary, the effect of reward was specific to the outward phase of the movement (the phase preceding acquisition of reward) and not present in the return phase after reward was acquired.

DISCUSSION

Reaching movements paired with reward exhibited reduced reaction time, higher peak velocity, shorter duration, and larger excursion. Despite increased vigor, movement variability remained largely intact, and in some cases was reduced. These changes were specific to the rewarded trials, with little transfer to temporally or spatially nearby nonrewarded movements.

Reward led to higher vigor. Increases in amplitude and speed of a reaching movement produce increases in the metabolic cost of that movement (Shadmehr et al. 2016). If we view metabolic cost as a proxy for effort, our results suggest that participants were willing to expend more effort when the goal was paired with reward: reaching in rewarding quadrants produced not only a 4% decrease in duration but also a 4% increase in excursion.

Why do subjects reach farther in the rewarded trials? A potential explanation is to increase probability of reward. All movements were rapid, out-and-back shooting movements, but reward was only acquired if the invisible cursor crossed the arc. We found no effect of reward on the proportion of trials where the reach turned around prematurely ($1.90 \pm 0.53\%$ for reward compared with $3.12 \pm 0.61\%$ for no reward, $P = 0.099$). However, it is possible that subjects reached farther to minimize the possibility, albeit unlikely, of turning around before reaching the arc, thus missing the reward.

Our results add to the significant literature demonstrating that movements that are paired with reward result in reduced reaction times (Bendiksbj and Platt 2006; Kawagoe et al. 1998; Milstein and Dorris 2007; Mosberger et al. 2016; Opris et al. 2011; Watanabe and Hikosaka 2005). However, we observed that in addition to reduction in the mean of the reaction time, reward also decreased the variance of the distribution, a fact that has not been noted before.

Reaction time is commonly explained using drift diffusion models (Ratcliff and Rouder 1998) in which evidence toward a decision accumulates until it reaches a threshold. The rate of evidence accumulation is influenced by properties of the stimulus, as well as attention invested toward that stimulus (Milstein and Dorris 2007). In the current paradigm, there were no reward specific environmental cues, suggesting that the strength of the stimulus that beckoned the movement did not affect the rate variable. However, attention may be selective toward rewarded targets (Milstein and Dorris 2007). In our experiment, reward could lead to greater attention toward those quadrants and away from nonrewarded quadrants, allowing for faster accumulation of evidence to initiate movement toward reward.

Our current task involved participants learning to control a robotic manipulandum to move an invisible cursor through alternating quadrants around a central point. In a majority of the trials, the only feedback of the movement was the outer ring changing color from red to gray. In a smaller fraction of trials, the movement outcome was increased with the outer ring flashing yellow while being paired with a short auditory stimulus. By altering the feedback associated with completing each reach, we may have altered the relative sense of agency or contingency between rewarded and nonrewarded movements (Behne et al. 2008; Elsner and Hommel 2004). In an effort to probe how contingency effects movement performance, Karsh and Eitam (2015) had participants press one of several keys on

a keyboard in response to cues. On a proportion of those trials, irrespective of the button selected, an added visual stimulus was displayed indicating that the trial was successfully completed. The researchers then estimated each participant's agency as a function of the number of trials paired with the stimulus and found that an increased sense of agency correlated with decreased reaction times. Manohar et al. (2017) reported that the presence of reward increases peak velocity for saccades, with the greatest effects observed when reward was highly contingent on saccade velocity (higher velocity = greater reward) rather than when reward was not contingent on velocity (reward delivered independently of velocity). The reward in our current study had minimal contingency with the reaching movement. Participants only needed to reach to the correct quadrant to receive reward. However, the additional audiovisual stimuli in movements paired with reward may have indirectly influenced the participants' sense of contingency, contributing to the observation that reward decreased reaction time.

A computational model of reaction time and vigor. A single computational framework may account for the observation that reward produced both a reduction in reaction time and an increase in movement vigor. Let us express utility of a reaching movement as reward minus effort, divided by duration of that movement. This utility is the net rate of reward, where metabolic cost serves as a proxy for effort (Shadmehr et al. 2016):

$$J = \frac{\alpha - aT - bd/T^2}{T} \quad (1)$$

In this expression, α represents the reward associated with the outcome of a successful movement. In the above model, α is represented in units of energy, specifically joules. Movement duration is represented as T , and movement distance as d . The remaining variables are constants that reflect the metabolic cost of reaching across a range of movement speeds and distances. Given the objective of maximizing net rate of reward, the optimum movement vigor is defined via duration T^* :

