

Current Biology

A Representation of Effort in Decision-Making and Motor Control

Highlights

- In reaching, like walking, there is a movement speed that minimizes energetic cost
- Reward makes it worthwhile to be energetically inefficient
- Effort may be represented objectively via energetic cost and discounted in time
- Neural control of decisions and movements may share a common utility

Authors

Reza Shadmehr, Helen J. Huang, Alaa A. Ahmed

Correspondence

alaa.ahmed@colorado.edu

In Brief

Both decisions and movements are influenced by reward and effort, suggesting that they may share a common utility. Shadmehr et al. demonstrate that a utility in which effort is objectively represented as energetic cost and discounted in time can account for both the choices animals make and the vigor of their movements.



A Representation of Effort in Decision-Making and Motor Control

Reza Shadmehr,¹ Helen J. Huang,^{2,3} and Alaa A. Ahmed^{2,*}

¹Laboratory for Computational Motor Control, Department of Biomedical Engineering, Johns Hopkins School of Medicine, 410 Traylor Building, 720 Rutland Avenue, Baltimore, MD 21205, USA

²Neuromechanics Laboratory, Department of Integrative Physiology, University of Colorado, 354 UCB, Boulder, CO 80309-0354, USA

³Present address: Department of Mechanical and Aerospace Engineering, University of Central Florida, Orlando, FL 32816-2450, USA

*Correspondence: alaa.ahmed@colorado.edu

<http://dx.doi.org/10.1016/j.cub.2016.05.065>

SUMMARY

Given two rewarding stimuli, animals tend to choose the more rewarding (or less effortful) option. However, they also move faster toward that stimulus [1–5]. This suggests that reward and effort not only affect decision-making, they also influence motor control [6, 7]. How does the brain compute the effort requirements of a task? Here, we considered data acquired during walking, reaching, flying, or isometric force production. In analyzing the decision-making and motor-control behaviors of various animals, we considered the possibility that the brain may estimate effort objectively, via the metabolic energy consumed to produce the action. We measured the energetic cost of reaching and found that, like walking, it was convex in time, with a global minimum, implying that there existed a movement speed that minimized effort. However, reward made it worthwhile to be energetically inefficient. Using a framework in which utility of an action depended on reward and energetic cost, both discounted in time, we found that it was possible to account for a body of data in which animals were free to choose how to move (reach slow or fast), as well as what to do (walk or fly, produce force F1 or F2). We suggest that some forms of decision-making and motor control may share a common utility in which the brain represents the effort associated with performing an action objectively via its metabolic energy cost and then, like reward, temporally discounts it as a function of movement duration.

RESULTS

Love goes toward love as schoolboys from their books.
But love from love, toward school with heavy looks.
—William Shakespeare

Suppose you are at the airport awaiting arrival of a passenger. As you scan the arrivals, you decide which person is your destina-

tion, and then walk to greet them. In this thought experiment, you will likely walk faster if the passenger is your child, rather than a colleague.

Our concern here is to describe a framework to account for both decision-making (which stimulus to acquire) and motor control (how fast to move). To do so, we need to consider the interaction between reward and effort. Let us consider the possibility that the brain represents effort objectively, via the energetic cost of the action.

We measured the rate of metabolic energy expenditure during reaching \dot{e}_r as a function of distance d and found that as reach duration T increased, \dot{e}_r decreased, approaching a non-zero asymptote (Figure 1A):

$$\dot{e}_r = am + b \frac{md^i}{T^{i-1}}. \quad (\text{Equation 1})$$

In the above equation, m is effective mass of the arm (Supplemental Mathematical Derivations). If we integrate Equation 1 with respect to T , we arrive at an estimate of the energy expended to reach distance d :

$$e_r = amT + b \frac{md^i}{T^{i-1}}. \quad (\text{Equation 2})$$

This finding illustrates that the energetic cost of reaching is convex in time and that there exists a reach duration that minimizes the cost. The energetics of walking a distance d are similar to that of reaching [8]:

$$e_w = amT + \frac{bmd^2}{T}. \quad (\text{Equation 3})$$

Therefore, for both human walking and reaching there exist optimum speeds of movement that minimize the energetic cost. In contrast, an optimal duration does not exist if we assign a cost that depends on the integral of squared forces (Supplemental Mathematical Derivations), because in that case the cost is proportional to m^2/T^3 .

However, minimizing energy expenditure cannot be the only concern of the nervous system (Supplemental Mathematical Derivations), as evidenced by the fact that reward modulates movement speed [1–5]. To incorporate the effect of reward, consider reaching for food that has reward value $\alpha > 0$. Passage of time discounts reward, $\alpha/(1 + \gamma T)$, where γ determines how rapidly reward is discounted [9, 10]. If we view duration of the movement as an implicit delay in the acquisition of reward [11],

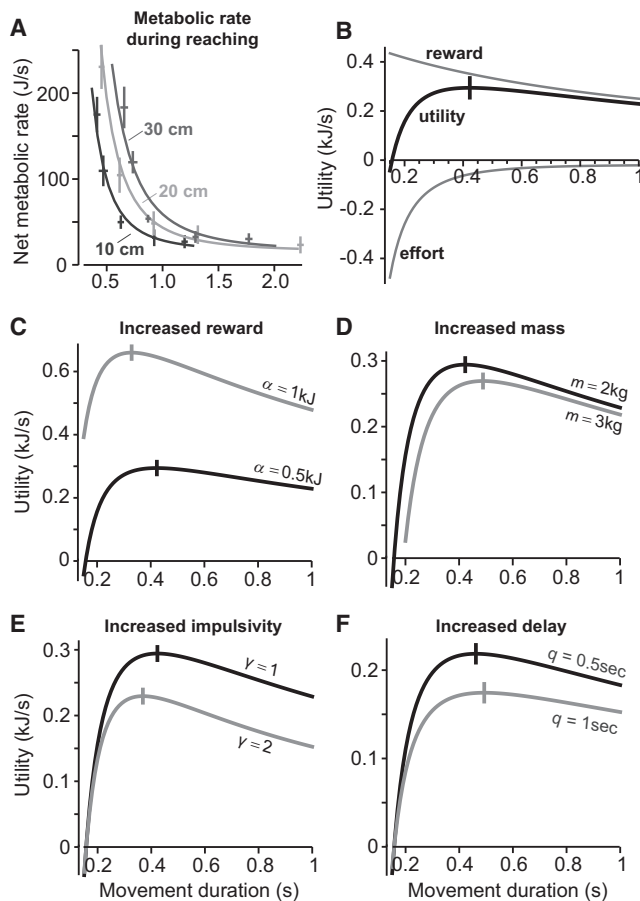


Figure 1. The Effects of Reward, Effort, and Time on Decision-Making and Movement Vigor

(A) Rate of metabolic energy expenditure during reaching as a function of reach duration for different distances. Fit to Equation 1, $R = 0.99$, $a = 15 \pm 4.8$ J/s/kg ($p < 10^{-2}$), $b = 77 \pm 16$ ($p < 10^{-3}$), $i = 1.1 \pm 0.2$ ($p < 10^{-3}$), and $j = 2.7 \pm 0.4$ ($p < 10^{-4}$). Error bars indicate the SEM.

