

Reward Prediction Error Modulates Saccade Vigor

 Ehsan Sedaghat-Nejad,  David J. Herzfeld, and  Reza Shadmehr

Laboratory for Computational Motor Control, Department of Biomedical Engineering, Johns Hopkins School of Medicine, Baltimore, Maryland 21205

Movement vigor, defined as the reciprocal of the latency from availability of reward to its acquisition, changes with reward magnitude: movements exhibit shorter reaction time and increased velocity when they are directed toward more rewarding stimuli. This invigoration may be due to release of dopamine before movement onset, which has been shown to be modulated by events that signal reward prediction error (RPE). Here, we generated an RPE event in the milliseconds before movement onset and tested whether there was a relationship between RPE and vigor. Human subjects (both sexes) made saccades toward an image. During execution of the primary saccade, we probabilistically changed the position and content of that image, encouraging a secondary saccade. On some trials, the content of the secondary image was more valuable than the first image, resulting in a positive RPE (+RPE) event that preceded the secondary saccade. On other trials, this content was less valuable (−RPE event). We found that reaction time of the secondary saccade was affected in an orderly fashion by the magnitude and direction of the preceding RPE event: the most vigorous saccades followed the largest +RPE, whereas the least vigorous saccades followed the largest −RPE. Presence of the secondary saccade indicated that the primary saccade had experienced a movement error, inducing trial-to-trial adaptation. However, this learning from movement error was not modulated by the RPE event. The data suggest that RPE events, which are thought to transiently alter the release of dopamine, modulate the vigor of the ensuing movement.

Key words: dopamine; vigor; saccade; latency; reward prediction error; motor control

Significance Statement

Does dopamine release in response to a stimulus serve to invigorate the ensuing movement? To test this hypothesis, we relied on the fact that reward prediction error (RPE) is a strong modulator of dopamine. Our innovation was a task in which an RPE event occurred precisely before onset of a stimulus-driven movement. We probabilistically produced a combination of large or small, negative or positive RPE events and observed that saccade vigor carried a robust signature of the preceding RPE event: high vigor saccades followed +RPE events, whereas low vigor saccades followed −RPE events. This suggests that in humans, vigor is partly controlled through release of dopamine in the moments before onset of a movement.

Introduction

We tend to move with shorter latency and greater velocity toward stimuli that we associate with greater value. For example, when the expected reward is large, saccades (Kawagoe et al., 1998; Milstein and Dorris, 2007; Xu-Wilson et al., 2009) and reaching movements (Summerside et al., 2018) toward the reward site have shorter reaction times and higher peak velocities than when the expected reward is smaller. That is, the reciprocal of the time

it takes to arrive at the reward site, which we can operationally define as vigor (Shadmehr et al., 2019), is modulated with reward magnitude. This link between expected reward and movement vigor may be partly due to the function of the basal ganglia (Kawagoe et al., 2004; Tachibana and Hikosaka, 2012) and release of dopamine (da Silva et al., 2018), raising the possibility that before every movement, the dopamine that is released in response to the stimulus partly controls the vigor of the ensuing movement.

Dopamine release appears to follow a simple rule. When the acquired reward is unexpectedly large, the neurons fire a burst, but if the same reward is expected, then the neurons no longer respond (Schultz et al., 1997; Bayer and Glimcher, 2005). Dopamine neurons encode the difference between the predicted stimulus value and the actually acquired value, termed reward prediction error (RPE). This transient encoding of RPE provides an interesting prediction: if dopamine release in the milliseconds before movement onset contributes to control of vigor, then movements that follow a positive RPE (+RPE) event should ex-

Received Feb. 11, 2019; revised April 16, 2019; accepted April 18, 2019.

Author contributions: E.S.-N., D.J.H., and R.S. designed research; E.S.-N. performed research; E.S.-N. analyzed data; E.S.-N. and D.J.H. edited the paper; R.S. wrote the paper.

The work was supported by the National Institutes of Health (Grant 5-R01-NS078311), the Office of Naval Research (Grant N00014-15-1-2312), and the National Science Foundation (Grant CNS-1714623).

The authors declare no competing financial interests.

D.J. Herzfeld's present address: Department of Neurobiology, Duke University School of Medicine, Durham, NC. Correspondence should be addressed to Ehsan Sedaghat-Nejad at e.sedaghatnejad@gmail.com or Reza Shadmehr at shadmehr@jhu.edu.

<https://doi.org/10.1523/JNEUROSCI.0432-19.2019>

Copyright © 2019 the authors

hibit high vigor and those that follow a negative RPE ($-RPE$) event should exhibit low vigor. Vigor modulation should depend on RPE, not reward itself.

Unfortunately, the hypothesis that RPE (and not reward per se) drives vigor has been difficult to test because in a typical experiment, the RPE event occurs after a movement has been completed and the reward acquired, not before the onset of a movement. Here, we designed a task that overcame this limitation.

In our experiment, we relied on the idea that viewing of images carries some of the hallmarks of reward: when given the option of choosing from various image categories, people prefer face images and are willing to spend a greater amount of effort in exchange for gazing at those images (Aharon et al., 2001; Yoon et al., 2018). Furthermore, viewing of face images activates the brain's reward system (O'Doherty et al., 2003). We used images as a proxy for reward and then probabilistically controlled the image content to induce RPE events. We investigated whether induction of an RPE event before a movement influenced vigor of that movement.

Subjects made saccades to view an image. However, upon initiation of the saccade, we probabilistically altered the position and content of the image. The position change encouraged the subjects to follow their initial saccade with a secondary saccade. Our concern was vigor of this secondary saccade; that is, its latency and velocity, which affected the total time it took from completion of the primary saccade to conclusion of the secondary saccade, thereby arriving at the reward site.

On some trials, the value of image A (primary image) was higher than that of image B (secondary image), whereas in other trials, the value of B was higher than that of A. As a result, in some trials, subjects expected to view a low-valued image, but upon completion of their primary saccade, were presented with the opportunity to gaze at a high-valued image. This resulted in conditions in which, during the milliseconds before the onset of the secondary saccade (as A was replaced by B), there was a $+RPE$ ($B > A$) or a $-RPE$ ($B < A$) event. We investigated whether the sign and magnitude of the RPE event altered vigor of the secondary saccade.