$$T^* = \left(\frac{3bd}{\alpha} \right)^{1/2} \quad (2)$$

The above expression implies that reward decreases the optimum movement duration, resulting in increased vigor. At the optimum duration, the resulting utility of the movement is

$$J^* = \frac{2\alpha^{3/2}}{3\sqrt{3bd}} - a \quad (3)$$

This implies that as reward increases, the utility of that option increases. During reaction time, decision making proceeds by integrating to threshold a random variable. If that random variable has a mean that is proportional to the rate specified by the utility of that action (Eq. 3), then the rate of rise increases as reward increases, producing an earlier reaction time. As a result, a utility that is defined as the rate of net reward, where effort is the metabolic cost of the action, can account for both the effect of reward on vigor and the effect of reward on reaction time.

Increase in vigor does not increase variability. We found that reward reduced variability of reaction time. Work by Takikawa et al. (2002), Manohar et al. (2015), and Manohar et

al. (2017) examined saccades and found that reward led to both an increase in vigor while reducing end-point variability. In reaching, Nikooyan and Ahmed (2015) observed that in an adaptation task, the addition of reward feedback led to greater reductions in reach end-point variability compared with visual feedback alone. In addition, Pekny et al. (2015) found that reward probability altered reach variability, with movements occurring under high probability of reward being less variable than movements under low probability. They found that reward-dependent control of variability was impaired in Parkinson's disease, suggesting a role for the basal ganglia.

A central source of variability may be the neural activity during the delay period when the movement is being planned. Churchland et al. (2007) noted that trial-to-trial variability in the activity of cells in the primary motor cortex and premotor cortex during the delay period accounted for roughly half of the variability in reach velocity. Although the effect of reward on the delay period activity of reach-related neurons is not well understood, pairing of a stimulus with reward tends to increase the delay period activity of neurons that direct a saccade toward that stimulus (Ikeda and Hikosaka 2003), an effect that is similar to changes associated with increased spatial attention (Ignashchenkova et al. 2004). On this basis, it is possible that the reward-related changes in reach variability may be associated with preferential allocation of spatial attention.

Neural correlates in reward-dependent modulation of vigor. Natural variations in dopamine levels can predict the amount of effort an individual will exert for reward (Wardle et al. 2011). In Parkinson's disease, dopamine levels deteriorate, slowing movement (bradykinesia) (Hallett and Khoshbin 1980). This symptom is traditionally believed to be due to increased signal-dependent noise in the motor system (Montgomery and Nissen 1990; Phillips et al. 1994). An alternative or perhaps complementary explanation of bradykinesia is that dopamine is essential in establishing vigor, with the pathology leading to a general decrease in motivation to move (Kojovic et al. 2014; Mazzoni et al. 2007; Salimpour et al. 2015) as well as decreased ability to adjust movements in response to changing reward landscapes (Kojovic et al. 2014; Pekny et al. 2015; Schmidt et al. 2008). Including individuals with parkinsonian symptoms in our current paradigm, with its low consequences on accuracy, may provide a promising platform in helping to further elucidate the role of dopamine in modulating both the vigor and variability of our movements.

Limitations. Our protocol only considered two conditions: reward and no reward. As a result, we did not quantify or modulate the value of reward. Adding auditory and visual reward coincided with an increase in reaching velocity of around 2%. Xu-Wilson et al. (2009) reported a 1% increase in saccade velocity toward images of human faces compared with other images. Nonhuman primates exhibit much greater changes in saccade velocity to obtain juice rewards (~25%) (Kawagoe et al. 1998; Takikawa et al. 2002). This difference may be due to reward modality. In our study, as well as the study by Xu-Wilson et al. (2009), reward had no explicit utility compared with the caloric rewards in the nonhuman primate studies.

Quantifying reward on the basis of its metabolic/energetic content predicts when starlings choose to walk and fly (Bautista et al. 2001). Studies on humans have used monetary rewards to study movement decisions under uncertainty

(O'Brien and Ahmed 2013, 2015, 2016); however, these rewards exhibit significant distortions from their actual value, which vary across individuals (Kahneman and Tversky 1979). Other intrinsic reward mediums are more difficult to quantify, such as the value of different images (Xu-Wilson et al. 2009). Furthermore, little is understood about how these intrinsic rewards compare with other extrinsic rewards such as food or money.

One potential method of developing a universal currency for reward may be through understanding how different rewards affect neural activity, specifically between regions of the prefrontal cortex and dopaminergic striatum (Levy and Glimcher 2011, 2012). This foundation has been considered in a model of motor control that predicts movement responses (lever presses) based on levels of dopamine (Niv et al. 2007). The model advances the role of environment by considering reward's influence on dopamine activity on both a phasic (quality of individual rewards) and tonic level (rate of reward). Understanding how the dopaminergic midbrain responds to reward may prove essential in explaining movement preference both across and within populations.