(B) Temporally discounted reward and metabolic cost of movement, plotted as a function of movement duration ($\alpha = 0.5$ kJ, $d = 0.1$ m, $m = 2$ kg, $\gamma = 1$).

(C) With increased reward, the utility of the movement increases and the optimal duration shifts to a smaller value. As a result, a stimulus that promises greater reward carries a greater utility and also produces movements that have greater velocity.

(D) The effort of the movement is increased by increasing the mass of the limb. This decreases the utility, but also shifts the optimal duration, thereby decreasing the velocity of the resulting movement.

(E) Increasing the rate of temporal discounting decreases the utility and shifts the optimal duration to a smaller value, thereby increasing movement velocity.

(F) Increased inter-trial delay decreases the utility of the movement and shifts the optimal duration to a larger value, thereby decreasing movement velocity. Therefore, although increased inter-trial delay and rate of temporal discounting both decrease utility of the movement, the former decreases movement vigor and the latter increases it.

See also Table S1.

then the act of moving fast or slow is a decision between acquisition of a large reward soon in exchange for payment of large effort and acquisition of smaller, discounted reward later in exchange for payment of small effort. We represent effort with

function $U(T)$ and write the utility of the movement as the sum of reward and effort:

$$J = \frac{\alpha}{1 + \gamma T} + U(T). \quad (\text{Equation 4})$$

For reaching, we arrive at the following representation of effort:

$$U(T) = -\frac{amT + bmd^i/T^{i-1}}{1 + \gamma T}. \quad (\text{Equation 5})$$

The resulting utility is plotted in Figure 1B. A fast movement results in small discounting of reward but requires larger effort. A slow movement requires a small effort but will produce large discounting of reward. The maximum utility is achieved at the optimum movement duration. Equation 4 makes four predictions: as reward increases, the utility of the movement increases, but its optimum duration decreases (Figure 1C). Animals should not only prefer stimuli that promise greater reward [12], but also move with greater speed toward them [1, 2, 4, 5]. As the metabolic cost of the movement increases, the utility of the movement decreases, but its duration increases (Figure 1D). Animals should not only prefer to move toward stimuli that require lower energetic cost, but move with greater speed toward them. As time discounts reward more steeply (larger γ), both the utility of the movement and its optimum duration decrease (Figure 1E). Individuals that are more impulsive should not only prefer the immediate reward, but also move faster than individuals who are more patient [13]. Finally, if we make someone wait period q before allowing them to make a movement, we are altering the time to acquisition of reward, extending it to $T + q$. The utility becomes

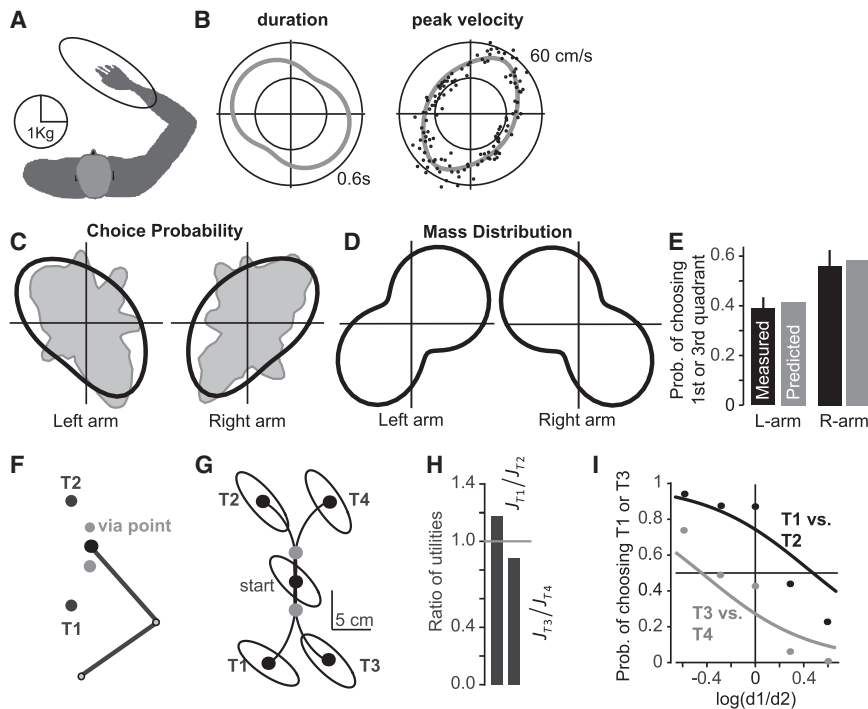
$$J = \frac{\alpha - amT - bmd^i/T^{i-1}}{1 + \gamma(T + q)}. \quad (\text{Equation 6})$$

As the imposed delay increases, movement duration also increases (Figure 1F). As a result, subjects would rather not wait to move toward a rewarding stimulus, but, if forced to do so, they will move slower [14].

We chose a utility in which effort additively combined with reward. In contrast, in previous approaches, reward and effort were combined multiplicatively [15–18]. A utility in which reward is multiplied by a function of effort is inconsistent with the observation that reward modulates movement speed (Supplemental Mathematical Derivations). We chose a utility in which time discounted effort and reward hyperbolically. Exponential temporal discounting fails to predict the effect of time delay on movement speed (Supplemental Mathematical Derivations).

A Common Utility for Some Forms of Decision-Making and Motor Control

Our framework should be able to predict the choice of which movement to perform, as well as the speed of the chosen movement. The arm has a mass distribution that resembles a heavy object when it moves in some directions (Figure 2A, major axis of the ellipse) and a light object when it moves in other directions. In [19], subjects reached without time constraints and chose a peak velocity that was around 55 cm/s for some directions, but only 35 cm/s for other directions (Figure 2B). We used the inertial properties of the arm to estimate its effective mass as a function



(G) The effective mass of the arm at the start position and at each of the various targets.

(H) The ratio of the utilities for targets T1 and T2, and targets T3 and T4, when all targets are 11 cm from the start point.

(I) The probability of choosing target T1 (or target T3) as a function of the log of the ratio of the distances for targets T1 and T2 (or targets T3 and T4). The curves are probabilities computed from the differences in utilities of the two targets.

See also Table S1.

of reach direction (Supplemental Mathematical Derivations), resulting in $m(\theta)$, where θ is direction of reach. Using Equation 4, we computed the duration that maximized the utility for each direction (Figure 2B). The resulting peak velocity was largest for the directions of smallest effective mass (Figure 2B; $R = 0.83$, $p < 10^{-34}$).

Volunteers were asked to reach but were not provided with a target [20, 22]. Rather, they were given the freedom to choose their own direction of movement (Figure 2C). We found that movement utility was a reasonable predictor of the distribution of directions that people had chosen (Figures 2D and 2E; right, $R = 0.67$, $p < 10^{-8}$; left, $R = 0.68$, $p < 10^{-9}$).