Materials and Methods

Subjects. A total of $N = 55$ healthy subjects (18–41 years of age, mean \pm SD = 23 ± 7 ; 34 females) participated in this study. The procedures were approved by the Johns Hopkins School of Medicine Institutional Review Board. All subjects signed a written consent form.

Data collection procedure. Subjects viewed an LED monitor (27-inch, 2560×1440 pixels, light gray background, refresh rate 144 Hz) placed at a distance of 35 cm while we measured their eye position at 1000 Hz (Eyelink 1000). Each trial began with presentation of a fixation spot (a green dot, $0.3^\circ \times 0.3^\circ$) that was randomly drawn near the center of the screen (the fixation spot was placed in a virtual box at -3° to $+1^\circ$ along the horizontal axis and -1.5 to $+1.5^\circ$ along the vertical axis, where 0,0 refers to center of the screen). After a random fixation interval of 250–750 ms (uniform distribution), the fixation spot was erased and a primary image was placed at 9° to the right along the horizontal axis. The size of this image was constant for each subject, but varied between subjects: $1.5^\circ \times 1.5^\circ$ for some ($n = 20$, 100 trials per block, 13 blocks of trials), $3.0^\circ \times 3.0^\circ$ for others ($n = 35$, 145 trials per block, 13 blocks of trials). A green fixation dot always appeared at the center of every image. Because the effect of RPE relied on within-subject analysis, data were combined in these two groups.

The removal of the central fixation dot and presentation of the primary image served as the go signal for the primary saccade. This saccade was detected in real time via a speed threshold ($20^\circ/s$) or an eye position change of 2° from fixation, whichever happened first. Each session contained 13 blocks of trials. In block 1, the primary image remained un-

changed after saccade onset. In blocks 2–13, after detecting saccade onset, the primary image was erased with probability of 50% and a new image was displayed at a distance of 3° from the original image. A green fixation dot also appeared at the center of the secondary image. As a result, after the completion of the primary saccade, subjects produced a secondary saccade to the secondary image. The location of the secondary image was random on each trial. The reason for this was to preclude accumulation of adaptation on the primary saccades that results from movement errors that subjects experience on each trial (Ethier et al., 2008; Pekny et al., 2011). For some subjects ($n = 33$), the secondary image was randomly located at either $+3^\circ$ or -3° along the horizontal axis with respect to the primary image. For other subjects ($n = 22$), the secondary image was randomly located at each trial at either $+3^\circ$ or -3° along the vertical axis with respect to the primary image. As a result, the location of the secondary image was random along the horizontal or vertical axis. The size of the secondary image was always the same as the primary image.

After completion of the secondary saccade, subjects were provided with 250 ms to view that image. At the end of this period the image was erased and a center fixation dot appeared at a random location near the center of the screen in the bounding box defined above. Each session contained 13 blocks of trials, with 100–145 trials per block. Subjects were provided with a 30 s rest period between each block.

Images were chosen from two categories: face and noise images. The facial images were gathered from the Internet (500 total images) and were modified in a way that the center of the two eyes was located at the center of the image. The noise images were constructed by shuffling the pixels of each face image (500×500 pixels). This ensured that the luminance and color content of the two categories were identical.

Magnitude of the RPE event. In our experiment, we presented a primary image (e.g., a face) and then at random trials replaced it with another image (noise). We hypothesized that viewing each image was a rewarding event and, as a result, a difference between the primary and secondary images would produce an RPE before the execution of the secondary saccade. We estimated the magnitude of the RPE from probability of each image and its relative value.

An objective estimate of the value of a face image with respect to a noise image can be attained from the choices that people make when given the option of viewing these images. For the image types that we used here, people on average chose the face image twice as often as the noise image (Yoon et al., 2018). This suggests that the relative value of face to noise is ~ 2 .

On a given trial, the primary image was changed with a probability of 50%. Assuming a prior probability that target of the saccade is unlikely to change, and the observed likelihood that on 50% of trials the primary image changes, we can write the predicted value of the primary image as follows:

$$\hat{F} = \alpha F + (1 - \alpha)N \quad 0.5 < \alpha < 1 \quad (1)$$

$$\hat{N} = \alpha N + (1 - \alpha)F$$

In the above expression, F and N represent the subjective value of face and noise images, and \hat{F} and \hat{N} are the predicted value. Because the primary image can change, the first equation in the above expression implies that the predicted value for a primary face image is less than its subjective value (the face can become noise). The second equation implies that the predicted value for a primary noise image is greater than its subjective value (noise can become face).

Once the primary saccade concludes, the subject is presented with a secondary image. This is the image that they will actually have the opportunity to gaze at. We define RPE as the value of the second image (reward we will receive) minus its predicted value (reward we had predicted). For example, if A is the primary image and B is the secondary image, then \hat{A} is the reward predicted, but B is the reward that will be received. That is, $RPE[A,B] = B - \hat{A}$.

There are four possible pairs of primary and secondary images. For each pair, we can compute the magnitude of the RPE event as follows:

$$RPE[F,F] = F - \hat{F} = (1 - \alpha)(F - N) \quad (2)$$

$$RPE [F,N] = N - \hat{F} = \alpha(N - F)$$

$$RPE [N,N] = N - \hat{N} = (1 - \alpha)(N - F)$$

$$RPE [N,F] = F - \hat{N} = \alpha(F - N)$$

If we assume that the subjective value of a face image is approximately twice that of noise, $F \approx 2N$, then we have the following:

$$RPE [N,F] = \frac{\alpha}{2}F \quad (3)$$

$$RPE [F,F] = \frac{1 - \alpha}{2}F$$

$$RPE [N,N] = \frac{\alpha - 1}{2}F$$

$$RPE [F,N] = -\frac{\alpha}{2}F$$

The above expressions imply that a noise–face (NF) trial is a very positive RPE event (first equation in the above expression), a face–face (FF) trial is a mildly positive RPE event (second equation), a noise–noise (NN) trial is a mildly negative RPE event (third equation), and a face–noise (FN) trial is a highly negative RPE event (fourth equation).