Our experiment did not control intertrial intervals. The only temporal constraint between trials was a short 150-ms period of time when the cursor was held in the start circle. Other than this delay, the pace of the experiment was limited only by how quickly participants completed their trials. Work focusing on intertrial intervals suggests that it is not just reward quality but also reward rate that alters movements (Haith et al. 2012; Niv et al. 2007). Not controlling reward rate, in principle, may explain the observed increase in vigor as the experiment progressed.

Conclusion. Humans reacted with shorter latency and produced faster and longer reaching movements when anticipating reward. In addition to modulating vigor, reward also led to more consistent movements, reducing the variance of the reaction times, compared with similar, nonrewarded movements. These results support the idea that vigor is not optimized solely by minimizing effort costs or error, but instead depends on a utility where reward discounts effort.

GRANTS

This work was supported by National Institutes of Health Grants 1R01-NS096083 and 1R01NS078311, National Science Foundation Grant 1723967 and CAREER Award 1352632, and Office of Naval Research Grant N00014-15-1-2312.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

E.S. and A.A.A. conceived and designed research; E.S. performed experiments; E.S., R.S., and A.A.A. analyzed data; E.S., R.S., and A.A.A. interpreted results of experiments; E.S., R.S., and A.A.A. prepared figures; E.S., R.S., and A.A.A. drafted manuscript; E.S., R.S., and A.A.A. edited and revised manuscript; E.S., R.S., and A.A.A. approved final version of manuscript.

REFERENCES

Alexander RM. A minimum energy cost hypothesis for human arm trajectories. *Biol Cybern* 76: 97–105, 1997. doi:10.1007/s004220050324.

- Bautista LM, Tinbergen J, Kacelnik A. To walk or to fly? How birds choose among foraging modes. *Proc Natl Acad Sci USA* 98: 1089–1094, 2001. doi:10.1073/pnas.98.3.1089.
- Behne N, Scheich H, Brechmann A. The left dorsal striatum is involved in the processing of neutral feedback. *Neuroreport* 19: 1497–1500, 2008. doi:10.1097/WNR.0b013e32830fe98c.
- Bendixby MS, Platt ML. Neural correlates of reward and attention in macaque area LIP. *Neuropsychologia* 44: 2411–2420, 2006. doi:10.1016/j.neuropsychologia.2006.04.011.
- Berret B, Jean F. Why don't we move slower? The value of time in the neural control of action. *J Neurosci* 36: 1056–1070, 2016. doi:10.1523/JNEUROSCI.1921-15.2016.
- Burdet E, Osu R, Franklin DW, Milner TE, Kawato M. The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature* 414: 446–449, 2001. doi:10.1038/35106566.
- Chen-Harris H, Joiner WM, Ethier V, Zee DS, Shadmehr R. Adaptive control of saccades via internal feedback. *J Neurosci* 28: 2804–2813, 2008. doi:10.1523/JNEUROSCI.5300-07.2008.
- Choi JE, Vaswani PA, Shadmehr R. Vigor of movements and the cost of time in decision making. *J Neurosci* 34: 1212–1223, 2014. doi:10.1523/JNEUROSCI.2798-13.2014.
