Effort cost of harvest affects decisions and movement vigor of marmosets during foraging

Paul Hage, In Kyu Jang, Vivian Looi, Mohammad Amin Fakharian, Simon P. Orozco,
Jay S. Pi, Ehsan Sedaghat-Nejad, and Reza Shadmehr

Laboratory for Computational Motor Control, Dept. of Biomedical Engineering
Johns Hopkins School of Medicine, Baltimore, Maryland USA

Correspondence: Paul Hage or Reza Shadmehr, Johns Hopkins School of Medicine, 410 Traylor Building, 720 Rutland Ave., Baltimore, MD 21205.
Email: paulhage@jhu.edu or shadmehr@jhu.edu

Acknowledgements: The work was supported by grants from the NIH (R01-EB028156, R01-NS078311), and the Office of Naval Research (N00014-15-1-2312).

Author contributions: P.H., S.P.O., E.S., J.S.P. and R.S. conceived and performed experiments. P.H., I.K.J., V.L., M.A.F., E.S., and J.S.P. analyzed data, P.H. made figures and performed statistical analysis. R.S. and P.H. wrote the manuscript.

Competing interests: None.

Abstract

We prefer to decline effortful rewards, but if the circumstances require it, we will move only slowly to harvest them. Why should economic variables such as reward and effort affect movement vigor? In theory, our decisions and movements both contribute to a measure of fitness in which the objective is to maximize rewards minus efforts, divided by time. To test this idea, we engaged marmosets in a foraging task in which on each trial they decided whether to work by making saccades to visual targets, thus accumulating food, or to harvest by licking. We varied the effort cost of harvest by moving the food tube with respect to the mouth. Theory predicted that the subjects would respond to the increased effort costs by working longer, stockpiling food before commencing harvest, but reduce their movement vigor to conserve energy. Indeed, in response to the increased effort costs of harvest, marmosets increased their work duration but reduced their movement vigor. These changes in decisions and movements coincided with changes in pupil size. As the effort cost of harvest declined, work duration decreased, the pupils dilated, and lick and tongue vigor increased. Thus, when acquisition of reward became effortful, there was a global change in the state of the brain: the pupils constricted, the decisions exhibited delayed gratification, and the movements displayed reduced vigor.

Significance statement

Why do economic variables such as reward and effort affect both the decision-making and the motor-control circuits of the brain? Our results suggest that as the brainstem neuromodulatory circuits that control pupil size respond to effort costs, they alter computations in brain regions that control decisions, encouraging work and delaying gratification, and brain regions that control movements, suppressing energy expenditure and reducing vigor. This coordinated response may improve a variable relevant to fitness: the capture rate.

Introduction

During foraging, animals work to locate a food cache and then spend effort harvesting what they have found. As they forage, their decisions appear to maximize a measure that is relevant to fitness: the sum of the rewards acquired, minus the efforts expended, divided by time, termed the capture rate (1–3). For example, a crow will spend effort extracting a clam from a sandy beach, but if the clam is small, it will abandon it because the additional time and effort required to harvest the small reward --dropping it repeatedly from a height onto rocks -- can be better spent finding a bigger prize (4). In other words, if going to the bank entails waiting in a long line, one should go infrequently, but make each transaction a large amount.

Intriguingly, reward expectation not only affects decisions, it also affects movements: we prefer the less effortful option, and move vigorously to obtain it (5, 6). This modulation of movement vigor can be justified if we consider that movements require expenditure of time and energy, which discount the value of the promised reward (7, 8). Thus, from a theoretical framework, it seems rational that the brain should have a mechanism to coordinate control of decisions with control of movements so that both contribute to maximizing a measure of fitness (9).

To study this coordination, we designed a task in which marmosets decided how long to work before they harvested their food. On a given trial, they made a sequence of saccades to visual targets and received an increment of food as their reward. However, the increment was small, and its harvest was effortful, requiring inserting the tongue inside a small tube. Theory predicted that in order to maximize the capture rate, harvest should commence when the reward magnitude justified the required effort. Indeed, the subjects chose to complete a few successive trials, stockpiling food, and only then initiated their harvest.

On some days the effort cost of harvest was low: the tube was placed close to the mouth. On other days the same amount of work, i.e., saccade trials, produced food that had a higher effort cost: the tube was located farther away. The theory made two interesting predictions: as the effort cost increased, the subjects should choose to work more trials, delaying their harvest so to stow more food, but reduce their movement vigor, thus saving energy. Indeed, when marmosets encountered an increased effort cost, they extended their work period, stockpiling food, but reduced their vigor. They slowed their saccades during the work period and slowed their licks during the harvest period.

What might be a neural basis for this coordinated response of the decision-making and the motor-control circuits? During the work and the harvest periods, momentary changes in pupil size closely tracked the changes in vigor: pupil dilation accompanied increases in vigor, while pupil constriction accompanied decreases in vigor, regardless of whether the movement that was being performed was a saccade, or a lick. Moreover, in response to the increased effort cost, the pupils exhibited a global change, constricting during both the work and the harvest periods.

If we view the changes in pupil size as a proxy for activity in the brainstem noradrenergic circuits (10), our results suggest that as these circuits respond to effort costs (11), they affect computations in the brain regions that control decisions, encouraging work and delayed gratification, and the brain regions that control movements, promoting sloth and energy conservation.

Results

We tracked the eyes and the tongue of head-fixed marmosets as they performed visually guided saccades in exchange for food (Fig. 1A). Each successful trial consisted of 3 visually guided saccades, at the end of which we delivered an increment of food (a slurry mixture of apple sauce and monkey chow). Because the reward amount was small (0.015-0.02 mL), the subjects rarely harvested following a single successful trial. Rather, they worked for a few trials, allowing the food to accumulate, then initiated their harvest by licking (Fig. 1B). The key variables were how many trials they chose to work before starting harvest, and how vigorously they moved their eyes and tongue during the work and the harvest periods.

Over the course of 2.5 years, we recorded 56 sessions in subject M (29 months) and 56 sessions in subject R (23 months). A typical work period lasted about 10 seconds, during which the subjects attempted ~8 trials, and succeeded in 4-5 trials (Fig. 1D) (a successful trial was when all three saccades were within 1.25° of the center of each target). The work period ended when the subjects stopped tracking the targets and initiated harvest, which lasted about 6 seconds, resulting in 16-18 licks. Subject M completed an average of 909.5 ± 61 successful trials per session (mean ± SEM), producing an average of 241 ± 13.9 work-harvest pairs, and subject R completed an average of 1431 ± 65 successful trials, producing an average of 263 ± 8.9 work-harvest pairs.

We delivered food via either the left or the right tube for 50-300 consecutive trials, and then switched tubes. We tracked the motion of the tongue using DeepLabCut (12), as shown for a typical session in Fig. 1B. The licks required precision because the tube was just large enough (4.4 mm diameter) to allow the tongue to penetrate. As a result, about 30% of the of the reward-seeking licks were successful and contacted food (30±1.6% for subject M, 28±2.5% for subject R), as shown in Video 1. Example of licks that failed to contact food are shown in Video 2, Video 3, and Video 4.

Theory and predictions

We imagined that our subjects coordinated their decisions and movements to maximize the sum of rewards acquired, minus efforts expended, divided by time, i.e., a capture rate. During a work period, they chose to complete a number of saccade trials n_s, a fraction β_s of which were successful, earning food increment α, but expending effort c_s and consuming time T_s for each trial. They then stopped working and initiated harvest, producing a number of licks n_l, a fraction β_l of which succeeded, expending effort c_l and consuming time T_l for each lick. These actions produced the following capture rate:

$$J = \frac{\alpha\beta_s n_s(1 - \frac{1}{1+\beta_l n_l}) - n_l c_l - n_s^2 c_s}{n_l T_l + n_s T_s} \quad (1)$$

In the numerator of Eq. (1), the first term represents the fact that the food cache increased linearly with successful trials and was then consumed gradually with successful licks. The second term represents the effort expenditure of licking, and the third term represents the effort expenditure of working. Notably, the effort expenditure of work, $n_s^2 c_s$, grows faster than linearly as a function of trials. This nonlinearity is essential to reflect the idea that following a long work period, the capture rate must be more negative than following a short work period (i.e., more work trials produce a greater loss).