Volunteers chose from two targets [21]. A via point was placed between the start position and each target, thereby constraining the reach trajectory (Figure 2F). This made it so that part of the trajectory was aligned with the major or minor axis of the mass ellipse. For example, consider a trial in which the options were targets T1 and T2, each placed at a distance of 11 cm from the start location (Figure 2G). Approaching T2 from the via point requires moving the hand along the major axis of the mass ellipse, whereas approaching T1 requires moving the hand along the minor axis. Because the effective mass of the arm is higher along the trajectory toward T2, its utility is lower. As a consequence, people should prefer to reach toward T1. We used Equation 4 to estimate the utility of each reach. The resulting ratio of utilities for targets T1 and T2, represented as J_{T1}/J_{T2} , and targets T3 and T4, represented as J_{T3}/J_{T4} , are shown in Figure 2H. As the dis-

tance to targets T1 and T3 increased, the preference shifted to targets T2 and T4 (Figure 2I).

In summary, if we define utility of an action as the temporally discounted sum of reward and effort, where effort is the energetic cost of the movement, then we may account for both movement vigor and movement choice across a few reaching experiments.

Temporal Discounting of Effort: Walking versus Flying

Starlings chose whether to walk or fly to obtain reward [23]. The value of reward was the caloric content of the food, and the metabolic costs of flying e_f and walking e_w were measured (flying required about 15 times more energy than walking). The birds chose between making multiple flying trips to receive reward and making multiple walking trips (to a closer site) to receive the same reward. On the final trip, the birds obtain the reward. For a fixed number of flying trips n_f , the number of walking trips n_w was incremented until a preference reversal was observed, indicating an indifference point (Figure 3A). Using Equation 4, we calculated the utility of each option. The utility of making n_w walking trips is

$$J_w = \frac{\alpha - e_w(n_w)}{1 + \gamma T_w(n_w)} \quad (\text{Equation 7})$$

To compute indifference points, we set $J_w(n_w) = J_f(n_f)$ and solve for n_w . The only free parameter is γ . In this experiment,

Figure 2. Effort Representation via Metabolic Cost May Account for Both the Decision of Which Movement to Make and the Speed of the Ensuing Movement

(A) The configuration of the arm at the start position and the associated effective mass $m(\theta)$. In this configuration, the greatest mass is associated with movements to targets aligned with the forearm and is three times the mass in the perpendicular direction.

(B) Subjects were instructed to reach to a target at 10 cm with no time constraints [19]. The resulting peak velocity as a function of direction is plotted as dots. The gray curve shows predictions of the utility function, where effort is objectively measured via metabolic cost of the reach (Figure 1A).

(C) Subjects performed an out-and-back movement but were free to choose the reach direction (no target provided). The gray region shows the probability distribution of the observed choices [20]. The black curve is the prediction.

(D) The mass matrix for an out-and-back movement as a function of reach direction for the left and right arm.

(E) Probability of choosing to reach toward first or third quadrant. Error bars indicate the SD.

(F) Subjects were presented with two targets [21] and chose to reach to one of the targets, moving their hand through a via point.

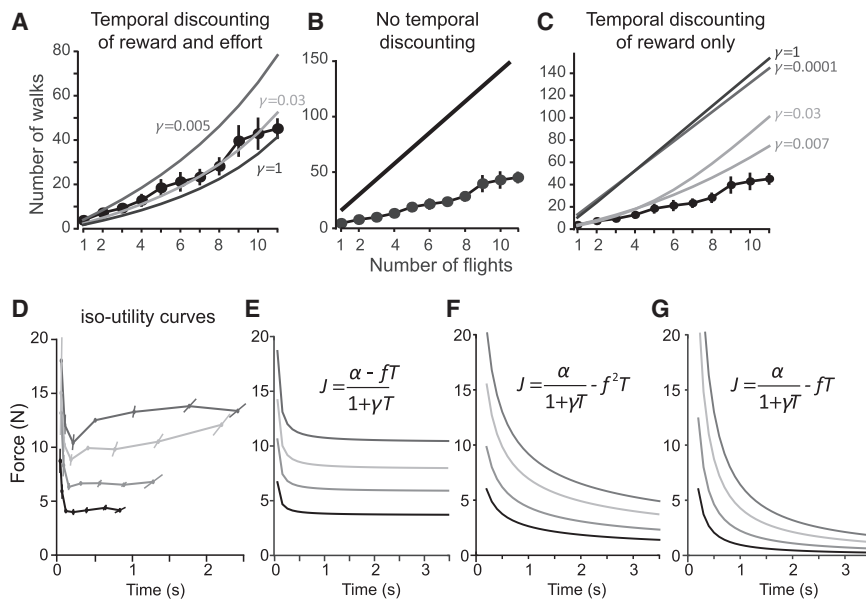


Figure 3. Decision-Making in Birds and Humans

(A) Birds chose between flying a number of times and walking a number of times to receive a constant reward. The data points represent the indifference values [23]. Error bars indicate the SEM. The solid curves are the predicted indifference curve for the utility function in which both reward (energetic content of the food) and metabolic cost are discounted by time. The effect of γ , the only unknown parameter in the model, is a monotonic change in the indifference curve.

(B) Predictions of a utility in which neither reward nor metabolic cost are discounted by time. There are no free parameters.

(C) Predictions of a utility in which only reward is discounted by time. The effect of γ is a non-monotonic change in the indifference curve and the optimal value produces a poor fit to the data.

(D) Iso-utility curves for a force production task [24], where each curve connects the force-time pairs that were judged to be equally effortful. For durations of greater than 0.5 s, force held for a short amount of time was judged to be approximately equal in effort to the same force magnitude

held for a longer amount of time. Error bars indicate the SEM.

(E) The iso-utility curves predicted by an effort cost in which the metabolic cost of force production (force-time integral; see Figure S1) is discounted as a hyperbolic function of time. The function reaches a plateau as the duration of force production increases.

(F) The indifference curves predicted by an effort cost that depends on the integral of the squared force (as is typical in optimal control models). The function goes to zero with increased duration.

(G) The indifference curves predicted by an effort cost that does not temporally discount the force-time integral. The function goes to zero with increased duration. See also Figure S1 and Tables S2 and S3.

where the value of reward and effort were objectively known, the model's performance was an excellent match with the decisions that the birds had made ($R = 0.99$, $p < 0.0001$).

Now consider a scenario in which reward and effort are not temporally discounted, $J_w = \alpha - e_w(n_w)$. This predicts that the birds should take many more walking trips than observed (Figure 3B). Alternatively, consider a scenario in which only reward is discounted,

$$J_w = \frac{\alpha}{1 + \gamma T_w(n_w)} - e_w(n_w). \quad (\text{Equation 8})$$

Here, the effect of γ on the indifference curve is non-monotonic, grossly over-estimating the measured data (Figure 3C).

In summary, when the values of reward and effort were known via their energetic contents, a utility in which effort is represented via temporally discounted movement energetics accounted for the decisions that birds made in flying versus walking.

Temporal Discounting of Effort: Isometric Force Production

The metabolic cost of producing an isometric force $f(t)$ for duration T is (Figure S1)

$$e = a_1 \int_0^T |f(t)| dt + a_2. \quad (\text{Equation 9})$$

When one is asked to produce a constant force F for period T , the utility of this action is

$$J = \frac{\alpha - a_1 FT - a_2}{1 + \gamma T}. \quad (\text{Equation 10})$$

The effort is

$$U = -\frac{1}{1 + \gamma T} (a_1 FT + a_2). \quad (\text{Equation 11})$$

The above result makes the surprising prediction that as duration of force production increases, effort reaches an asymptote. That is, as duration of action increases, subjects become increasingly indifferent to duration.