- Churchland MM, Yu BM, Sahani M, Shenoy KV. Techniques for extracting single-trial activity patterns from large-scale neural recordings. *Curr Opin Neurobiol* 17: 609–618, 2007. doi:10.1016/j.conb.2007.11.001.
- Elsner B, Hommel B. Contiguity and contingency in action-effect learning. *Psychol Res* 68: 138–154, 2004. doi:10.1007/s00426-003-0151-8.
- Esteves PO, Oliveira LA, Nogueira-Campos AA, Saunier G, Pozzo T, Oliveira JM, Rodrigues EC, Volchan E, Vargas CD. Motor planning of goal-directed action is tuned by the emotional valence of the stimulus: a kinematic study. *Sci Rep* 6: 28780, 2016. doi:10.1038/srep28780.
- Haith AM, Reppert TR, Shadmehr R. Evidence for hyperbolic temporal discounting of reward in control of movements. *J Neurosci* 32: 11727–11736, 2012. doi:10.1523/JNEUROSCI.0424-12.2012.
- Hallett M, Khoshbin S. A physiological mechanism of bradykinesia. *Brain* 103: 301–314, 1980. doi:10.1093/brain/103.2.301.
- Harris CM, Wolpert DM. Signal-dependent noise determines motor planning. *Nature* 394: 780–784, 1998. doi:10.1038/29528.
- Ignashchenkova A, Dicke PW, Haarmeier T, Thier P. Neuron-specific contribution of the superior colliculus to overt and covert shifts of attention. *Nat Neurosci* 7: 56–64, 2004. doi:10.1038/nn1169.
- Ikeda T, Hikosaka O. Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron* 39: 693–700, 2003. doi:10.1016/S0896-6273(03)00464-1.
- Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica* 47: 263–291, 1979. doi:10.2307/1914185.
- Karsh N, Eitam B. I control therefore I do: judgments of agency influence action selection. *Cognition* 138: 122–131, 2015. doi:10.1016/j.cognition.2015.02.002.
- Kawagoe R, Takikawa Y, Hikosaka O. Expectation of reward modulates cognitive signals in the basal ganglia. *Nat Neurosci* 1: 411–416, 1998. doi:10.1038/1625.
- Kojovic M, Mir P, Trender-Gerhard I, Schneider SA, Pareés I, Edwards MJ, Bhatia KP, Jahanshahi M. Motivational modulation of bradykinesia in Parkinson's disease off and on dopaminergic medication. *J Neurol* 261: 1080–1089, 2014. doi:10.1007/s00415-014-7315-x.
- Levy DJ, Glimcher PW. Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. *J Neurosci* 31: 14693–14707, 2011. doi:10.1523/JNEUROSCI.2218-11.2011.
- Levy DJ, Glimcher PW. The root of all value: a neural common currency for choice. *Curr Opin Neurobiol* 22: 1027–1038, 2012. doi:10.1016/j.conb.2012.06.001.
- Manohar SG, Chong TT, Apps MA, Batla A, Stamelou M, Jarman PR, Bhatia KP, Husain M. Reward pays the cost of noise reduction in motor and cognitive control. *Curr Biol* 25: 1707–1716, 2015. doi:10.1016/j.cub.2015.05.038.
- Manohar SG, Finzi RD, Drew D, Husain M. Distinct motivational effects of contingent and noncontingent rewards. *Psychol Sci* 28: 1016–1026, 2017. doi:10.1177/0956797617693326.
- Mazzoni P, Hristova A, Krakauer JW. Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. *J Neurosci* 27: 7105–7116, 2007. doi:10.1523/JNEUROSCI.0264-07.2007.
- Milstein DM, Dorris MC. The influence of expected value on saccadic preparation. *J Neurosci* 27: 4810–4818, 2007. doi:10.1523/JNEUROSCI.0577-07.2007.