If vigor is modulated by RPE, then the secondary saccades should exhibit their highest vigor in NF trials and lowest vigor in FN trials. In comparison, FF trials should show smaller vigor compared with NF trials despite the fact that, in both trials, the secondary saccade is toward a face. Finally, NN trials should show a greater vigor than FN trials despite the fact that, in both trials, the secondary saccade is toward noise.

Data analysis. Eye position data were filtered with a second-order Butterworth low-pass filter with cutoff frequency of 100 Hz. Eye velocity data in offline analysis were calculated as the first derivative of the filtered position data. Saccades were identified with a speed magnitude threshold of $20^\circ/s$ and minimum hold time of 10 ms at saccade end (i.e., velocity magnitude could not exceed the cutoff for a minimum 10 ms after the end point). We measured reaction time of the secondary saccade via the time period between offset of the primary saccade and onset of the secondary saccade. Secondary saccades onset and offset were detected identically to the primary saccades using $20^\circ/s$ threshold on velocity magnitude. The saccade duration was considered as the time between saccade onset and offset.

Statistical analyses were performed using SPSS and general linear models, with stimulus value (e.g., face or noise) serving as the within-subject factor. We reported results of tests of within-subject effects under the assumption of sphericity. We tested this assumption via Mauchly's test, which was confirmed in every case reported. We also performed two-sided *t* tests on the between-subject effect of learning from movement errors.

Results

To produce a +RPE event, the trial began with a noise image (Fig. 1A, first column). As subjects initiated their primary saccade, we probabilistically erased that image and replaced it with a face image at a new location (NF trials; Fig. 1A, first column). The control condition for the +RPE event was a trial in which both the primary and secondary images were faces (FF control; Fig. 1A). Similarly, to produce a –RPE event, the trial began with presentation of a face image, which, after saccade onset, was probabilistically replaced with a noise image (FN trials; Fig. 1A, third column). The control condition for the –RPE event was a trial in which both the primary and secondary images were noise (NN control; Fig. 1A). Therefore, the secondary saccade in both the control and RPE trials was made toward the same image. Based on probability of the events, the trials produced four magnitudes of RPE (see Materials and Methods): highly positive (NF

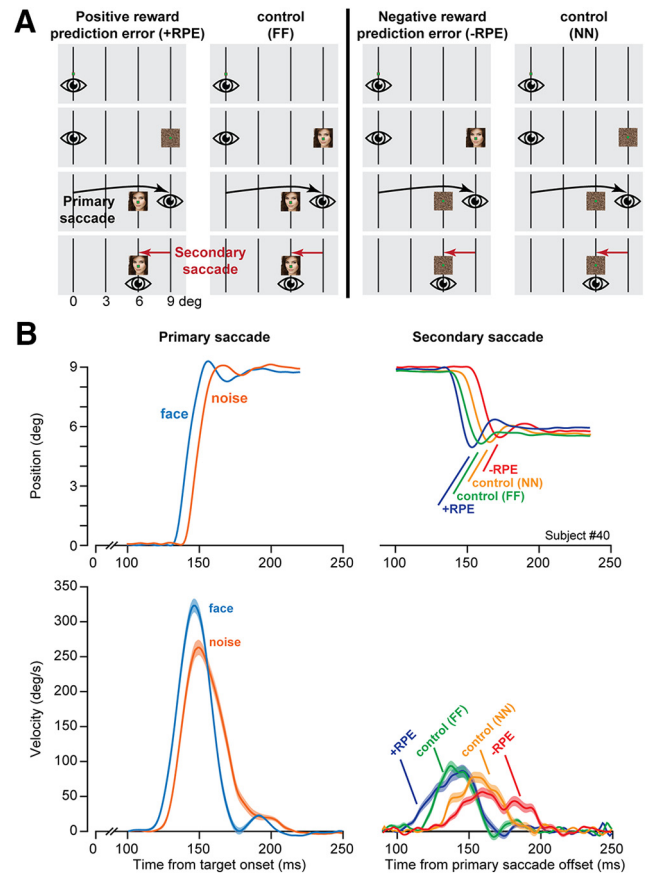


Figure 1. Experiment design and data from a representative subject. **A**, Each trial began with a fixation dot near the center. After a random fixation interval, we presented a primary image at 9° to the right along the horizontal axis. As the primary saccade took place, we erased the primary image and replaced it with a secondary image. In +RPE trials, a noise primary image was replaced with a face secondary image. A face–face trial served as control for the +RPE trial. In –RPE trials, a face primary image was replaced with a noise secondary image. A noise–noise trial served as control for the –RPE trial. **B**, Saccade position and velocity traces for a representative subject. Position traces are for individual saccades. Velocity traces are within-subject average of all saccades. Primary saccade exhibited a shorter reaction time and a higher velocity in response to a face image. Data for the secondary saccade are plotted with respect to termination of the primary saccade in the same trial. The secondary saccades exhibited the shortest reaction times in +RPE trials and the longest reaction times in –RPE trials. Error bars indicate SEM.

trials), slightly positive (FF trials), slightly negative (NN trials), and highly negative (FN trials) (Eq. 3).