Volunteers produced F_1 for T_1 , and then F_2 for T_2 [24]. They chose which force-time pair they would like to experience again ("choose the force-time pair that you judge to be less effortful"). By increasing F_2 , the authors determined the indifference point and observed that the iso-utility curves reached a plateau with increased duration (Figure 3D). Our theory provides an explanation. We solved the equality $J(F_1, T_1) = J(F_2, T_2)$ for F_2 , arriving at the following expression for the iso-utility curve:

$$F_2 = \frac{\gamma(\alpha - a_2)(T_1 - T_2) + a_1 F_1 T_1 (1 + \gamma T_2)}{a_1 (1 + \gamma T_1) T_2}. \quad (\text{Equation 12})$$

As $T_2 \rightarrow \infty$, F_2 reaches an asymptote that depends on F_1 (Figure 3E; $R = 0.92$, $p < 10^{-11}$). In contrast, let us consider existing models in which effort is represented as the undiscounted sum of squared forces. In that case,

$$J = \frac{\alpha}{1 + \gamma T} - a F^2 T. \quad (\text{Equation 13})$$

The iso-utility curves become

$$F_2^2 = \frac{F_1^2 T_1}{T_2} - \frac{\alpha}{a T_2 (1 + \gamma T_1)} + \frac{\alpha}{a T_2 (1 + \gamma T_2)}. \quad (\text{Equation 14})$$

As $T_2 \rightarrow \infty$, $F_2 \rightarrow 0$ (Figure 3F), a fact inconsistent with measurements [24]. Indeed, this inconsistency remains whether forces are quadratically penalized or not (Figure 3G).

In summary, if we represent effort as the temporally discounted metabolic cost of action, then the perceived effort associated with generating an isometric force does not grow unbounded as a function of duration, but rather approaches an asymptote. This counter-intuitive prediction provides the first explanation that we are aware of for empirical data regarding decisions that people made in an isometric force task.

DISCUSSION

We measured the energetic cost of reaching and parameterized it as a function of movement duration, mass, and distance. We found that energetics of reaching, like that of walking, is a convex function of time with a global minimum. This demonstrates that for both reaching and walking there are movement speeds that are energetically optimum. Indeed, people and other animals walk and reach at speeds that are near energetic minimums (Supplemental Mathematical Derivations). However, animals move faster when there is greater reward at stake [1, 5]. That is, reward makes it worthwhile to be energetically inefficient.

Earlier work had suggested that the brain may assign a utility for each potential action, where utility contains an interaction between reward and effort [6, 7]. Here, we represented effort objectively via the energy needed to produce the movement and then temporally discounted it like reward. The resulting equations made the following predictions: animals should prefer the more rewarding stimulus and move with greater vigor toward it [5]; animals should prefer the less effortful movement [21] and make that movement with greater vigor [19]; and animals should prefer not to wait before acquiring reward, but, if forced to do so, they will move with reduced vigor [14].

In our framework, effort and reward interact additively and are discounted by time. In decision-making, it is often assumed that effort discounts reward multiplicatively [15, 16, 18]. This is inconsistent with the observation that reward modulates movement vigor. In motor control, effort is represented as the undiscounted sum of squared forces [25–28]. This makes the incorrect prediction that effort grows unbounded with action duration [24]. In contrast, our results connect motor control with optimal foraging theory [29], where decisions depend on the energy gained via the reward, minus the energy spent performing the act, divided by time.

We did not consider decision-making in tasks where evidence is accumulated over time [30]. A recent study demonstrated that the urgency with which a decision is made (i.e., decision speed) affects the vigor of the ensuing movement [31]. We also did not measure the relation between mass and energetic cost of reaching. When people walk with backpack loads, energetic cost increases linearly with mass [32]. Thus, our framework assumes a linear relationship between mass and energetic cost, but this awaits experimental validation.

Our utility cannot account for certain behaviors. When people train to reach in a force field in which a straight trajectory requires more force than a curved trajectory, they choose the straight trajectory [33] (but see [27, 34, 35]). When people walk with an exoskeleton that alters the relationship between step frequency and metabolic cost, they persist on producing their now higher cost natural step frequency, until coached to explore [36]. These examples illustrate instances in which the brain chooses an action that requires greater effort, despite availability of a lower effort option.

Dopamine depletion alters the interaction between reward and effort, shifting preferences toward actions that are less effortful [37, 38]. In the striatum, dopamine interacts with medium spiny neurons (MSNs). MSNs with D1-type receptors project via the direct pathway of the basal ganglia, whereas MSNs with D2-type receptors project via the indirect pathway. Bilateral activation of MSNs in the indirect pathway reduces movement vigor, whereas stimulation of MSNs in the direct pathway increases vigor [39]. This led Hwang [40] to propose that the indirect pathway of the basal ganglia is involved in computing effort costs, whereas the direct pathway is involved in computing the expected reward. In this framework, the utility of an action may be computed via the convergence of the direct and indirect pathways at the substantia nigra pars reticulata (for control of saccades) or the internal segment of globus pallidus (for control of reaching).

Our theory provides a different way to consider experiments in which subjects are free to choose the stimulus to reach for, as well as the arm to use. A current approach asks whether the stimulus is chosen first, or the arm [41]. In contrast, our theory suggests that each potential stimulus/arm pairing is assigned a utility, and then the option is chosen that has the highest utility. Our model suggests a strong coupling between the neural circuits responsible for generating an action and the circuits involved in deciding between actions. This aligns with the hypothesis that decision-making involves sensorimotor areas, where each potential movement is represented simultaneously and competes against other potential movements [42]. The intriguing possibility is that effort is associated with an internal model of the energetic cost of action and, like reward, is discounted in time.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, mathematical derivations, one figure, and three tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cub.2016.05.065>.

AUTHOR CONTRIBUTIONS

H.J.H. and A.A.A. designed, collected, and analyzed the data for the metabolics of reaching. R.S. and A.A.A. developed the mathematical framework, performed simulations, prepared figures, and wrote the manuscript. All authors approved the final version of the manuscript.

ACKNOWLEDGMENTS

This work was supported by grants from the NIH (NS078311), the ONR (N00014-15-1-2312), and the NSF (SES1230933 and SES1352632). We thank Peter Dayan, Mickey Goldberg, Konrad Kording, and Rodger Kram for insightful comments, as well as use of metabolic equipment (Rodger Kram). The University of Colorado Institutional Review Board approved the metabolics of reaching experimental protocol, and all subjects gave written, informed consent.