- Montgomery EB Jr, Nuessen J.** The movement speed/accuracy operator in Parkinson's disease. *Neurology* 40: 269–272, 1990. doi:10.1212/WNL.40.2.269.
- Mosberger AC, de Clauser L, Kasper H, Schwab ME.** Motivational state, reward value, and Pavlovian cues differentially affect skilled forelimb grasping in rats. *Learn Mem* 23: 289–302, 2016. doi:10.1101/lm.039537.115.
- Nikooyan AA, Ahmed AA.** Reward feedback accelerates motor learning. *J Neurophysiol* 113: 633–646, 2015. doi:10.1152/jn.00032.2014.
- Niv Y, Daw ND, Joel D, Dayan P.** Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 191: 507–520, 2007. doi:10.1007/s00213-006-0502-4.
- O'Brien MK, Ahmed AA.** Does risk-sensitivity transfer across movements? *J Neurophysiol* 109: 1866–1875, 2013. doi:10.1152/jn.00826.2012.
- O'Brien MK, Ahmed AA.** Threat affects risk preferences in movement decision making. *Front Behav Neurosci* 9: 150, 2015. doi:10.3389/fnbeh.2015.00150.
- O'Brien MK, Ahmed AA.** Rationality in human movement. *Exerc Sport Sci Rev* 44: 20–28, 2016. doi:10.1249/JES.0000000000000066.
- Opris I, Lebedev M, Nelson RJ.** Motor planning under unpredictable reward: modulations of movement vigor and primate striatum activity. *Front Neurosci* 5: 61, 2011. doi:10.3389/fnins.2011.00061.
- Pasquereau B, Nadjar A, Arkadir D, Bezdard E, Goillandeau M, Bioulac B, Gross CE, Boraud T.** Shaping of motor responses by incentive values through the basal ganglia. *J Neurosci* 27: 1176–1183, 2007. doi:10.1523/JNEUROSCI.3745-06.2007.
- Pekny SE, Izawa J, Shadmehr R.** Reward-dependent modulation of movement variability. *J Neurosci* 35: 4015–4024, 2015. doi:10.1523/JNEUROSCI.3244-14.2015.
- Phillips JG, Martin KE, Bradshaw JL, Iansek R.** Could bradykinesia in Parkinson's disease simply be compensation? *J Neurol* 241: 439–447, 1994. doi:10.1007/BF00900963.
- Ratcliff R, Rouder JN.** Modeling response times for two-choice decisions. *Psychol Sci* 9: 347–356, 1998. doi:10.1111/1467-9280.00067.
- Reppert TR, Lempert KM, Glimcher PW, Shadmehr R.** Modulation of saccade vigor during value-based decision making. *J Neurosci* 35: 15369–15378, 2015. doi:10.1523/JNEUROSCI.2621-15.2015.
- Ridderinkhof KR, Scheres A, Oosterlaan J, Sergeant JA.** Delta plots in the study of individual differences: new tools reveal response inhibition deficits in AD/HD that are eliminated by methylphenidate treatment. *J Abnorm Psychol* 114: 197–215, 2005. doi:10.1037/0021-843X.114.2.197.
- Rigoux L, Guigon E.** A model of reward- and effort-based optimal decision making and motor control. *PLoS Comput Biol* 8: e1002716, 2012. doi:10.1371/journal.pcbi.1002716.
- Salimpour Y, Mari ZK, Shadmehr R.** Altering effort costs in Parkinson's disease with noninvasive cortical stimulation. *J Neurosci* 35: 12287–12302, 2015. doi:10.1523/JNEUROSCI.1827-15.2015.
- Schmidt L, d'Arc BF, Lafargue G, Galanaud D, Czernecki V, Grabli D, Schüpbach M, Hartmann A, Lévy R, Dubois B, Pessiglione M.** Disconnecting force from money: effects of basal ganglia damage on incentive motivation. *Brain* 131: 1303–1310, 2008. doi:10.1093/brain/awn045.
- Schweighofer N, Xiao Y, Kim S, Yoshioka T, Gordon J, Osu R.** Effort, success, and nonuse determine arm choice. *J Neurophysiol* 114: 551–559, 2015. doi:10.1152/jn.00593.2014.
- Shadmehr R, Huang HJ, Ahmed AA.** A representation of effort in decision-making and motor control. *Curr Biol* 26: 1929–1934, 2016. doi:10.1016/j.cub.2016.05.065.
- Shadmehr R, Orban de Xivry JJ, Xu-Wilson M, Shih TY.** Temporal discounting of reward and the cost of time in motor control. *J Neurosci* 30: 10507–10516, 2010. doi:10.1523/JNEUROSCI.1343-10.2010.
- Smith MA, Shadmehr R.** Intact ability to learn internal models of arm dynamics in Huntington's disease but not cerebellar degeneration. *J Neurophysiol* 93: 2809–2821, 2005. doi:10.1152/jn.00943.2004.
- Takikawa Y, Kawagoe R, Itoh H, Nakahara H, Hikosaka O.** Modulation of saccadic eye movements by predicted reward outcome. *Exp Brain Res* 142: 284–291, 2002. doi:10.1007/s00221-001-0928-1.
- Trommershäuser J, Maloney LT, Landy MS.** Statistical decision theory and the selection of rapid, goal-directed movements. *J Opt Soc Am A Opt Image Sci Vis* 20: 1419–1433, 2003. doi:10.1364/JOSAA.20.001419.
- van Beers RJ, Haggard P, Wolpert DM.** The role of execution noise in movement variability. *J Neurophysiol* 91: 1050–1063, 2004. doi:10.1152/jn.00652.2003.
- Wang C, Xiao Y, Burdet E, Gordon J, Schweighofer N.** The duration of reaching movement is longer than predicted by minimum variance. *J Neurophysiol* 116: 2342–2345, 2016. doi:10.1152/jn.00148.2016.
- Wardle MC, Treadway MT, Mayo LM, Zald DH, de Wit H.** Amping up effort: effects of d-amphetamine on human effort-based decision-making. *J Neurosci* 31: 16597–16602, 2011. doi:10.1523/JNEUROSCI.4387-11.2011.
- Watanabe K, Hikosaka O.** Immediate changes in anticipatory activity of caudate neurons associated with reversal of position-reward contingency. *J Neurophysiol* 94: 1879–1887, 2005. doi:10.1152/jn.00012.2005.
- Xu-Wilson M, Zee DS, Shadmehr R.** The intrinsic value of visual information affects saccade velocities. *Exp Brain Res* 196: 475–481, 2009. doi:10.1007/s00221-009-1879-1.