Effects of RPE on vigor

Data from a representative subject are shown in Figure 1B. Position traces for single saccades are shown in the top panel of Figure 1B and averaged velocity profiles are shown in the bottom panel. As expected, the primary saccade had a shorter reaction time and higher peak velocity when made toward a face image. During the primary saccade, on some trials, the face image was changed to noise (–RPE event, FN trial). Similarly, on some trials, the noise image was changed to face (+RPE event, NF trial). Because of the change in image location, at 100–150 ms after completion of the primary saccade, the subject generated a secondary saccade. We measured the reaction time of the secondary saccade with respect to end of the primary saccade. The reaction time and peak velocity of the secondary saccade were affected by not just the image at the destination (i.e., the secondary image), but, more importantly, by the sign of the RPE event. Reaction time of the second-

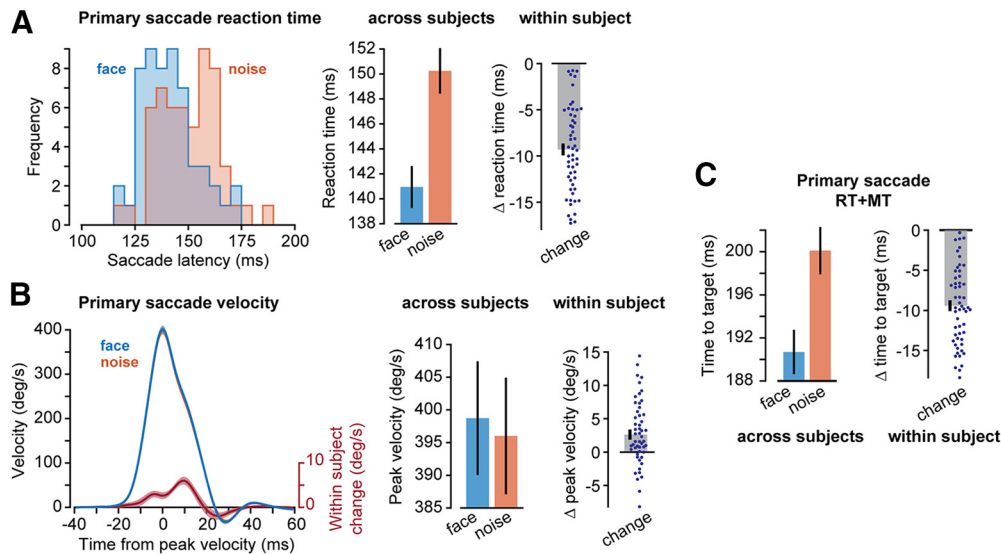


Figure 2. Primary saccades exhibited shorter reaction time and higher velocity in response to face images. **A**, Distribution of mean reaction times and within-subject change in the mean reaction time (face minus noise). Each dot is a single subject. **B**, Saccade velocity traces and within-subject change in velocity (face minus noise). The trace for noise is largely hidden behind the trace for face. **C**, Total time to target (reaction time plus movement duration) and within-subject change. Error bars and shaded traces indicate SEM.

ary saccade appeared shortest after the +RPE event and longest after the -RPE event. Indeed, the properties of the secondary saccade appeared to follow a consistent pattern: shortest reaction time and highest velocity for the most positive RPE event (NF), longest reaction time and lowest velocity for the most negative RPE event (NN), and in between for the mildly positive (FF) and mildly negative (NN) RPE events.

These results were repeated in our population of subjects. The opportunity to view a face image strongly affected the vigor of the primary saccade (Fig. 2). The distribution of reaction times shifted earlier (Fig. 2A), from a mean of 150.2 to 140.95 ms, resulting in a within-subject reduction of 9.30 ± 0.63 ms (mean \pm SEM) (within-subject change, $F_{(1,54)} = 217, p < 10^{-4}$). The face image also induced an increase in the velocity of the primary saccade (Fig. 2B), particularly in the second half of that movement. As a result, saccade peak velocity increased by 2.63 ± 0.76 %/s (within-subject change, $F_{(1,54)} = 12.1, p = 0.001$). Maximum change occurred 10 ms after peak velocity with 6.01 ± 0.84 %/s difference between two categories (within-subject change, $F_{(1,54)} = 51.511, p < 10^{-4}$).

We tried to minimize the changes in saccade amplitude that may arise from changes in image content by presenting a green dot at the center of every image. We observed a very small effect of image type on saccade amplitude: the primary saccades were $8.19 \pm 0.01^\circ$ toward face images and $8.10 \pm 0.01^\circ$ toward noise images, a within-subject change of $0.09 \pm 0.015^\circ$ ($F_{(1,54)} = 32, p < 10^{-4}$). However, this change was less than the resolution of our measurement instrument. Overall, the opportunity to view a face image produced a significant reduction in the total time it took for the eyes to arrive at the location of the primary target (within-subject change, $F_{(1,54)} = 199, p < 10^{-4}$; Fig. 2C).

As the primary saccade started, we displaced the primary image to a new location, encouraging the subjects to produce a secondary saccade. We found that the reaction times for the secondary saccade (Fig. 3B) were shortest in the NF trials (+RPE event) and longest in the FN trials (-RPE event). Indeed, there was an orderly increase in the reaction times in the precise pattern predicted by the RPE events [repeated-measures (RM) ANOVA $F_{(3,52)} = 34.7, p < 10^{-4}$; Fig. 3C]. On average, the secondary

saccade that followed the +RPE event had a reaction time that was 19.15 ± 1.95 ms less than the -RPE event (148.6 ± 3.1 ms in +RPE trial compared with 167.75 ± 3.3 ms in -RPE trial). *Post hoc* pairwise comparisons indicated that reaction time of +RPE trials were 6 ± 1.09 ms smaller than FF control trials ($p < 10^{-4}$) and reaction time of -RPE trials were 6.03 ± 0.95 ms greater than NN control trials ($p < 10^{-4}$). Peak velocity of the secondary saccade appeared affected by the various events (RM ANOVA $F_{(3,52)} = 14.65, p < 10^{-4}$; Fig. 3A,D). However, *post hoc* pairwise comparisons did not dissociate the \pm RPE events from their respective control trials. Overall, the RPE events significantly affected the total time it took for the eyes to respond and acquire the secondary target (Fig. 3E): the time to target, measured from completion of the primary saccade to arrival at the target, was smallest after the +RPE event (NF trials) and largest after the -RPE event (FN trials, RM ANOVA $F_{(3,52)} = 34.88, p < 10^{-4}$). *Post hoc* pairwise comparisons indicated that times to target of +RPE trials were 6.06 ± 1.14 ms smaller than those of FF control trials ($p < 10^{-4}$) and times to target of -RPE trials were 6.15 ± 0.93 ms larger than NN control trials ($p < 10^{-4}$). Therefore, the magnitude of the RPE event that preceded a saccade affected the vigor of that saccade.