A control policy describes how long to work and harvest, and an optimal policy produces periods of working and harvesting, n_s^*, n_l^*, that maximize Eq. (1). A closed-form solution for the optimal policy

can be obtained (*Mathematica* notebook simulations.nb) and Fig. 2A provides an example. As the work period concludes and the harvest period beings ($n_l = 0$), the capture rate is negative. This reflects the fact that the subject has performed a few trials and stockpiled food. The capture rate rises when licking commences, but the peak capture rate is not an increasing function of the work period. Rather, there is an optimal work period (n_s^* , red trace, Fig. 2A) associated with a given effort cost of licking c_l. If we now move the tube further, that is, increase c_l, the peak of the capture rate shifts and the optimal work period changes: the proper response to an increased effort cost of licking is to work longer, stowing more food.

Notably, the higher cost of licking inevitably reduces the maximum capture rate (Fig. 2A). This should impact movement vigor: animals tend to respond to a reduced capture rate by slowing their movements (9), which can be viewed as an effective way to save energy (8). To incorporate vigor into the capture rate, we tried to define the effort cost of a single lick c_l in terms of its metabolic cost, a relationship that is currently unknown. Fortunately, other movements provide a clue: the metabolic cost of reaching (8, 13), as well as the metabolic cost of walking (Ralston, 1958; Bastien et al., 2005), are both concave upward functions of the movement's duration. That is, from an energetic standpoint, there is a reach speed, and a walking speed, that minimizes the energetic cost of each type of movement. If we generalize these empirical observations to licking, a reasonable energetic cost associated with a single lick would be as follows:

$$c_l(T_l) = \frac{d^2}{T_l} + kT_l \tag{2}$$

In Eq. (2), the lick is aimed at a tube located at distance d, and has a duration T_l. The lick duration that minimizes the energetic cost is $d/\sqrt{k}$. Thus, the energetically optimal lick duration grows linearly with tube distance (Fig. 2B). However, our objective is not to minimize the cost of licking, but to maximize the capture rate. To do so, we insert Eq. (2) into Eq. (1) and find the optimal policy (n_s^*, n_l^*, T_l^*), which now depends on the distance of the food tube to the mouth (*Mathematica* notebook simulations.nb).

The theory predicts that to maximize the capture rate (Eq. 1), the response to an increased effort cost of harvest (i.e., tube distance) should be as follows: n_s^* should increase (Fig. 2C), n_l^* should decrease (Fig. 2C), and T_l^* should increase (Fig. 2D). Notably, the rate of increase in T_l^* as a function of tube distance is faster than linear, while from an energetic point of view (Eq. 2), increase in distance should produce a linear increase in lick duration. Thus, as the harvest becomes more effortful, the subject should work longer to stockpile food, but move slower.

To test our theory further, we thought it useful to have a way to alter decisions in one direction (say work longer) but change movement vigor in the opposite direction (move faster). In theory, this is possible: if the subject is hungry (darker lines in Fig. 2C and 2D), i.e., the reward is more valuable, then they should work longer before initiating harvest. Paradoxically, they should also move faster.

In summary, this simple theory made two sets of predictions: in response to increased cost of harvest, one should work longer, but move with reduced vigor. In response to increased reward value, as in hunger, one should also work longer, but now move with increased vigor.

Increased effort cost of harvest promoted work but reduced saccade vigor

To vary the effort cost of harvest, we altered the tube distance to the mouth but kept this distance constant during each session. Varying tube distance affected the decisions of the subjects: when the

tube was placed farther, they lengthened their work duration before starting harvest (Fig. 3A, left subplot): they attempted more trials during each work period (ANOVA, subject M: $F(2,7908)=41.5$ $p=5.2\times10^{-25}$, subject R: $F(2,10948)=88.2$ $p=7\times10^{-50}$), and produced more successful trials per work period (ANOVA, subject M: $F(2,7908)=63$ $p=2.8\times10^{-24}$, subject R: $F(2,10948)=163$ $p<10^{-50}$). This policy of delayed gratification was present throughout the recording session (Fig. 3A, middle plot). That is, when the harvest required more effort, the subjects worked longer, stockpiling food before initiating their harvest (Fig. 3A, right plot, effect of tube distance on food cached: subject M: $F(2,9566)=176$ $p<10^{-50}$, subject R: $F(2,8907)=204$ $p<10^{-50}$).

During the work period the subjects made saccades to visual targets and accumulated their food. They also made saccades that were not toward visual targets and thus were not eligible for reward. For each animal we computed the relationship between peak saccade velocity and saccade amplitude across all sessions and then calculated the vigor of each saccade: defined as the ratio of the actual peak velocity with respect to the expected peak velocity for that amplitude (16, 17). For example, a saccade that exhibited a vigor of 1.10 had a peak velocity that was 10% greater than the average peak velocity of the saccades of that amplitude for that subject. As expected, the reward-relevant saccades, i.e., saccades made to visual targets (primary, corrective, and center saccades) were more vigorous than other saccades (Fig. 3B, 2-way ANOVA, effect of saccade type, subject M: $F(1,391459)=7248$ $p<10^{-50}$, subject R: $F(1,355839)=13641$ $p<10^{-50}$).

As a work period began, the reward-relevant saccades exhibited high vigor, but then trial-by-trial, this vigor declined, reaching a low vigor value just before the work period ended (Fig. 3C vigor). Remarkably, on days in which the tube was placed farther, saccade vigor was lower (RMANOVA, effect of tube distance, subject M: $F(2,59033)=224$ $p<10^{-50}$, subject R: $F(2,50103)=75.51$ $p=1.8\times10^{-33}$). Thus, increasing the effort cost of extracting food during the harvest period produced reduced saccade vigor during the work period.

By definition, a more vigorous saccade had a greater peak velocity. This might imply that high vigor saccades should suffer from inaccuracy due to signal dependent noise (18). However, we observed the opposite tendency: as saccade vigor increased, both the magnitude and the variance of the endpoint error decreased (Fig. 3D, 2-way ANOVA, effect of vigor on error magnitude, subject M: $F(8,59046)=480$ $p<10^{-50}$, subject R: $F(8,50184)=252$ $p<10^{-50}$, effect of vigor on error variance, subject M: $F(8,2673)=18200$ $p<10^{-50}$, subject R: $F(8,2673)=4170$ $p<10^{-50}$). That is, reducing the effort costs of harvest not only promoted movement vigor, it also facilitated accuracy (19).

Cognitive signals such as effort requirements of the task and stimulus value are associated with changes in pupil size (10), as well as transient activation of brainstem neuromodulatory circuits in locus coeruleus (11). We wondered if the changes in tube position altered the output of these neuromodulatory circuits, as inferred via pupil size. For each reward-relevant saccade, we measured the pupil size during a ±250 ms window centered on saccade onset, and then normalized this measure based on the distribution of pupil sizes that we had measured during the entire recording for that session in that subject, resulting in a z-score.

At the onset of each work period the pupils were dilated, but as the subjects performed more trials, the pupils constricted, exhibiting a trial-by-trial reduction that paralleled the changes in saccade vigor (Fig. 3C). Notably, the effort cost of harvest affected pupil size: during the work period the pupils were more dilated if the tube was placed closer to the mouth (Fig. 3C, RMANOVA, effect of tube

distance, subject M: $F(2,60502)=20$ $p=2x10^{-9}$, subject R: $F(2,50431)=23.8$ $p=4.9x10^{-11}$). That is, when the effort cost of harvest was lower, the pupils dilated, and the saccades were invigorated.

In summary, when we increased the effort cost of harvest, the subjects responded by changing their decisions and movements. They delayed harvest onset, stockpiling more food before initiating their licking bout, and reduced their saccade vigor. These changes accompanied constriction of the pupil.

Increased effort cost of harvest reduced lick vigor

The work period ended when the subject chose to stop tracking the target and initiated harvest via a licking bout. As in saccades, we defined lick vigor via the ratio of the actual peak velocity of the lick with respect to the expected velocity for that lick amplitude. Lick peak velocity increased with amplitude during both protraction and retraction (Fig. 4A). Some of the licks were reward seeking and directed toward the tube, while others were grooming licks, cleaning the tongue and the area around the mouth (Video 5). Reward-seeking licks were more vigorous than grooming licks (2-way ANOVA, effect of lick type, protraction, subject M: $F(1,272233)=66$ $p=4.5x10^{-16}$, subject R: $F(1,229052)=698$ $p<10^{-50}$), and retraction was more vigorous than protraction (reward seeking licks, retraction vs. protraction, subject M: $t(241145)=532$ $p<10^{-50}$, subject R: $t(213674)=665$ $p<10^{-50}$).