Received: March 1, 2016

Revised: April 19, 2016

Accepted: May 26, 2016

Published: June 30, 2016

REFERENCES

- Kawagoe, R., Takikawa, Y., and Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* 1, 411–416.
- Xu-Wilson, M., Zee, D.S., and Shadmehr, R. (2009). The intrinsic value of visual information affects saccade velocities. *Exp. Brain Res.* 196, 475–481.
- Opris, I., Lebedev, M., and Nelson, R.J. (2011). Motor planning under unpredictable reward: modulations of movement vigor and primate striatum activity. *Front. Neurosci.* 5, 61.
- Sackaloo, K., Strouse, E., and Rice, M.S. (2015). Degree of preference and its influence on motor control when reaching for most preferred, neutrally preferred, and least preferred candy. *OTJR (Thorofare N J)* 35, 81–88.
- Reppert, T.R., Lempert, K.M., Glimcher, P.W., and Shadmehr, R. (2015). Modulation of saccade vigor during value-based decision making. *J. Neurosci.* 35, 15369–15378.
- Niv, Y., Daw, N.D., Joel, D., and Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl.)* 191, 507–520.
- Rigoux, L., and Guigon, E. (2012). A model of reward- and effort-based optimal decision making and motor control. *PLoS Comput. Biol.* 8, e1002716.
- Ralston, H.J. (1958). Energy-speed relation and optimal speed during level walking. *Int. Z. Angew. Physiol.* 17, 277–283.
- Jimura, K., Myerson, J., Hilgard, J., Braver, T.S., and Green, L. (2009). Are people really more patient than other animals? Evidence from human discounting of real liquid rewards. *Psychon. Bull. Rev.* 16, 1071–1075.
- Kobayashi, S., and Schultz, W. (2008). Influence of reward delays on responses of dopamine neurons. *J. Neurosci.* 28, 7837–7846.
- Shadmehr, R., Orban de Xivry, J.J., Xu-Wilson, M., and Shih, T.Y. (2010). Temporal discounting of reward and the cost of time in motor control. *J. Neurosci.* 30, 10507–10516.
- Walton, M.E., Kennerley, S.W., Bannerman, D.M., Phillips, P.E., and Rushworth, M.F. (2006). Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making. *Neural Netw.* 19, 1302–1314.
- Choi, J.E., Vaswani, P.A., and Shadmehr, R. (2014). Vigor of movements and the cost of time in decision making. *J. Neurosci.* 34, 1212–1223.
- Haith, A.M., Reppert, T.R., and Shadmehr, R. (2012). Evidence for hyperbolic temporal discounting of reward in control of movements. *J. Neurosci.* 32, 11727–11736.
- Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M.L., and Dreher, J.C. (2010). Separate valuation subsystems for delay and effort decision costs. *J. Neurosci.* 30, 14080–14090.
- Sugiwaka, H., and Okouchi, H. (2004). Reformatory self-control and discounting of reward value by delay or effort. *Jpn. Psychol. Res.* 46, 1–9.
- Pasquereau, B., and Turner, R.S. (2013). Limited encoding of effort by dopamine neurons in a cost-benefit trade-off task. *J. Neurosci.* 33, 8288–8300.
- Klein-Flügge, M.C., Kennerley, S.W., Saraiva, A.C., Penny, W.D., and Bestmann, S. (2015). Behavioral modeling of human choices reveals dissociable effects of physical effort and temporal delay on reward devaluation. *PLoS Comput. Biol.* 11, e1004116.
- Gordon, J., Ghilardi, M.F., Cooper, S.E., and Ghez, C. (1994). Accuracy of planar reaching movements. II. Systematic extent errors resulting from inertial anisotropy. *Exp. Brain Res.* 99, 112–130.
- Wang, W., and Dounskaia, N. (2012). Load emphasizes muscle effort minimization during selection of arm movement direction. *J. Neuroeng. Rehabil.* 9, 70.
- Cos, I., Bélanger, N., and Cisek, P. (2011). The influence of predicted arm biomechanics on decision making. *J. Neurophysiol.* 105, 3022–3033.
- Goble, J.A., Zhang, Y., Shimansky, Y., Sharma, S., and Dounskaia, N.V. (2007). Directional biases reveal utilization of arm's biomechanical properties for optimization of motor behavior. *J. Neurophysiol.* 98, 1240–1252.
- Bautista, L.M., Tinbergen, J., and Kacelnik, A. (2001). To walk or to fly? How birds choose among foraging modes. *Proc. Natl. Acad. Sci. USA* 98, 1089–1094.
- Körding, K.P., Fukunaga, I., Howard, I.S., Ingram, J.N., and Wolpert, D.M. (2004). A neuroeconomics approach to inferring utility functions in sensorimotor control. *PLoS Biol.* 2, e330.
- Todorov, E., and Jordan, M.I. (2002). Optimal feedback control as a theory of motor coordination. *Nat. Neurosci.* 5, 1226–1235.
- O'Sullivan, I., Burdet, E., and Diedrichsen, J. (2009). Dissociating variability and effort as determinants of coordination. *PLoS Comput. Biol.* 5, e1000345.
- Izawa, J., Rane, T., Donchin, O., and Shadmehr, R. (2008). Motor adaptation as a process of reoptimization. *J. Neurosci.* 28, 2883–2891.
- Salimpour, Y., and Shadmehr, R. (2014). Motor costs and the coordination of the two arms. *J. Neurosci.* 34, 1806–1818.
- Emlen, J.M. (1966). The role of time and energy in food preference. *Am. Nat.* 100, 611–617.
- Gold, J.I., and Shadlen, M.N. (2002). Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* 36, 299–308.
- Thura, D., Cos, I., Trung, J., and Cisek, P. (2014). Context-dependent urgency influences speed-accuracy trade-offs in decision-making and movement execution. *J. Neurosci.* 34, 16442–16454.
- Bastien, G.J., Willems, P.A., Schepens, B., and Heglund, N.C. (2005). Effect of load and speed on the energetic cost of human walking. *Eur. J. Appl. Physiol.* 94, 76–83.
- Kistemaker, D.A., Wong, J.D., and Gribble, P.L. (2010). The central nervous system does not minimize energy cost in arm movements. *J. Neurophysiol.* 104, 2985–2994.
- Huang, H.J., Kram, R., and Ahmed, A.A. (2012). Reduction of metabolic cost during motor learning of arm reaching dynamics. *J. Neurosci.* 32, 2182–2190.
- Farshchiansadegh, A., Melendez-Calderon, A., Ranganathan, R., Murphey, T.D., and Mussa-Ivaldi, F.A. (2016). Sensory agreement guides kinetic energy optimization of arm movements during object manipulation. *PLoS. Comput. Biol.* 12, e1004861.
- Selinger, J.C., O'Connor, S.M., Wong, J.D., and Donelan, J.M. (2015). Humans Can Continuously Optimize Energetic Cost during Walking. *Curr. Biol.* 25, 2452–2456.
- Salamone, J.D., Steinpreis, R.E., McCullough, L.D., Smith, P., Grebel, D., and Mahan, K. (1991). Haloperidol and nucleus accumbens dopamine depletion suppress lever pressing for food but increase free food consumption in a novel food choice procedure. *Psychopharmacology (Berl.)* 104, 515–521.
- Correa, M., Carlson, B.B., Wisniecki, A., and Salamone, J.D. (2002). Nucleus accumbens dopamine and work requirements on interval schedules. *Behav. Brain Res.* 137, 179–187.
- Kravitz, A.V., Freeze, B.S., Parker, P.R., Kay, K., Thwin, M.T., Deisseroth, K., and Kreitzer, A.C. (2010). Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature* 466, 622–626.
- Hwang, E.J. (2013). The basal ganglia, the ideal machinery for the cost-benefit analysis of action plans. *Front. Neural Circuits* 7, 121.
- Herbort, O., and Rosenbaum, D.A. (2014). What is chosen first, the hand used for reaching or the target that is reached? *Psychon. Bull. Rev.* 21, 170–177.
- Cisek, P., and Pastor-Bernier, A. (2014). On the challenges and mechanisms of embodied decisions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369, 20130479.