Amplitude of the secondary saccade on average varied by $< 0.09^\circ$ across the range of the various conditions: +RPE $3.02 \pm 0.04^\circ$, FF $3.08 \pm 0.04^\circ$, NN $3.04 \pm 0.05^\circ$, and FN $3.0 \pm 0.05^\circ$. Similarly, metrics of the primary saccade were not affected by whether the image was changed midflight. To check for this, we grouped NF and FN trials (image changed) and compared them with FF and NN trials (image unchanged). We found no difference between primary saccades in these trials (RM ANOVA for the effect of stimulus on peak velocity, within-subject change, $F_{(1,54)} = 0.022, p = 0.884$; reaction time within-subject change, $F_{(1,54)} = 0.302, p = 0.585$; time to target within-subject change, $F_{(1,54)} = 0.919, p = 0.342$; and amplitude within-subject change, $F_{(1,54)} = 0.499, p = 0.483$).

Effect of RPE on learning

Presence of a secondary (or corrective) saccade indicates presence of a motor error: at the end of the primary saccade, the target

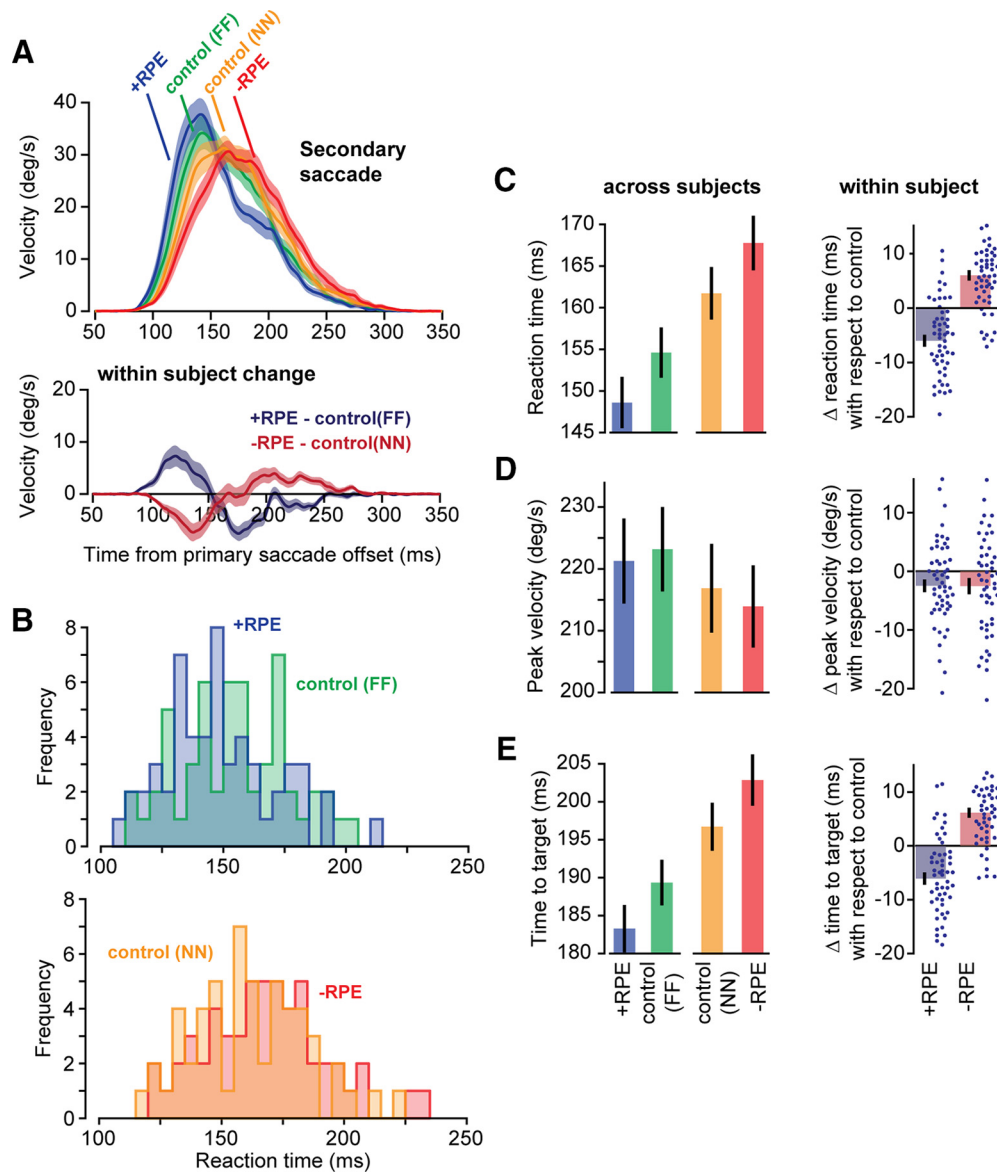


Figure 3. Secondary saccades were influenced by the preceding RPE event. **A**, Saccade velocity and within-subject change in velocity. **B**, Distribution of mean reaction times across subjects. Reaction times are measured as the latency with respect to offset of the preceding primary saccade. **C**, Mean reaction times across subjects and the within-subject change in reaction times. The bars are the +RPE condition with respect to control FF and the -RPE condition with respect to control NN. Dots are individual subjects. **D**, Peak velocity and within-subject change in peak velocity. **E**, Total time to target measured from completion of the primary saccade to conclusion of the secondary saccade; that is, the reaction time plus movement duration of the secondary saccade. The error bars and shaded traces indicate SEM. The dots show within-subject change with respect to control.

was not on the fovea. The resulting movement error should induce plasticity in the cerebellum (Herzfeld et al., 2018), affecting the subsequent primary saccade. We wondered whether presence of the RPE event modulated learning from the motor error. In particular, would a +RPE event enhance learning?