As the harvest began, the first lick was very low vigor, but lick after lick, the movements gathered velocity, reaching peak vigor by the 3rd or the 4th lick (Fig. 4B). As the harvest continued, lick vigor gradually declined. Like saccades, licks had a lower vigor in sessions in which the tube was placed farther from the mouth (RMANOVA, effect of tube distance, subject M: $F(2,59033)=222.5$ $p<10^{-50}$, subject R: $F(2,133502)=224$ $p<10^{-50}$), and this pattern was present during the entire recording session (Fig. 4C, left subplot). Thus, an increased effort cost of harvest promoted sloth: reduced vigor of saccades during the work period, reduced vigor of licks during the harvest period.

For each reward seeking lick, we measured pupil size during a ±250ms window centered on the moment of peak tongue displacement. During licking, the pupil size changed with a pattern that closely paralleled lick vigor: as the harvest began, pupil size was small, but it rapidly increased during the early licks, then gradually declined as the harvest continued (Fig. 4B, right subplot). Importantly, the pupils were more dilated in sessions in which the tube was closer to the mouth (Fig. 4C, right subplot, effect of tube distance, subject M: $F(2,166742)=583$ $p<10^{-50}$, subject R: $F(2,130493)=118$ $p<10^{-50}$). As a result, when the effort cost of reward increased, the pupils constricted, and the vigor of both saccades and licks decreased.

While the theory predicted that moving the tube farther would result in a longer work period and reduced movement vigor, it also predicted that the subjects would reduce their harvest duration (reduced licks, Fig. 2C). That is, it predicted that the subjects would work longer, stowing more food, but leave more of it behind. This last prediction did not agree with our data (see Discussion). For Subject R, the number of licks were approximately the same across the various tube distances, and for Subject M the number of licks increased with tube distance (Supplementary Fig. 1).

In summary, within a harvest period, lick vigor rapidly increased and then gradually declined. Simultaneous with the changes in vigor, the pupils rapidly dilated and then gradually constricted. In sessions where the tube was placed farther from the mouth, the licks had lower vigor, and the pupils were more constricted.

Expectation of greater reward increased lick vigor

During the work period, as the subjects performed trials and accumulated food, they increased the magnitude of the available reward. To check whether reward magnitude affected movement vigor, for each tube distance we computed the average lick vigor during the harvest as a function of the number of trials completed in the preceding work period. Indeed, if a work period had included many completed trials, then the movements in the ensuing harvest period were more vigorous (Fig. 4D, 2-way ANOVA, effect of trials, subject M: $F(4, 164242)=353$ $p<10^{-50}$, subject R: $F(4,123411)=152$ $p<10^{-50}$). Thus, the licks were invigorated by the amount of food that awaited harvest.

Because the tube was small, many of the licks missed their goal and failed to contact the food. The success or failure of a lick affected both the vigor of the subsequent lick, and the change in the size of the pupil. Following a successful lick there was a large increase in lick vigor (Fig. 4E, subject M: $t(85182)=40$ $p<10^{-50}$, subject R: $t(81378)=104$ $p<10^{-50}$), and a large increase in pupil size (subject M: $F(84969)=57$ $p<10^{-50}$, subject R: $t(80318)=94$ $p<10^{-50}$). In contrast, following a failed lick the subjects either reduced or did not increase their lick vigor (Fig. 4E, subject M: $t(114159)=0.88$ $p=0.37$, subject R: $t(97164)=-44$ $p<10^{-50}$). This failure also produced a smaller increase in pupil size (comparison to successful lick, two sample t-test, subject M: $t(198722)=14.8$ $p=4.3\times10^{-50}$, subject R: $t(176044)=53$ $p<10^{-50}$). Thus, a single successful lick led to acquisition of reward, which then was followed by a relatively large increase in pupil size, and an invigorated subsequent lick.

Hunger promoted work and increased vigor

Our theory predicted that it should be possible to change decisions in one direction (say work longer), while altering movement vigor in the opposite direction (move faster). An increase in the subjective value of reward, as might occur when the subject is hungry, should have two effects: increase the number of trials that one chooses to work before commencing harvest, and increase movement vigor.

We did not explicitly manipulate the weight of the subjects. Indeed, to maintain their health, we strived to keep their weights constant during the roughly 2.5-year period of these experiments. However, there was natural variability, which allowed us to test the predictions of the theory.

We found that when their weight was lower than average, the subjects chose to work a greater number of trials before commencing harvest (Fig. 5A, two sample t-test, subject M: $t(11052)=7.9$ $p=3.4\times10^{-25}$, subject M: $t(12549)=10.1$ $p=9.3\times10^{-24}$). This result was similar to the effect that we had seen when the effort cost of harvest was increased. However, the theory had predicted that the effect on vigor should be in the opposite direction: if hunger increased reward valuation, then one should speed up to hasten food acquisition. Notably, weight did not have a consistent effect on saccade vigor across the two subjects (Fig. 5A), yet during the harvest, both subjects licked with greater vigor when their weight was lower (Fig. 5A, subject M: $t(219752)=88$ $p<10^{-50}$, subject R: $t(205163)=22$ $p<10^{-50}$).

Thus, while both the effort cost of reward and hunger promoted greater work and delay of harvest, effort promoted sloth while hunger promoted lick vigor.

Pupil size variations strongly correlated with changes in decisions and movements

Finally, we considered the data across both the work and the harvest periods and asked how well movement vigor tracked pupil size. The results demonstrated that in both the work and the harvest

periods, for both saccades and licks, an increase in pupil size was associated with an increase in vigor (Fig. 5B, reward-relevant saccades, subject M: $r=0.989$ $p=7.7\times10^{-9}$, subject R: $r=0.97$ $p=7.1\times10^{-7}$, reward-seeking protraction licks, subject M: $r=0.969$ $p=9.8\times10^{-7}$, subject R: $r=0.989$ $p=6.3\times10^{-9}$). Moreover, when the pupil was dilated, the work periods tended to be shorter (Fig. 5C, subject M: $r=-0.90$ $p=0.00014$, subject R: $r=-0.97$ $p=6.3\times10^{-7}$), while harvest durations tended to be longer (Fig. 5C, subject M: $r=0.894$ $p=0.00021$, subject R: $r=0.935$ $p=2.4\times10^{-5}$). Thus, pupil dilation was associated with choosing to work less, while moving faster.

Discussion

What we choose to do is the purview of the decision-making circuits of our brain, while the implicit vigor with which we perform that action is the concern of the motor-control circuits. From a theoretical perspective (9), these two forms of behavior should be coordinated because both the act that we select, and its vigor, dictate expenditure of time and energy, contributing to a capture rate that is critical for longevity and fecundity (20). Does the brain coordinate decisions and movements to maximize a capture rate? If so, how might the brain accomplish this coordination?

Here, we designed a foraging task in which marmosets worked by making saccades, accumulating food for each successful trial, then stopped working and harvested their cache by licking. On every trial they decided whether to work, or to harvest. Their decision was carried out by the motor system, producing either a visually guided saccade, or a lick, each exhibiting a particular vigor. The theory predicted that to maximize the capture rate, the appropriate response to an increased effort cost of harvest was to do two things: on the one hand work longer to store more food, but on the other hand reduce vigor to conserve energy.

We varied the effort costs by moving the food tube with respect to the mouth. This changed the effort cost of harvest but not the effort cost of work. The subjects responded by altering how they worked as well as how they harvested. When the harvest was more effortful, they performed more saccade trials to stockpile food. They also slowed their movements, reducing saccade velocity during the work period, and lick velocity during the harvest period. Notably, the most vigorous saccades were also the most accurate: as saccade vigor increased, so did endpoint accuracy.

The theory made a second prediction: as the value of reward increased, the subjects should again choose to work a longer period before initiating harvest, but unlike the effort costs, now respond by moving more vigorously. We did not directly manipulate the subjective value of reward, but rather relied on the natural fluctuations in body weight and assumed that when the weight was relatively low, there would be a greater hunger for reward. Indeed, when their weight was low, the subjects chose to not only work longer before initiating harvest, but also elevate their vigor during their harvest.