Current Biology, Volume 26

Supplemental Information

**A Representation of Effort
in Decision-Making and Motor Control**

Reza Shadmehr, Helen J. Huang, and Alaa A. Ahmed

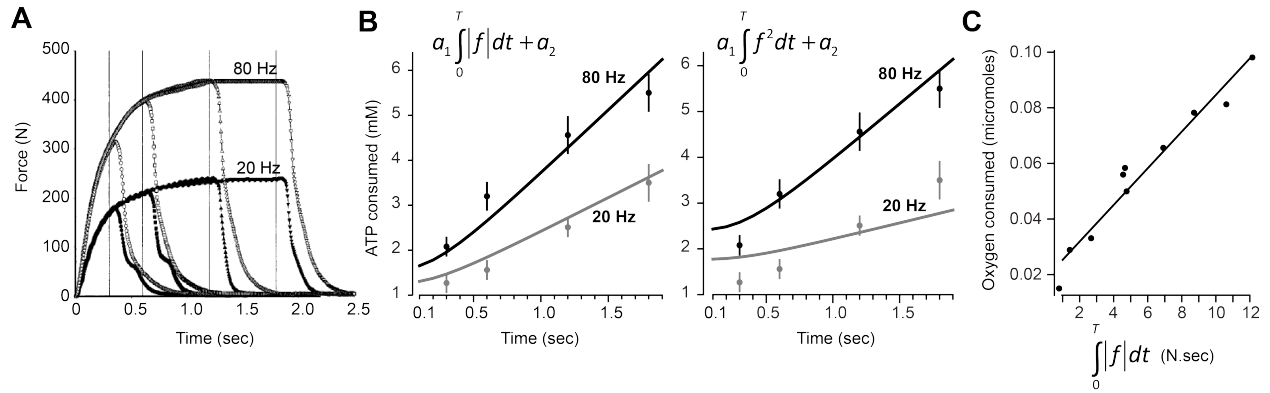


Figure S1. Related to Figure 3D-G. Energy consumed during isometric contraction grows linearly with the force-time integral. **A.** The measured force during electrical stimulation of the human gastrocnemius muscle at 20 or 80Hz (data from [S1]). The 20Hz stimulation produced a force that increased with a time constant of around 0.25sec, reaching a plateau of approximately 230N. The 80Hz stimulation produced a force that had a similar time constant, but reached a plateau of approximately 430N. **B.** Energy expended by the muscle as measured via consumption of ATP. The data points are from [S1]. Error bars are SEM. In the left subplot, the curves depict a model where energetic costs grow linearly with the force-time integral. In the right subplot, the curves depict a model where energetic costs grow with the squared force-time integral. The fit with the force-time integral has about half of the errors as in the squared force-time integral. **C.** Oxygen consumed during electrical stimulation of a frog muscle plotted as a function of the force-time integral. Data from [S2].

Supplemental Mathematical Derivations

Speed of walking

Considerable research has approached the question of natural walking speed of various animals. Here, we briefly summarize this work, demonstrating that like reaching, energetics of human walking are convex in time, suggesting an optimum speed. However, this speed is slower than one that humans choose during locomotion. We suggest that the reason for this is that reward interacts with effort, making it worthwhile to be energetically inefficient.

The current view in walking research is that the energetic cost of locomotion defines the optimal speed of motion [S3]. For example, Ralston [S4] measured rate of energy expenditure in human walking and found that when people were not walking, the rate of energy expenditure per unit mass was $\dot{e}_0 = a_0$, and when they were walking $\dot{e}_w = \dot{e}_0 + a_1 + bv^2$ per unit mass, where v is average speed of walking. Suppose that during an arbitrary period of time T a person spends some time $T_w < T$ walking a distance d , and is otherwise not walking. During walking their average speed is d/T_w . The rate of energy spent during walking (per unit mass) becomes:

$$\dot{e}_w = a_0 + a_1 + \frac{bd^2}{T_w^2} \quad (S1)$$

If that individual has mass m , total energy spent during that T period is:

$$\begin{aligned} e_T &= a_0(T - T_w)m + (a_0 + a_1)mT_w + \frac{bmd^2}{T_w} \\ &= Ta_0m + a_1mT_w + \frac{bmd^2}{T_w} \end{aligned} \quad (S2)$$

To find the speed of walking that minimizes energy expenditure, we compute the derivative of the above expression with respect to T_w , and find the optimum walk duration T_w^* :

$$T_w^* = d \frac{\sqrt{b}}{\sqrt{a_1}} \quad (S3)$$

The optimum walking speed is:

$$v^* = \frac{d}{T_w^*} = \frac{\sqrt{a_1}}{\sqrt{b}} \quad (S4)$$

Therefore, the energetic cost of walking, similar to that of reaching, describes a convex function of time which exhibits a minimum. This means that in this energetics-only formulation of movement, there is an optimal speed and duration that minimizes the energetic cost of walking. Indeed, Hoyt and Taylor [S5] concluded that humans and other animals “select speed within a gait in a manner that minimizes energy consumption”.

However, this energetics-only framework has a number of limitations. As Srinivasan [S3] has pointed out, the optimum speed predicted by Eq. (S4) is considerably slower than the speed that people naturally choose to walk. The reason for this, our theory suggests, is that reward discounts effort, making it worthwhile to expend energy to acquire a rewarding goal (Figure 1). One caveat is that the optimum walking speed depends on how energetic cost is represented. A common approach in the locomotion literature [S4] is to represent it as the total energetic cost to walk a unit distance, referred to as the cost of transport. The optimum speed predicted by minimizing the total cost of transport provides a reasonable match to the natural speeds observed. However, an argument against optimization of total cost of transport is that it relies on the unrealistic assumption that one will always be walking and nothing else.

Ultimately, regardless how energetic cost is calculated, an energetics-only framework does not consider a role for reward. As a result, this framework cannot account for the observation that people and other animals move faster toward stimuli that promise greater reward [S6-S10]. It cannot account for the observation that more impulsive people may move faster than less impulsive people [S11]. It cannot account for the observation that movement speed declines with increasing inter-trial intervals [S12]. It cannot account for the observation that people in different cities walk at different speeds [S13].

Because we found that energetics of walking and reaching were both convex functions of time, our results regarding effects of reward, mass, distance, etc. on vigor of reaching predict similar effects on vigor of walking.

In summary, our framework extends the existing approach in the field of walking research in a critical way: using energetics of action as a measure of effort, we define a utility in which there is interplay between reward, effort, and time. We arrive at a theory that not only describes where one should walk (toward the choice that offers the greatest utility), but also how fast one should walk (via a speed that maximizes the utility). As a result, the theory provides an account of the observation that when there is greater reward at stake, one should walk faster. When time is more valuable, one should walk faster. If mass increases, one should walk slower.