In our experiment, we displaced the primary image along four directions: positive and negative along the horizontal axis (H+ and H-) and positive and negative along the vertical axis (V+ and V-). In all cases, the magnitude of the displacement was 3°. Each displacement resulted in a motor error, which in principle may have induced trial-to-trial learning. To measure this learning, we considered two consecutive trials in which the primary saccades were made to the same visual stimulus type and then further divided the trials based on the direction of motor error. For example, suppose that, on trial n , the primary saccade was toward face and the subject experienced an H+ error on that trial

and that, on trial $n + 1$, the primary saccade was again toward face. In all such consecutive pairs of trials, we measured the change in the primary saccade made in trial $n + 1$ with respect to trial n . This trial-to-trial change in the primary saccade is plotted in Figure 4A for each type of motor error. We found that, across all error types, the largest change in the velocity profile was ~15 ms after the saccade peak velocity (Fig. 4A). After an H+ error, the tail of velocity trace (15 ms after peak velocity) increased by $6.62 \pm 1.02^\circ/\text{s}$ (two-sided t test, $t_{(32)} = 6.485$, $p < 10^{-4}$) along the horizontal direction. The trial-to-trial change in saccade amplitude showed $0.17 \pm 0.018^\circ$ (two-sided t test, $t_{(32)} = 9.098$, $p < 10^{-4}$) increase after an H+ error. Similarly, after an H- error, the subsequent primary saccade exhibited a $4.96 \pm 1.09^\circ/\text{s}$ (two-sided t test, $t_{(32)} = 4.550$, $p < 10^{-4}$) decrease in the tail of velocity trace along the horizontal direction and $0.17 \pm 0.016^\circ$ (two-sided t test, $t_{(32)} = 10.747$, $p < 10^{-4}$) reduction in amplitude.

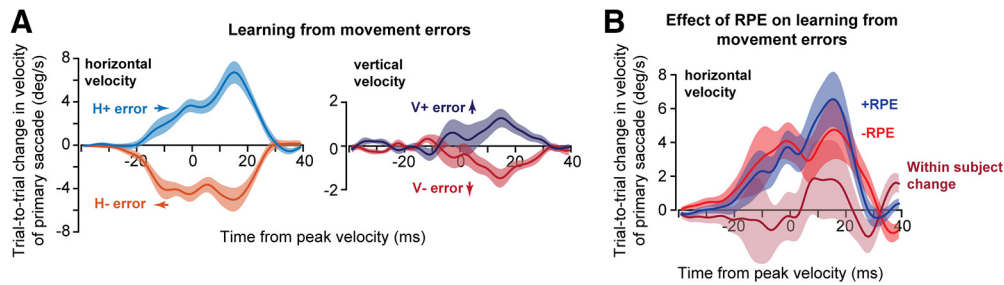


Figure 4. Learning from movement errors and the effect of RPE on learning. **A**, Change in the velocity of the primary saccade from the trial in which the movement error was experienced to the next trial (the primary image type in both saccades were the same). Saccades are grouped based on the movement error experienced at the conclusion of the first primary saccade. Movement error is defined as the position of the secondary image with respect to the primary image. That is, H+ errors imply that the secondary image was further to the right along the horizontal axis than the primary image. The motor error along each axis produced learning along that axis. **B**, Trial-to-trial changes for the horizontal motor-errors-only grouped based on the RPE event that followed the first primary saccade. Data for H– were flipped and averaged with H+ trials. Within-subject change in learning was marginally larger after a +RPE event (compared with a –RPE event), but this effect was not significant.

Learning was also present after V+ and V– errors. After a V+ error, there was a $1.28 \pm 0.42^\circ/\text{s}$ (two-sided *t* test, $t_{(21)} = 3.038$, $p = 0.006$) change in velocity tail and a $0.027 \pm 0.012^\circ$ (two-sided *t* test, $t_{(21)} = 2.348$, $p = 0.029$) change in amplitude of vertical component of primary saccade in the subsequent trial. After a V– error, there was a $1.41 \pm 0.42^\circ/\text{s}$ (two-sided *t* test, $t_{(21)} = 3.368$, $p = 0.003$) change in velocity tail and a $0.033 \pm 0.0086^\circ$ (two-sided *t* test, $t_{(21)} = 3.809$, $p = 0.001$) change in vertical amplitude. These results demonstrated that experience of a motor error on a given trial induced learning, resulting in an error-dependent change in the subsequent primary saccade.

In some trials, a given movement error occurred in the context of a +RPE event, whereas in other trials, the same movement error occurred in the context of a –RPE event. To measure the effect of the RPE events on learning, we focused on the horizontal error trials because the vertical error trials produced substantially less learning. The trial-to-trial changes in the primary saccade in the +RPE and –RPE trials are plotted in Figure 4B. To make this figure, we considered the H+ and H– pairs of trials together (with the positive axis now reflecting trial-to-trial change in velocity along direction of the movement error). We found that learning after a +RPE event ($6.46 \pm 1.59^\circ/\text{s}$ change in velocity tail, $t_{(32)} = 4.067$, $p < 10^{-3}$) was marginally stronger than after a –RPE event ($4.64 \pm 1.73^\circ/\text{s}$, $t_{(32)} = 2.677$, $p = 0.012$), but the effect did not reach statistical significance ($1.83 \pm 2.32^\circ/\text{s}$ within-subject change, $F_{(1,32)} = 0.620$, $p = 0.23$).

Discussion

It is possible that dopamine release in the milliseconds before onset of a movement serves to invigorate the ensuing movement. For example, in mice, self-initiated movements in an empty field that have higher than average acceleration tend to be preceded by higher than average activity in nigral dopamine cells (da Silva et al., 2018). Here, we attempted to indirectly test the link between dopamine release and vigor of self-generated movements in humans. Our approach was to probabilistically change the potential reward for completing a movement, thereby producing a combination of large or small, negative or positive RPE events. Because dopamine release is affected by the magnitude and direction of the RPE event (Schultz et al., 1997; Bayer and Glimcher, 2005), we posited that vigor of the movement that followed the RPE event should exhibit a distinct pattern: highest vigor after a large +RPE events and lowest vigor after a large –RPE event.