Finally, we quantified the effect of reward magnitude on vigor. In a given session, lick vigor increased robustly as a function of the number of trials completed in the preceding work period.

Notably, some of the predictions of the theory did not agree with the experimental data. An increased effort cost did not accompany a reduction in the duration of harvest, and hunger did not increase saccade vigor robustly. Indeed, earlier experiments have shown that if the effort cost of harvest increases, animals who expend the effort will then linger longer to harvest more of the reward that they have earned (2). This mismatch between observed behavior and theory highlights some of the limitations of our formulation. For example, our capture rate reflected a single work-harvest period,

rather than a long sequence. Moreover, the capture rate did not consider the fact that the food tube had finite capacity, beyond which the food would fall and be wasted. This constraint would discourage a policy of working more but harvesting less. Finally, if we assume that a reduced body weight is a proxy for increased subjective value of reward, it is notable that we observed a robust effect on vigor of licks, but not saccades. A more realistic capture rate formulation awaits simulations, possibly one that describes capture rate not as the ratio of two sums (sum of gains and losses with respect to sum of time), but rather the expected value of the ratio of each gain and loss with respect to time (21, 22).

What might be a neural basis for this coordination of decisions and movements? A clue was the fact that the pupils were more constricted in sessions in which the effort cost of harvest was greater. This global change in pupil size accompanied delayed harvest and reduced vigor across sessions, but surprisingly, even within a session, transient changes in vigor accompanied changes in pupil size. During the work period the trial-to-trial reduction in saccade vigor accompanied trial-to-trial constriction of the pupil, and within a harvest period, the rapid rise and then the gradual fall in lick vigor paralleled a rapid dilatation then gradual constriction of the pupil.

Pupil dilation is a proxy for activity in the brainstem neuromodulatory system (23) and is a measure of arousal (24). Control of pupil size is dependent on spiking of norepinephrine neurons in locus coeruleus (LC-NE): an increase in the activity of these neurons produces pupil dilation (25, 26). Some of these neurons show a transient change in their activity when acquisition of reward requires expenditure of physical effort (11). It is possible that in the present task, as the effort cost of harvest increased, LC-NE neurons decreased their activity, producing pupil constriction. If so, the reduced NE release may have had two simultaneous effects: encourage work and promote delayed gratification in brain regions that control decisions, discourage energy expenditure and promote sloth in brain regions that control movements. Thus, the idea that emerges is that the response of NE to economic variables might act as a bridge to coordinate the computations in the decision-making circuits with the computations in the motor-control circuits, aiming to improve the capture rate.

In addition to NE, the basal ganglia, and in particular the neurotransmitter dopamine, are likely the key contributors to the coordination of decisions with actions (27, 28). When the effort price of a preferred food increases, animals choose to work longer, pressing a lever a greater number of times (29, 30). This desire to expend effort to acquire a valuable reward is reduced if dopamine is blocked in the ventral striatum (31–33). Hunger activates circuits in the hypothalamic nuclei, disinhibiting dopamine release in response to food cues (34). Dopamine concentrations in the striatum drop when the effort price of a food reward increases (35), and dopamine release before onset of a movement tends to invigorate that movement (36). Thus, the presence of dopamine is likely to not only alter decisions by encouraging expenditure of effort, but also modify movements by promoting vigor.

Experiments of Hayden et al. (37) and Barack et al. (38) suggest that the decision of when to stop work and commence harvest may rely on computations that are carried out in the cingulate cortex. They found that as monkeys deliberated between the choice of staying and acquiring diminishing rewards, or leaving and incurring a travel cost, these neurons encoded a decision-variable that reflected the value of leaving the patch. The prediction that emerges from our work is that the rate of rise of these decision variables may be modulated by the presence of NE.

From a motor control perspective, a surprising aspect of our results was that an increase in saccade vigor accompanied improved accuracy (i.e., reduced endpoint variance). The high vigor

saccades were produced at a time when the pupils were dilated, possibly implying an increased release of NE. Control of saccade accuracy appears to depend on the coordinated activity of Purkinje cells in the oculomotor region of the cerebellar vermis (39). LC projects to the cerebellum, and stimulation of LC neurons increases the sensitivity of Purkinje cells to their inputs (40). An intriguing possibility that remains to be tested is that movement accuracy is controlled in the cerebellum via activities among populations of Purkinje cells, which in turn are modulated by the NE projections from LC.

Methods

Behavioral and neurophysiological data were collected from two marmosets (*Callithrix Jacchus*, male and female, 350-390 g, subjects R and M, 6 years old). The neurophysiological data focused on the cerebellum and are described elsewhere (39, 41). Here, our focus is on the behavioral data.

The marmosets were born and raised in a colony that Prof. Xiaoqin Wang has maintained at the Johns Hopkins School of Medicine since 1996. The procedures on the marmosets were evaluated and approved by the Johns Hopkins University Animal Care and Use Committee in compliance with the guidelines of the United States National Institutes of Health.

Data acquisition

Following recovery from head-post implantation surgery, the animals were trained to make saccades to visual targets and rewarded with a mixture of apple sauce and lab diet (41). They were placed in a monkey chair and head-fixed while we presented visual targets on an LCD screen (Curved MSI 32" 144 Hz - model AG32CQ) and tracked both eyes at 1000 Hz using an EyeLink-1000 system (SR Research, USA). Timing of target presentation on the video screen was measured using a photo diode. Tongue movements were tracked with a 522 frame per second Sony IMX287 FLIR camera, with frames captured at 100 Hz.

Each trial began with a saccade to the center target followed by fixation for 200 ms, after which a primary target (0.5x0.5 deg square) appeared at one of 8 randomly selected directions at a distance of 5-6.5 deg. Onset of the primary target coincided with presentation of a tone. As the animal made a saccade to the primary target, that target was erased, and a secondary target was presented at a distance of 2-2.5 deg, also at one of 8 randomly selected directions. The subject was rewarded if following the primary saccade, it made a corrective saccade to the secondary target, landed within 1.5 deg radius of the target center, and maintained fixation for at least 200 ms. Onset of reward coincided with presentation of another tone. Following an additional 150-250 ms period (uniform random distribution), the secondary target was erased, and the center target was displayed, indicating the onset of the next trial. Thus, a successful trial was comprised of a sequence of 3 saccades: center, primary, and corrective, after which the subject received a small increment of food (0.015 mL).

The food was provided in two small tubes (4.4 mm diameter), one to the left and the other to the right of the animal (Fig. 1A). A successful trial produced a food increment in one of the tubes and would continue to do so for 50-300 consecutive trials, then switch to the other tube. Because the food increment was small, the subjects naturally chose to work for a few consecutive trials, tracking the visual targets and allowing the food to accumulate, then stopped tracking and harvested the food via a licking bout. The licking bout typically consisting of 15-40 licks. The subjects did not work while harvesting. As a result, the behavior consisted of a work period of targeted saccades, following by a harvest period of targeted licking, repeated hundreds of times per session.

Data analysis

All saccades, regardless of whether they were instructed by presentation of a visual target or not, were identified using a velocity threshold. Saccades to primary, secondary, and central targets were labeled as reward-relevant saccades, while all remaining saccades were labeled as task irrelevant.

We analyzed tongue movements using DeepLabCut (12). Our network was trained on 89 video recordings of the subjects with 15-25 frames extracted and labeled from each recording. The network was built on the ResNet-152 pre-trained model, and then trained over 1.03×10^6 iterations with a batch size of 8, using a GeForce GTX 1080Ti graphics processing unit (42). A Kalman filter was further applied to improve quality and smoothness of the tracking, and the output was analyzed in MATLAB to quantify varying lick events and kinematics.

We tracked the tongue tip and the edge of the food in the tube, along with control locations (nose position and tube edges). We tracked all licks, regardless of whether they were aimed toward the tube (reward seeking), or not (grooming). Reward seeking licks were further differentiated based on whether they aimed to enter the tube (inner-tube licks), hit the outer edge of the tube (outer-edge licks), or fell below the tube (under tube). If any of these licks successfully contacted the tube, we labeled that lick as a success (otherwise, a failed lick).

Pupil area was measured during a 500 ms period centered at the onset of each reward-relevant saccade, and the onset of each lick. We then normalized the pupil measurements by representing it as a z-score over that session.