Combining effort and reward additively vs. multiplicatively

We suggested that effort discounted reward additively. This is in contrast to earlier work where it was assumed that effort discounts reward multiplicatively [S14-S16]. Let us compare these two approaches. In the case of multiplicative interaction between reward and effort, we have the following representation for utility:

$$J = \frac{\alpha}{1 + \gamma T} U(T) \quad (S5)$$

In the above formulation increasing reward increases the utility of the action, but has no effect on the optimal movement duration (because the effect of reward is to scale the utility function, which has no effect on the value of time that maximizes the utility). Therefore, such a formulation is inconsistent with the observation that reward makes movements faster. That is, any utility in which reward is multiplied by a function of effort will fail to predict sensitivity of movement vigor to reward. This is in contrast to experimental evidence [S10]. In contrast, an additive interaction between reward and effort correctly predicts that increased reward increases movement vigor.

Hyperbolic vs. exponential temporal discounting

We suggested that like reward, effort is discounted hyperbolically. In reinforcement learning [S17], future rewards are discounted exponentially, largely due to mathematical convenience. Exponential discounting has also been suggested in models of human decision making [S18]. Hyperbolic discounting, however, is more consistent with decision-making data in humans [S19] and monkeys [S20]. In our expression for movement utility, we opted to discount reward and effort as a hyperbolic function of time. The reason for this is that exponential discounting makes the incorrect prediction that changes in inter-trial intervals should have no effect on movement vigor. To illustrate this, suppose that time discounts utility exponentially. For an arbitrary effort $e(T)$, we have:

$$J = (\alpha + e(T)) \exp\left(-\frac{T+q}{\tau}\right) \quad (S6)$$

If we take the derivative of the above equation with respect to T , we find the following expression:

$$\frac{dJ}{dT} = \left[\frac{de(T)}{dT} - \frac{\alpha + e(T)}{\tau} \right] \exp\left(-\frac{T+q}{\tau}\right) \quad (S7)$$

To find the optimum duration, we set the above expression equal to zero, and find that the inter-trial delay q has no consequence on the optimum duration. Thus, in the case of exponential discounting we find that vigor is independent of inter-trial delay, something that is inconsistent with experimental evidence [S12].

We also considered the possibility that reward and effort were discounted exponentially in time in our analysis of the choices that birds made in walking vs. flying, and the choices people made in isometric force production. For the walking vs. flying data set, we found that exponential discounting of reward and effort provided as good a fit to the data as hyperbolic discounting ($R=0.99$, $p<10^{-8}$). For isometric force production (Figure 3), we found that exponential discounting did a better job than no discounting at all, but still underperformed hyperbolic discounting ($R=0.84$, $p<10^{-7}$).

In Figure 3D, the indifference curves dip downward, and then move slightly upward with increasing durations. This is unexpected because it suggests that the utility of producing a force for a given duration is equal to producing a larger force for a larger duration. Intriguingly, exponential discounting can account for the curious dip observed in the indifference curves, although the overall fit to the data was still worse than hyperbolic discounting. The dip could also be explained by a utility in which time was discounted hyperbolically, where discounting of time is due to a power function of time:

$$U = \frac{\alpha - a_1 FT - a_2}{1 + T^\gamma} \quad (S8)$$

While the above utility could account for the curious dip, its overall fit to the data ($R=0.65$, $p<10^{-3}$) was not as good as the utility where effort is hyperbolically discounted by time.

Integral of squared motor commands vs. energetic cost

Assuming that temporal discounting of effort is hyperbolic, is there a fundamental difference between representing effort via energetic cost as compared to sum of squared forces? We will consider this question in this section and show that sum of squared forces presents a cost that is convex in time but has no global minimum. In contrast, energetics of both walking and reaching are convex in time but also have a global minimum. The implications of this result are presented.

Let us describe a utility that represents effort via integral of squared forces. How does this utility depend on duration of movement? Suppose a movement is made in period T , along trajectory $q(t)$. We compute the integral of the squared forces for that movement, represented by s . Now suppose that the system makes the movement in period T/a . What is integral of the squared forces produced in this movement? We note that:

$$\begin{aligned}\bar{q}(t) &= q(r(t)) \\ r(t) &= at\end{aligned}\tag{S9}$$

The velocity of the movement $\bar{q}(t)$ is related to the movement $q(t)$ by the following:

$$\frac{d\bar{q}(t)}{dt} = \frac{dq(r)}{dt} = \frac{dq(r)}{dr} \dot{r} = q'(r) \dot{r} = a q'(r)\tag{S10}$$

For acceleration we have:

$$\begin{aligned}\frac{d^2\bar{q}(t)}{dt^2} &= \frac{d(q'(r)\dot{r})}{dt} = \frac{dq'(r)}{dt} \dot{r} + q'(t)\ddot{r} = \frac{dq'(r)}{dr} \dot{r}^2 + q'(t)\ddot{r} = q''(r)\dot{r}^2 + q'(t)\ddot{r} \\ &= a^2 q''(r)\end{aligned}\tag{S11}$$

For a typical system that has dynamics described by inertial and coriolis/centripetal forces, we have the following relationship between forces and motion:

$$\tau(t) = I(\mathbf{q}(t))\ddot{\mathbf{q}}(t) + C(\mathbf{q}(t))\dot{\mathbf{q}}^2(t)\tag{S12}$$

For the sped-up trajectory, we have:

$$\begin{aligned}\bar{\tau}(t) &= I(\bar{\mathbf{q}}(t))\frac{d^2\bar{\mathbf{q}}(t)}{dt^2} + C(\bar{\mathbf{q}}(t))\left(\frac{d\bar{\mathbf{q}}(t)}{dt}\right)^2 \\ &= I(\mathbf{q}(r))a^2\mathbf{q}''(r) + C(\mathbf{q}(r))a^2\mathbf{q}'(r) \\ &= a^2\tau(r)\end{aligned}\tag{S13}$$

The above result implies that when we scale time to go from one trajectory that takes duration T to another that takes duration T/a , the forces scale by a^2 . Let us now compare the integral of the squared forces in the two trajectories.

We indicate this integral with the symbol $i(T)$ for one trajectory, and $i(T/a)$ for the other trajectory.

$$\begin{aligned}i(T) &= \int_0^T \tau^2(t) dt = s \\ i(T/a) &= \int_0^{T/a} \bar{\tau}^2(t) dt = \int_0^T a^4 \tau^2(r) \frac{dr}{a} = a^3 s\end{aligned}\tag{S14}$$

These two equations imply that there exists a function $i(T)$ such that:

$$\frac{i(T)}{i(T/a)} = \frac{1}{a^3}\tag{S15}$$

A function that satisfies the above condition is:

$$i(T) = \frac{c}{T^3}\tag{S16}$$

This result implies that the sum of squared forces required to move an inertial system along an arbitrary trajectory will decreased as the third power of the duration of movement. These forces scale approximately linearly with mass. As a result, we conclude that if our measure of effort is sum of squared forces, then the utility of a constant amplitude movement as a function of duration T and mass m is:

$$J = \frac{\alpha - cm^2/T^3}{1 + \gamma T}\tag{S17}$$

In comparison, if our measure of effort is the energetic cost of that movement, then the utility is:

$$J = \frac{\alpha - amT - bm/T^2}{1 + \gamma T} \quad (\text{S18})$$

We find that a fundamental difference between representations of effort via sum of squared forces vs. energetic cost is that the latter is a convex function of time with a minimum, whereas the former is not. The minimum arises from the bias term, am , in Eq. (1) that represents an energetic cost that grows with time, penalizing movements with longer durations. The bias term is not present if effort is sum of squared forces or even sum of forces. It also cannot be explained by the cost of supporting the arm against gravity, as the arm was supported throughout the movement. Interestingly, a similar bias is present in walking, and also leads to energetics that are convex in time with a minimum [S4].