Our innovation was a behavioral paradigm in which the RPE events occurred just before the onset of a movement. Subjects were presented with the opportunity to view face or noise images.

Once they initiated their primary saccade, we probabilistically changed the location and content of the image. As a result, after completion of the primary saccade, the image value was higher or lower than expected. This resulted in a range of RPEs: highly negative (face change to noise, FN), slightly negative (noise not changed, NN), slightly positive (face not changed, FF), and highly positive (noise changed to face, NF). This RPE event took place just before the onset of the secondary saccade, which had a latency of ~ 150 ms with respect to termination of the primary saccade. We found that reaction times were shortest after the large +RPE event and longest after the large –RPE event. The time to target, defined as sum of reaction time and movement duration, was shortest for the largest +RPE event and longest for the largest –RPE event. That is, the magnitude and direction of the RPE event modulated vigor of the ensuing saccade.

Although there are currently no commonly accepted definitions of movement vigor, in the context of elementary, stimulus-driven movements such as saccades and reaching, one useful definition is the inverse of the time from stimulus onset to movement completion, conditioned on distance (Shadmehr et al., 2019). This definition is based on the empirical observation that both reaction time and movement duration are influenced by the subjective value of the reward at the destination (Kawagoe et al., 1998; Milstein and Dorris, 2007; Xu-Wilson et al., 2009; Haith et al., 2012; Manohar et al., 2015; Reppert et al., 2015; Summerside et al., 2018).

Because saccade reaction time and velocity depend on activity of neurons in the superior colliculus (Smalianchuk et al., 2018) and these neurons are influenced by luminance and other low-level properties of the visual stimulus (Marino et al., 2015), the differences in vigor may have arisen not from presence of an RPE event, but rather because of other variables associated with differences in properties of the secondary image. Therefore, we included control trials in which the saccade of interest was made to the same image as in RPE trials, but without the benefit of an RPE event. We found that in +RPE trials, saccades had shorter reaction times compared with image-matched control trials. Similarly, in –RPE trials, saccades had longer reaction times compared with image-matched control trials. Furthermore, because the primary and secondary images were always presented at different locations with respect to the fovea, and often in opposite directions, we would expect little or no overlap between regions of collicular activity associated with the primary and secondary saccades. This dissociation between magnitude of primary and secondary saccades reduces the possibility of an interaction be-

tween the saccades at the level of colliculus. As a result, it seems likely that the vigor differences in the secondary saccades were not solely due to differences in the intrinsic response of the collicular neurons to information that they received from the retina. Rather, the link between RPE events in our experiments and modulation of vigor may have been because of reward-dependent regions that project to the colliculus, such as the basal ganglia and the frontal eye field (FEF).

Saccades toward a rewarding stimulus exhibit greater vigor partly because the opportunity for reward reduces the inhibition that the colliculus receives from the substantia nigra reticulata (Yasuda et al., 2012). In addition, the opportunity for reward also increases the excitation that the colliculus receives from the FEF (Heitz and Schall, 2012; Glaser et al., 2016). It seems likely that dopamine plays an important role in controlling this drive. Just before the onset of a spontaneous movement, there is a diversity of responses among dopamine neurons: some show a transient increase, whereas others show a transient decrease (da Silva et al., 2018). For the dopamine cells that increase their activity, the amount of increase is positively corrected with the acceleration of the upcoming movement (da Silva et al., 2018).

In monkeys, dopamine neurons respond to presentation of a saccade target within 100 ms and dissociate between reward and nonrewarding stimuli within 150 ms (Matsumoto and Hikosaka, 2007). For unconditioned, aversive stimuli, dopamine response can dissociate between various magnitudes in <100 ms (Matsumoto and Hikosaka, 2009). In our data, saccades that were affected by the RPE event were generated at extremely short reaction times of ~150 ms. Therefore, in principle, the time range of dopamine response to reward is near the window of vigor modulation that we observed in saccades. Because RPE events have a robust effect on dopamine release, it is possible that the RPE-event-driven change in saccade vigor is linked through dopamine-dependent drive to the basal ganglia and FEF, affecting saccade-related discharge in the superior colliculus.

Reward-dependent modulation of learning from movement error

If the primary saccade ends but the target is not on the fovea, the result is a sensory prediction error that produces unexpected activity on the superior colliculus (Kojima and Soetedjo, 2017a, 2018). This activity engages the inferior olive, resulting in modulation of complex spikes in the Purkinje cells of the cerebellum (Soetedjo et al., 2008; Kojima et al., 2010; Herzfeld et al., 2015). The result is plasticity in the Purkinje cells (Herzfeld et al., 2018), altering the primary saccade on the subsequent trial. In our experiment, we found robust evidence for this trial-to-trial learning: after a secondary saccade, the velocity of the next primary saccade was adjusted in the direction of the motor error.

Previous work has suggested that reward magnitude can modulate the rate of saccade adaptation. Kojima and Soetedjo (2017b) observed that, when monkeys were rewarded for making saccades for one direction but not for another, they adapted their primary saccades more strongly in the direction of the rewarded trials. This reward-dependent effect developed slowly, accumulating over the course of ~400 trials.

Here, we did not see a significant effect of RPE on learning from movement errors. This may have been because we focused on adaptation in response to random errors rather than consistent errors. Randomness of errors downregulates learning from error (Herzfeld et al., 2014), which makes it more difficult to measure any modulatory influence that reward may have had on learning from error.

Limitations

Although we observed robust effects of RPE on saccade latencies, we did not observe an effect on saccade velocities. One reason for this may have been that, in our experiment, the saccades of interest (the secondary saccades) were only 3° in amplitude. Although larger saccades produce higher velocities, they require presentation of stimuli farther from the fovea, which may make identification of that image more difficult. Regardless, the question of whether RPE events alter velocity of the ensuing movement can benefit from further exploration.