Saccade and tongue vigor

We relied on previous work to define vigor of a movement (5, 9, 16, 17). Briefly, if the amplitude of a movement is x and the peak speed of that movement is v, then for each subject the relationship between the two variables was assumed to be:

$$v = \alpha\left(1 - \frac{1}{1 + \beta x}\right) \tag{3}$$

In the above expression, $\alpha, \beta \geq 0$ and are subject-specific parameters. For a movement with amplitude x, its vigor was defined as the ratio of the actual peak speed with respect to the expected value of its peak speed, i.e., $\frac{v}{E[v(x)]}$. Expected value was computed by fitting Eq. (3) to all the data acquired across all sessions. When vigor is greater than 1, the movement had a peak velocity that was higher than the mean value associated with that amplitude.

Simulations

The optimal policy specifies the decisions and movements that for the effort cost defined in Eq. (2), maximizes the capture rate defined in Eq. (1). This policy selects the number of saccade trials n_s to perform during the work period, the number of licks n_l to perform during the harvest period, and the vigor of each lick, represented by the average duration of a lick T_l. To compute the optimal policy, we found the derivative of the capture rate with respect to each policy variable n_s, n_l, and T_l, then set each derivative equal to zero, producing three simultaneous nonlinear equations. In all three cases, we were able to solve for the relevant control variable analytically (see Mathematica notebook *simultatons.nb* for the derivations). We found that if the solution was a real, an increase in d produced an increase in n_S^*, decrease in n_L^*, and increase in T_L^*. To generate the plots in Fig. 2A, we used the following parameter values: $\alpha = 20$, $\beta_s = 0.5$, $\beta_L = 0.3$, $c_s = 0.5$, $T_s = 1$, $T_L = 0.2$, $c_L = 0.5$ (low effort), $c_L = 2.5$ (high effort). For the plots in Fig. 2C and 2D, we used the same parameter values, but c_L was defined via Eq.

(2). Thus, tube distance d varied, and T_L was unknown and was solved for. In Eq. (2), $k_L = 1$. In the simulations, to describe state of hunger, we set $\alpha = 20$ for a sated state and $\alpha = 25$ for a hungry state.

Statistical analysis

Hypothesis testing was performed using functions provided by the *MATLAB Statistics and Machine Learning Toolbox*. For t-tests, across the one-sample, paired-sample, and two-sample conditions, p values were computed using the *ttest* and *ttest2* functions with data that was combined across sessions, separated by condition. For ANOVA tests, in the one-way condition, p values were computed using a nonparametric Kruskal-Wallis test, using the *kruskalwallis* function. In the 2-way condition, the *anovan* function was used to compute p values, accounting for an unbalanced design resulting from a varied number of samples across conditions. In both cases, like in the t-tests, data was combined across sessions, separated by condition. In the repeated measures condition, each session was treated as a subject with multiple repeated measures representing a given variable (i.e lick vigor per lick in a harvest period). To fit a repeated measures model, the *fitrm* function was used, then analyzed using the *ranova* function. In all cases of repeated measures ANOVA, compound symmetry assumptions were tested using the Mauchly sphericity test with the *maulchy* function. In cases where the assumption was violated (Maulchy test $p < 0.05$), epsilon adjustments were used, with the *epsilon* function, to compute corrected p values (for $\varepsilon > .75$, use Huynh-Feldt p value and for $\varepsilon < .75$, use Greenhouse-Geisser p values). For correlation analyses, Pearson's correlation coefficient, r, and corresponding p values were computed using the *corrcoef* function.

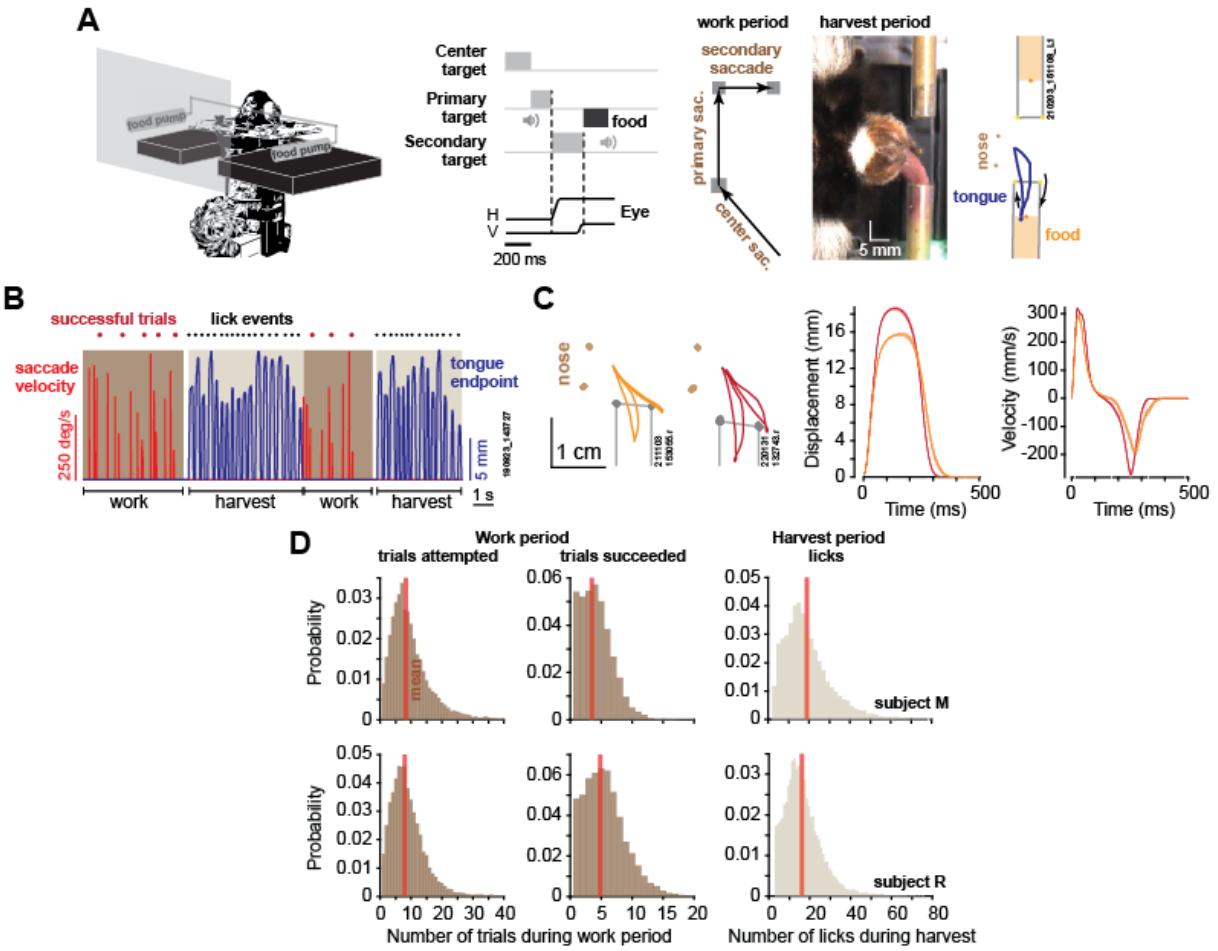


Figure 1. Foraging task. **A**. During the work period, marmosets made a sequence of saccades to visual targets. A trial consisted of 3 consecutive saccades, at the end of which the subject was rewarded by an increment of food. We tracked the eyes, the tongue, and the food. **B**. An example of two consecutive work-harvest periods, showing reward-relevant saccades (eye velocity) and tongue endpoint displacement with respect to the mouth. **C**. The figure shows data for two sessions, one where the tube was placed close to the mouth (orange trace), and one where it was placed farther away (red trace). Two types of licks are shown: inner-tube licks, and outer-tube licks. Depending on food location, both types of licks can contact the food. Data on the right two panels show endpoint displacement and velocity of the tongue during inner-tube licks. Error bars are SEM. **D**. During the work period, the subjects attempted about 8 trials on average, succeeding in 4-5 trials before starting harvest, and then licked about 18 times to extract the food.

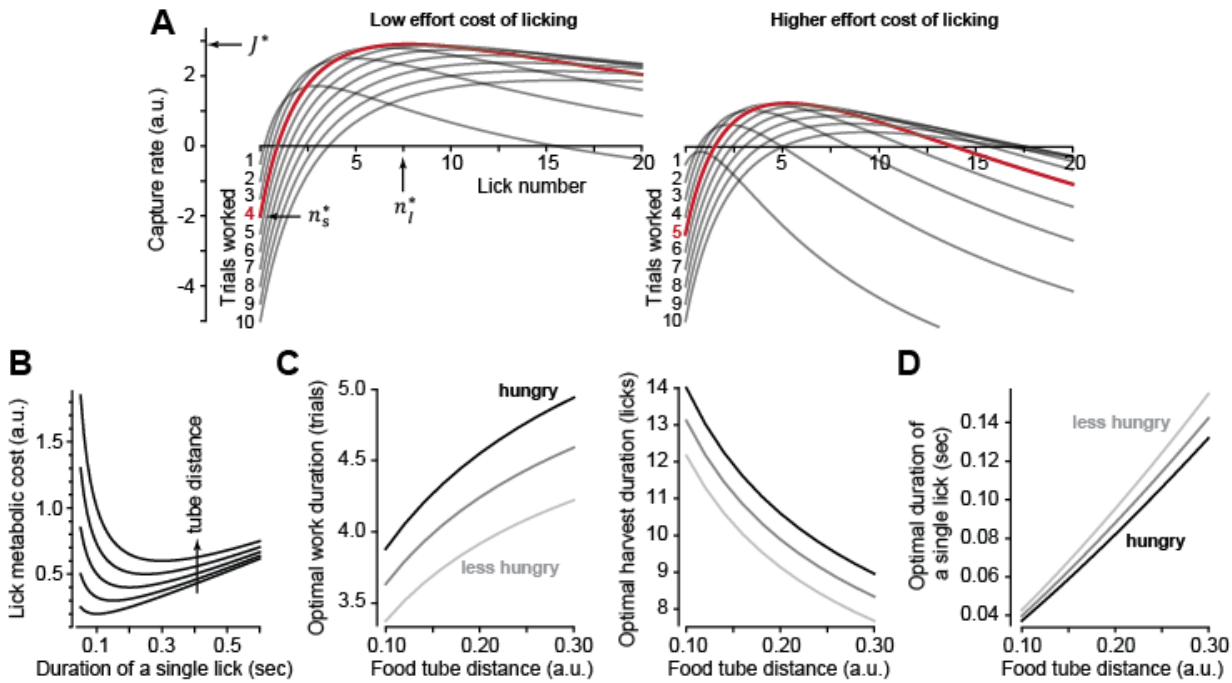


Figure 2. Simulation results. **A**. Capture rate (Eq. 1) is plotted during the harvest period as a function of lick number n_l following various number of work trials n_s. When the effort cost of licking is low (left plot, $c_l = 0.5$), the optimal work period is $n_s^* = 4$ (red trace). When the effort cost is higher (right plot, $c_l = 2$), it is best to work longer before initiating harvest. **B**. The metabolic cost of licking (Eq. 2) is minimized when a lick has a specific duration. Tube distance varied from 0.1 to 0.3. Optimal duration that minimizes lick cost grows linearly with tube distance. **C**. Optimal number of work trials n_s^* and licks n_l^* as a function food tube distance d. As the effort cost of harvest increases, one should respond by working longer, delaying harvest. **D**. Optimal lick duration T_l^* as a function of food tube distance. The lick duration T_l^* that maximizes the capture rate is smaller than the one that minimizes the lick metabolic cost (Fig. 2B). That is, it is worthwhile moving vigorously to acquire reward. However, T_l^* grows faster than linearly as a function of tube distance. Thus, as the tube moves farther, it is best to reduce lick vigor. Hunger should promote work and increase vigor, while effort cost of harvest (tube distance) should promote work but reduce vigor. Parameter values for all simulations: $\beta_s = 0.5$, $\beta_l = 0.3$, $T_s = 1$, $k = 1$, $\alpha = 20$ (low food value, less hunger), $\alpha = 25$ (high food value, hungry).

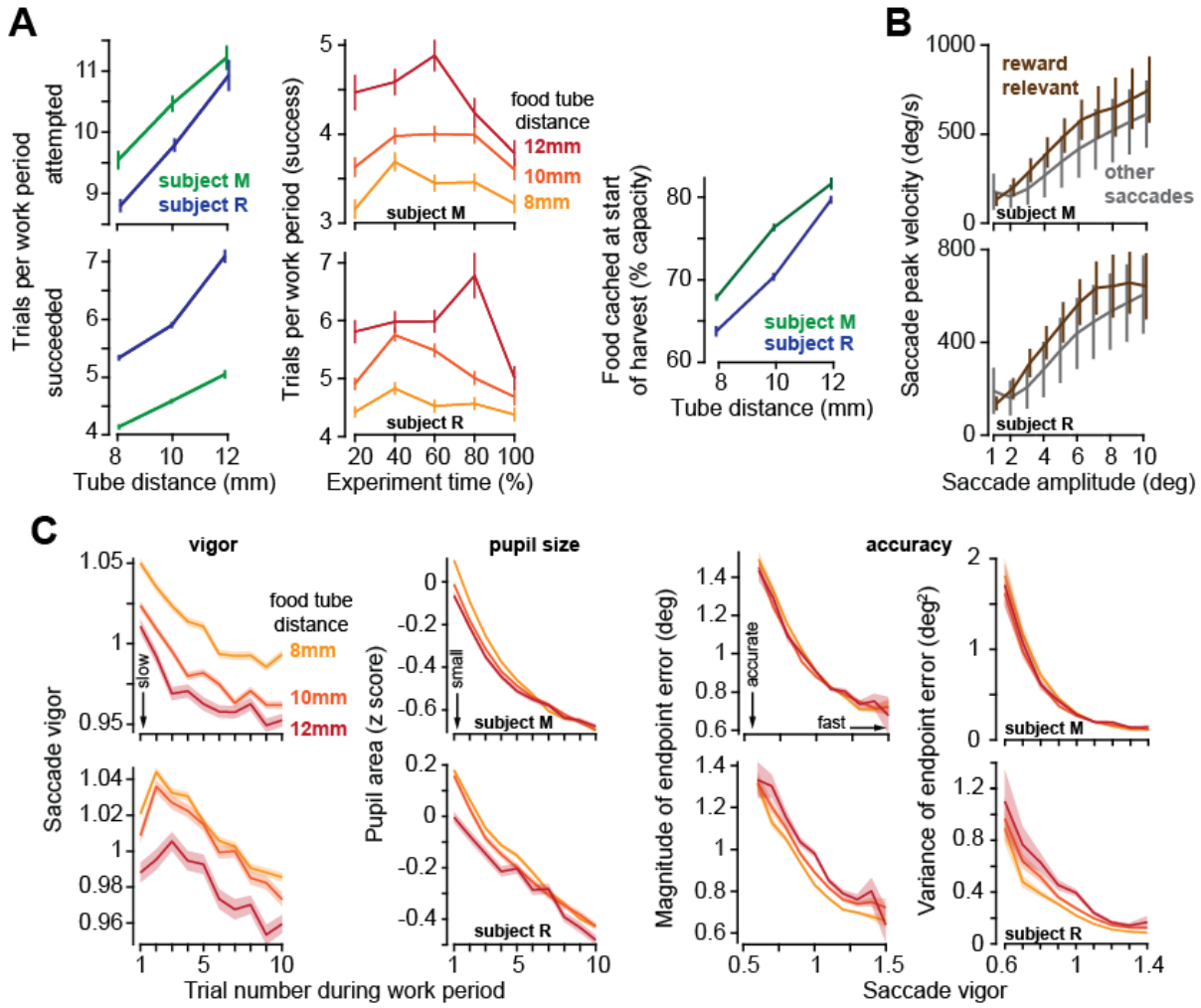


Figure 3. As the effort cost of harvest increased, subjects worked longer to stow more food, and saccade vigor during the work period declined. **A**. Left: the number of trials attempted and succeeded per work period as a function of tube distance. Middle: successful trials per work period as a function of time during the recording session. Tube distance is with respect to a marker on the nose. Right: food available in the tube at the start of the harvest. **B**. Peak saccade velocity as a function of amplitude for reward-relevant and other saccades. **C**. Vigor of reward relevant saccades as a function of trial number during the work period. Vigor was greater when the tube was closer. Pupil size is quantified during the same work periods. Accuracy is quantified as the magnitude of the saccade's endpoint error vector (to the target) and the variance of that error vector (determinant of the variance-covariance matrix), plotted as a function of the vigor of the saccade (bin size=0.05 vigor units). Error bars are SEM.