Our interpretations regarding RPE events relied on the assumption that the opportunity to gaze at an image served as proxy for reward acquisition. This assumption is based on the observation that, for humans, gazing at images follows many of the behavioral characteristics associated with acquisition of primary rewards (i.e., food). For example, people make saccades that are faster toward images that they prefer (Xu-Wilson et al., 2009), they gaze for a longer period of time at those images (Yoon et al., 2018), and they are willing to pay a greater effort cost to have the opportunity to view their preferred images (Yoon et al., 2018). Furthermore, viewing of images activates the reward system of the brain (O'Doherty et al., 2003). However, despite these observations, the question of whether gazing at an image engages the dopamine system remains to be explored.

In summary, whereas earlier work had demonstrated that movements are more vigorous toward more rewarding stimuli, here, we found that the RPE event that takes place in the moments before onset of a movement, and not reward in itself, is necessary for modulation of movement vigor.

References

- Aharon I, Etcoff N, Arieli D, Chabris CF, O'Connor E, Breiter HC (2001) Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron* 32:537–551.
- Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129–141.
- da Silva JA, Tecuapetla F, Paixão V, Costa RM (2018) Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature* 554:244–248.
- Ethier V, Zee DS, Shadmehr R (2008) Changes in control of saccades during gain adaptation. *J Neurosci* 28:13929–13937.
- Glaser JL, Wood DK, Lawlor PN, Ramkumar P, Kording KP, Segraves MA (2016) Role of expected reward in frontal eye field during natural scene search. *J Neurophysiol* 116:645–657.
- Haith AM, Reppert TR, Shadmehr R (2012) Evidence for hyperbolic temporal discounting of reward in control of movements. *J Neurosci* 32:11727–11736.
- Heitz RP, Schall JD (2012) Neural mechanisms of speed-accuracy tradeoff. *Neuron* 76:616–628.
- Herzfeld DJ, Vaswani PA, Marko MK, Shadmehr R (2014) A memory of errors in sensorimotor learning. *Science* 345:1349–1353.
- Herzfeld DJ, Kojima Y, Soetedjo R, Shadmehr R (2015) Encoding of action by the Purkinje cells of the cerebellum. *Nature* 526:439–442.
- Herzfeld DJ, Kojima Y, Soetedjo R, Shadmehr R (2018) Encoding of error and learning to correct that error by the Purkinje cells of the cerebellum. *Nat Neurosci* 21:736–743.
- Kawagoe R, Takikawa Y, Hikosaka O (1998) Expectation of reward modulates cognitive signals in the basal ganglia. *Nat Neurosci* 1:411–416.
- Kawagoe R, Takikawa Y, Hikosaka O (2004) Reward-predicting activity of dopamine and caudate neurons—a possible mechanism of motivational control of saccadic eye movement. *J Neurophysiol* 91:1013–1024.
- Kojima Y, Soetedjo R (2017a) Change in sensitivity to visual error in superior colliculus during saccade adaptation. *Sci Rep* 7:9566.
- Kojima Y, Soetedjo R (2017b) Selective reward affects the rate of saccade adaptation. *Neuroscience* 355:113–125.
- Kojima Y, Soetedjo R (2018) Elimination of the error signal in the superior colliculus impairs saccade motor learning. *Proc Natl Acad Sci U S A* 115: E8987–E8995.

- Kojima Y, Soetedjo R, Fuchs AF (2010) Changes in simple spike activity of some Purkinje cells in the oculomotor vermis during saccade adaptation are appropriate to participate in motor learning. *J Neurosci* 30:3715–3727.
- Manohar SG, Chong TT, Apps MA, Batla A, Stamelou M, Jarman PR, Bhatia KP, Husain M (2015) Reward pays the cost of noise reduction in motor and cognitive control. *Curr Biol* 25:1707–1716.
- Marino RA, Levy R, Munoz DP (2015) Linking express saccade occurrence to stimulus properties and sensorimotor integration in the superior colliculus. *J Neurophysiol* 114:879–892.
- Matsumoto M, Hikosaka O (2007) Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* 447:1111–1115.
- Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459:837–841.
- Milstein DM, Dorris MC (2007) The influence of expected value on saccadic preparation. *J Neurosci* 27:4810–4818.
- O'Doherty J, Winston J, Critchley H, Perrett D, Burt DM, Dolan RJ (2003) Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia* 41:147–155.
- Pekny SE, Criscimagna-Hemminger SE, Shadmehr R (2011) Protection and expression of human motor memories. *J Neurosci* 31:13829–13839.
- Reppert TR, Lempert KM, Glimcher PW, Shadmehr R (2015) Modulation of saccade vigor during value-based decision making. *J Neurosci* 35:15369–15378.
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Shadmehr R, Reppert TR, Summerside EM, Yoon T, Ahmed AA (2019) Movement vigor as a reflection of subjective economic utility. *Trends Neurosci* 42:323–336.
- Smalianchuk I, Jagadisan UK, Gandhi NJ (2018) Instantaneous midbrain control of saccade velocity. *J Neurosci* 38:10156–10167.
- Soetedjo R, Kojima Y, Fuchs AF (2008) Complex spike activity in the oculomotor vermis of the cerebellum: a vectorial error signal for saccade motor learning? *J Neurophysiol* 100:1949–1966.
- Summerside EM, Shadmehr R, Ahmed AA (2018) Vigor of reaching movements: reward discounts the cost of effort. *J Neurophysiol* 119:2347–2357.
- Tachibana Y, Hikosaka O (2012) The primate ventral pallidum encodes expected reward value and regulates motor action. *Neuron* 76:826–837.
- Xu-Wilson M, Zee DS, Shadmehr R (2009) The intrinsic value of visual information affects saccade velocities. *Exp Brain Res* 196:475–481.
- Yasuda M, Yamamoto S, Hikosaka O (2012) Robust representation of stable object values in the oculomotor basal ganglia. *J Neurosci* 32:16917–16932.
- Yoon T, Geary RB, Ahmed AA, Shadmehr R (2018) Control of movement vigor and decision making during foraging. *Proc Natl Acad Sci U S A* 115:E10476–E10485.