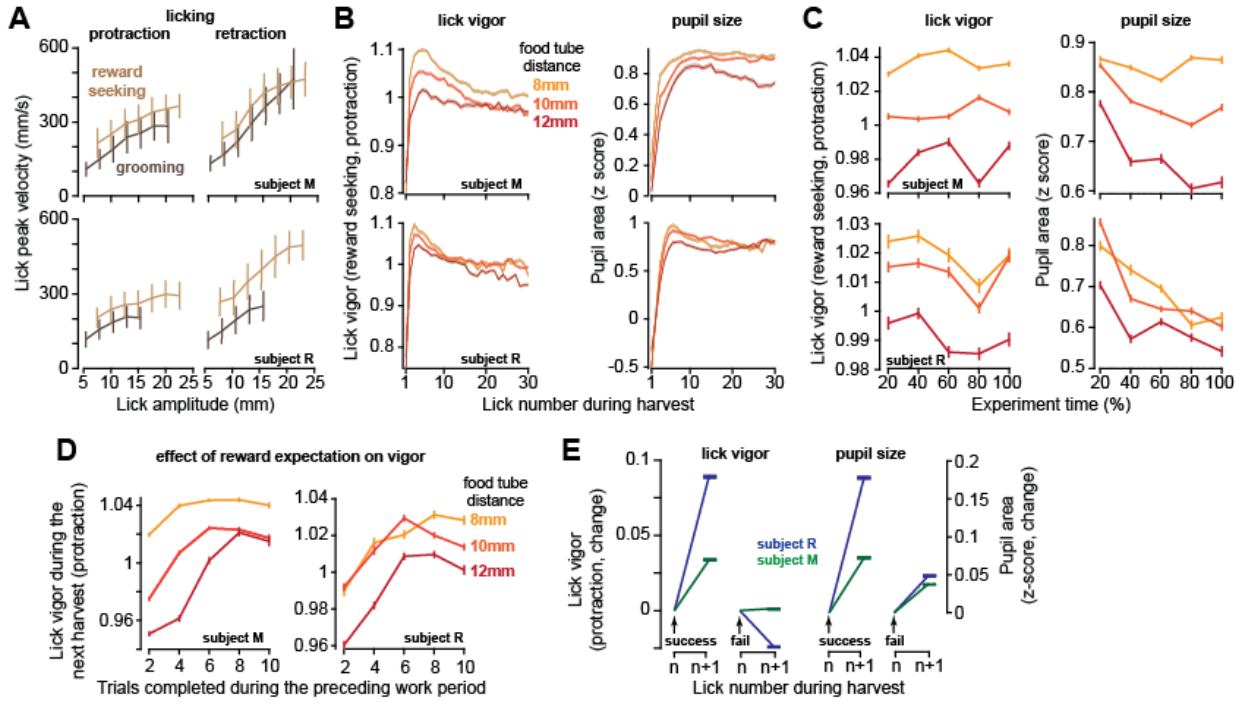


Figure 4. As the effort cost of harvest increased, lick vigor declined, and the pupils constricted. **A**. Peak speed of reward-seeking and grooming licks during protraction and retraction as a function of lick amplitude. **B**. Vigor of reward seeking licks (protraction) and pupil size as a function of lick number during harvest at various tube distances. **C**. Lick vigor and pupil size as a function of time during the entire recording session. Line colors depict tube distance as in part B. **D**. Average lick vigor and pupil size during a harvest as a function of number of trials successfully completed in the previous work period. Lick vigor and pupil size were greater during the harvest when more food had been stored. **E**. Following a successful lick (contact with food), the next lick was more vigorous, and pupils dilated. Following a failed lick, the next lick was slowed, and pupils were less dilated. Error bars are SEM.

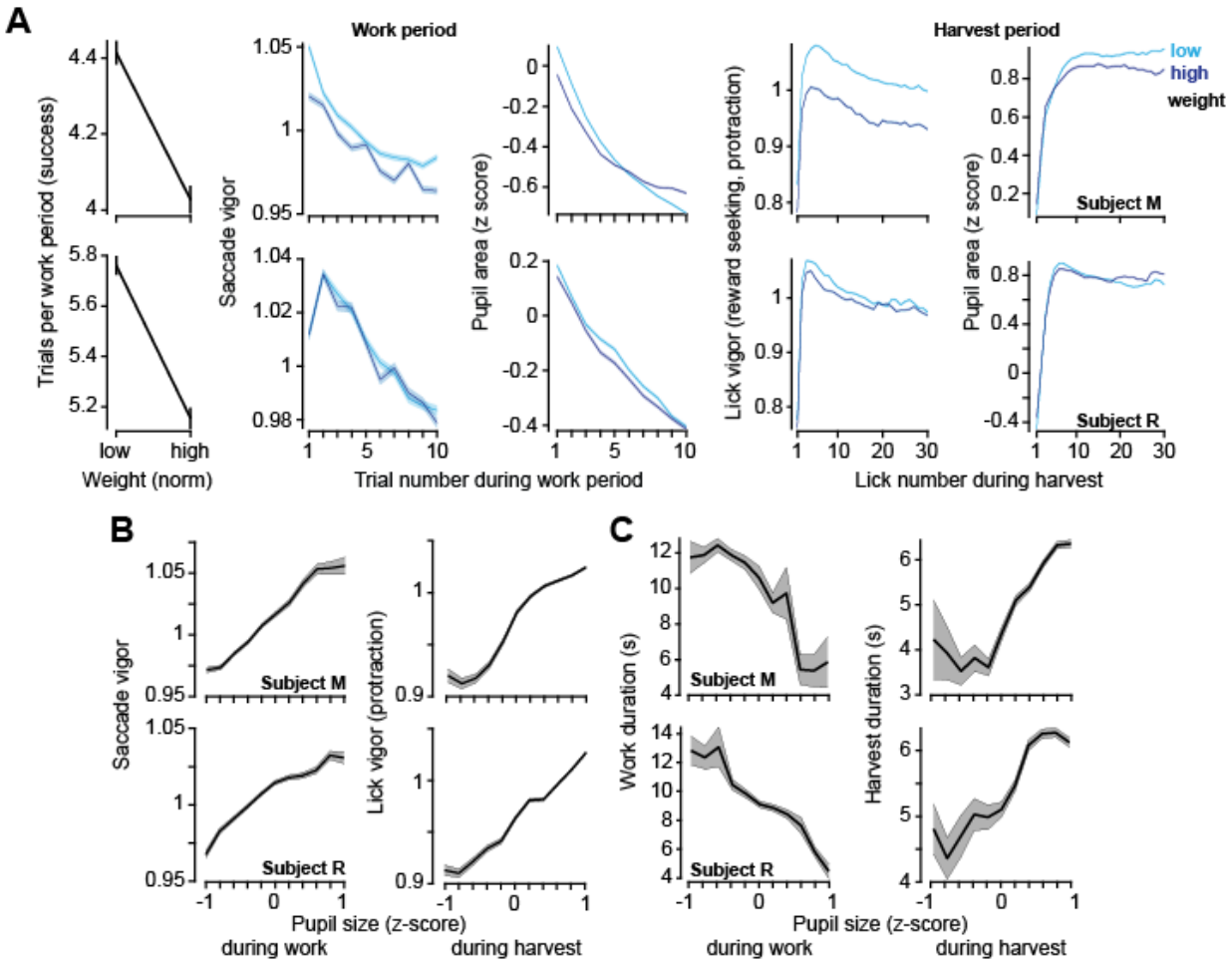
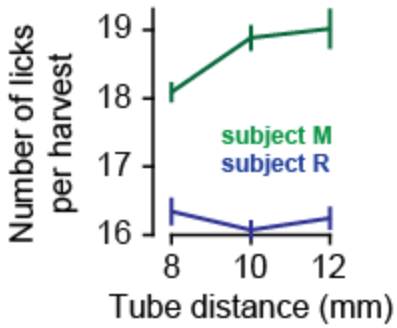


Figure 5. Hunger promoted vigor, and pupil size correlated with both vigor and decisions. **A**. Trials successfully completed during a work period as a function of normalized body weight at the start of the session. **B**. Left: lick vigor as a function of lick number as a function of body weight. Right: lick vigor as a function of body weight and tube distance. **C**. Saccade vigor during the work period, and lick vigor during the harvest period, as a function of pupil size. **D**. Work duration and harvest duration as a function of pupil size. Error bars are SEM.



Supplementary Fig. 1. Number of licks per harvest as a function of tube distance.

Video 1. Example of a successful inner-tube lick.
Video 2. Example of an under-tube lick that failed to contact food.
Video 3. Example of an outer-tube lick that failed to contact food.
Video 4. Example of a lick that hit the outer edge of the tube and failed to contact food.
Video 5. Example of a grooming lick.

1. E. L. Charnov, Optimal foraging, the marginal value theorem. *Theor.Popul.Biol.* **9**, 129–136 (1976).

2. R. J. Cowie, Optimal foraging in great tits (Parus major). *Nature* **268**, 137–139 (1977).

3. R. Shadmehr, A. A. Ahmed, *Vigor: neuroeconomics of movement control* (MIT Press, 2020).

4. H. Richardson, N. A. M. Verbeek, Diet selection and optimization by northwestern crows feeding on Japanese littleneck clams. *Ecology* **67**, 1219–1226 (1986).

5. T. Yoon, A. Jaleel, A. A. Ahmed, R. Shadmehr, Saccade vigor and the subjective economic value of visual stimuli. *J. Neurophysiol.* **123**, 2161–2172 (2020).

6. C. C. Korbisch, D. R. Apuan, R. Shadmehr, A. A. Ahmed, Saccade vigor reflects the rise of decision variables during deliberation. *Curr. Biol.* (2022) https:/doi.org/10.1016/j.cub.2022.10.053 (December 4, 2022).

7. R. Shadmehr, J. J. Orban de Xivry, M. Xu-Wilson, T. Y. Shih, Temporal discounting of reward and the cost of time in motor control. *J.Neurosci.* **30**, 10507–10516 (2010).

8. R. Shadmehr, H. J. Huang, A. A. Ahmed, A representation of effort in decision-making and motor control. *Curr.Biol.* **26**, 1929–1934 (2016).

9. T. Yoon, R. B. Geary, A. A. Ahmed, R. Shadmehr, Control of movement vigor and decision making during foraging. *Proc.Natl.Acad.Sci.U.S.A.* **115**, E10476–E10485 (2018).

10. S. Joshi, J. I. Gold, Pupil Size as a Window on Neural Substrates of Cognition. *Trends Cogn. Sci.* **24**, 466–480 (2020).

11. P. Bornert, S. Bouret, Locus coeruleus neurons encode the subjective difficulty of triggering and executing actions. *PLOS Biol.* **19**, e3001487 (2021).

12. A. Mathis, *et al.*, DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).

13. H. J. Huang, A. A. Ahmed, Older adults learn less, but still reduce metabolic cost, during motor adaptation. *J.Neurophysiol.* **111**, 135–144 (2014).

14. H. J. RALSTON, Energy-speed relation and optimal speed during level walking. *Int.Z.Angew.Physiol.* **17**, 277–283 (1958).

15. G. J. Bastien, P. A. Willems, B. Schepens, N. C. Heglund, Effect of load and speed on the energetic cost of human walking. *Eur.J.Appl.Physiol* **94**, 76–83 (2005).

16. T. R. Reppert, K. M. Lempert, P. W. Glimcher, R. Shadmehr, Modulation of Saccade Vigor during Value-Based Decision Making. *J.Neurosci.* **35**, 15369–15378 (2015).

17. T. R. Reppert, *et al.*, Movement vigor as a trait-like attribute of individuality. *J.Neurophysiol.* **120**, 741–757 (2018).

18. C. M. Harris, D. M. Wolpert, Signal-dependent noise determines motor planning. *Nature* **394**, 780–784 (1998).

19. C. Wang, Y. Xiao, E. Burdet, J. Gordon, N. Schweighofer, The duration of reaching movement is longer than predicted by minumum variance. *J.Neurophysiol.* **116**, 2342–2345 (2016).

20. W. C. Lemon, Fitness consequences of foraging behaviour in the zebra finch. *Nature* **352**, 153–155 (1991).

21. M. Bateson, A. Kacelnik, Preferences for fixed and variable food sources: variability in amount and delay. *J.Exp.Analysis.Behav.* **63**, 313–329 (1995).

22. M. Bateson, A. Kacelnik, Rate currencies and the foraging starling: the fallacy of the averages revisited. *Behav. Ecol.* **7**, 341–352 (1996).

23. E. M. Vazey, D. E. Moorman, G. Aston-Jones, Phasic locus coeruleus activity regulates cortical encoding of salience information. *Proc. Natl. Acad. Sci.* **115**, E9439–E9448 (2018).

24. S. Mathot, Pupillometry: Psychology, Physiology, and Function. *J. Cogn.* **1**, 1–23 (2018).

25. S. Joshi, Y. Li, R. M. Kalwani, J. I. Gold, Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron* **89**, 221–234 (2016).

26. V. Breton-Provencher, M. Sur, Active control of arousal by a locus coeruleus GABAergic circuit. *Nat. Neurosci.* **22**, 218–228 (2019).

27. D. Thura, P. Cisek, The Basal Ganglia Do Not Select Reach Targets but Control the Urgency of Commitment. *Neuron* **95**, 1160–1170 (2017).

28. D. M. Herz, *et al.*, Dynamic control of decision and movement speed in the human basal ganglia. *Nat. Commun.* **13**, 7530 (2022).

29. J. D. Salamone, *et al.*, Haloperidol and nucleus accumbens dopamine depletion suppress lever pressing for food but increase free food consumption in a novel food choice procedure. *Psychopharmacol. Berl* **104**, 515–521 (1991).

30. J. E. Aberman, J. D. Salamone, Nucleus accumbens dopamine depletions make rats more sensitive to high ratio requirements but do not impair primary food reinforcement. *Neuroscience.* **92**, 545–552 (1999).

31. M. Koch, A. Schmid, H. U. Schnitzler, Role of nucleus accumbens dopamine D1 and D2 receptors in instrumental and Pavlovian paradigms of conditioned reward. *Psychopharmacol. Berl* **152**, 67–73 (2000).

32. A. M. Farrar, *et al.*, Nucleus accumbens and effort-related functions: behavioral and neural markers of the interactions between adenosine A2A and dopamine D2 receptors. *Neuroscience* **166**, 1056–1067 (2010).

33. S. E. Yohn, *et al.*, The role of dopamine D1 receptor transmission in effort-related choice behavior: Effects of D1 agonists. *Pharmacol.Biochem.Behav.* **135**, 217–226 (2015).

34. R. M. Cassidy, Q. Tong, Hunger and Satiety Gauge Reward Sensitivity. *Front. Endocrinol.* **8** (2017).

35. S. A. Schelp, *et al.*, A transient dopamine signal encodes subjective value and causally influences demand in an economic context. *Proc.Natl.Acad.Sci.U.S.A* **114**, E11303–E11312 (2017).

36. J. A. da Silva, F. Tecuapetla, V. Paixao, R. M. Costa, Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature* **554**, 244–248 (2018).

37. B. Y. Hayden, J. M. Pearson, M. L. Platt, Neuronal basis of sequential foraging decisions in a patchy environment. *Nat.Neurosci.* **14**, 933–939 (2011).

38. D. L. Barack, S. W. C. Chang, M. L. Platt, Posterior Cingulate Neurons Dynamically Signal Decisions to Disengage during Foraging. *Neuron* **96**, 339–347 (2017).

39. E. Sedaghat-Nejad, J. S. Pi, P. Hage, M. A. Fakharian, R. Shadmehr, Synchronous spiking of cerebellar Purkinje cells during control of movements. *Proc. Natl. Acad. Sci.* **119**, e2118954119 (2022).

40. H. C. Moises, B. D. Waterhouse, D. J. Woodward, Locus coeruleus stimulation potentiates Purkinje cell responses to afferent input: The climbing fiber system. *Brain Res.* **222**, 43–64 (1981).

41. E. Sedaghat-Nejad, *et al.*, Behavioral training of marmosets and electrophysiological recording from the cerebellum. *J.Neurophysiol.* **122**, 1502–1517 (2019).

42. K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition in (2016), pp. 770–778.