What is the implication of a movement duration that minimizes energetic cost? Consider a thought experiment in which one is given the option of doing nothing (and receiving no reward), vs. performing a movement for $T = 10$ seconds to receive reward α . If we set α large enough, then both utilities produce a positive value, which implies that one should choose to perform the movement and receive the reward. Now suppose that we increase the required movement duration. If the utility of these choices is based on squared forces, as $T \rightarrow \infty$, $J \rightarrow 0$, but always remains positive. This means that according to the squared forces model, if we accept to perform a short duration movement for some reward, we must also accept it despite the requirement of extremely long movement durations. This unreasonable prediction arises because representation of effort via sum of squared forces does not exhibit a minimum as a function of time.

However, if the utility of the choice is based on energetic cost, then as $T \rightarrow \infty$, then $J \rightarrow -\frac{am}{1 + \gamma T}$. This means that regardless of reward, the utility of doing an action for a very long period is always less than zero. As a result, while we will accept to perform the movement for a short period of time in exchange for the reward, as the movement duration becomes longer, we will reject the offer of reward and opt to do nothing. This reasonable prediction arises because the energetic cost has a global minimum.

Effective mass of the hand as a function of movement direction

In Figure 2 we considered data from experiments in which people reached to various directions. To compute the utility of each movement, we needed to estimate the effective mass of the arm as it moved in each direction. Here, we show how to estimate this effective mass.

We begin with the inertia of the arm, which for the planar configuration is a 2x2 position-dependent matrix $I(\theta)$, where $\theta = \begin{bmatrix} \theta_s & \theta_e \end{bmatrix}^T$, representing angular position of the shoulder and elbow joints. At rest, inertia represents the relationship between a vector of joint accelerations and the resulting torque:

$$\tau = I(\theta)\ddot{\theta} \quad (\text{S19})$$

We are interested in computing the mass matrix $M(\theta)$ which represents the relationship between the acceleration vector \ddot{x} and the force vector as measured at the hand at rest:

$$f = M(\theta)\ddot{x} \quad (\text{S20})$$

We use the Jacobian matrix:

$$\Lambda = \frac{dx}{d\theta} \quad (\text{S21})$$

and the principle of virtual work to relate force to torque, and acceleration in joint coordinates to hand coordinates:

$$\begin{aligned} \tau &= \Lambda^T f \\ \dot{x} &= \Lambda \dot{\theta} \\ \ddot{x} &= \frac{d\Lambda}{d\theta} \dot{\theta} \dot{\theta} + \Lambda \ddot{\theta} \end{aligned} \quad (\text{S22})$$

Using the above equalities, we can write the relationship between hand acceleration and force:

$$f = \left(\Lambda^{-1}\right)^T I(\theta) \Lambda^{-1} \ddot{x} - \left(\Lambda^{-1}\right)^T I(\theta) \Lambda^{-1} \frac{d\Lambda}{d\theta} \dot{\theta} \dot{\theta} \quad (\text{S23})$$

When the hand is at rest, the velocity-dependent term on the right side of the equation is zero, providing the following relationship between acceleration and force:

$$f = (\Lambda^{-1})^T I(\theta) \Lambda^{-1} \ddot{x} \quad (\text{S24})$$

As a result, we can define the mass matrix at the hand as:

$$M(\theta) = (\Lambda^{-1})^T I(\theta) \Lambda^{-1} \quad (\text{S25})$$

In the case of the planar arm that we are considering, the mass matrix $M(\theta)$ is a 2x2, describing the relationship between accelerations and forces at the hand. To compute the effective mass $m(\theta)$ we apply an acceleration of length unity in a given direction and compute the length of the resulting force vector.

Energetic cost of isometric force production

We considered data from an experiment in which energy consumption was estimated via the change in ATP concentration (or alternatively, oxygen uptake) during the electrical stimulation of an isometric muscle (Figure S1A). We modeled the actual forces produced by the muscle, computed its integral, and then fitted the following equation to the measured data, with the results shown in Figure S1B (left panel):

$$e = a_1 \int_0^T |f(t)| dt + a_2 \quad (\text{S26})$$

We found that this model produced a mean-squared error that was about half as large as those produced if we had assumed that energy was related to the integral of the squared force:

$$e(f(t)) = a_1 \int_0^T f^2(t) dt + a_2 \quad (\text{S27})$$

The results are shown in the right subplot of Figure S1B. Other experimental data [S2] provided further evidence for the conjecture that the energetic cost associated with isometric force production is a linear function of the force-time integral (Figure S1C).

Let us consider the iso-utility curves in Figure 3D and ask what would happen if effort was the energetic cost of force production, but not discounted in time. This means that the utility of producing a constant force F for duration T is:

$$J = \frac{\alpha}{1 + \gamma T} - a_1 F T - a_2 \quad (\text{S28})$$

To find iso-utility curves, we solve the equality $J(F_1, T_1) = J(F_2, T_2)$ and obtain the following expression:

$$F_2 = \frac{a_1 F_1 T_1}{a_1 T_2} - \frac{\alpha}{a_1 T_2 (1 - \gamma T_1)} + \frac{\alpha}{a_1 T_2 (1 - \gamma T_2)} \quad (\text{S29})$$

This equation tells us that if effort is not temporally discounted, then the iso-utility curves will asymptote to zero with increasing duration T_2 . Therefore, undiscounted energetic cost, like the time-integral of force (Figure 3G), will lead to iso-utility curves that go to zero, which is inconsistent with the experimental data.

Supplemental Tables

Table S1. Simulation Parameters and Goodness of Fit for Figures 1 and 2

		Parameters										Goodness of Fit
Simulation	Fig	a (J/s/kg)	b	i	j	α (kJ)	γ	m (kg)	d (m)	θ (sec)	k	
Movement Utility	1B	15	77	1.1	3	0.5	1	2	0.1	0	n/a	n/a
Effect of reward	1C	15	77	1.1	3	0.5 1.0	1	2	0.1	0	n/a	n/a
Effect of mass	1D	15	77	1.1	3	0.5	1	2 3	0.1	0	n/a	n/a
Effect of impulsivity	1E	15	77	1.1	3	0.5	1 2	2	0.1	0	n/a	n/a
Effect of inter-trial delay	1F	15	77	1.1	3	0.5	1	2	0.1	0.5 1	n/a	n/a
Effect of mass on movement duration (Gordon et al. 1994)	2B	15	77	1.1	3	0.31	1	$m(\theta)$	0.1	0	n/a	$R = 0.83, p < 10^{-34}$
Effect of mass on movement direction (Wang and Dounskaia 2012)	2C	15	77	1.1	3	0.1	1	$m(\theta)$	0.15	0	n/a	Right: $R=0.67, p < 10^{-8}$ Left: $R=0.68, p < 10^{-9}$
Effect of mass on movement direction (Cos et al. 2011)	2H 2I	15	77	1.1	3	0.1	1	$m(\theta)$	0.11 0.075-0.145	0	28	T1/T2: $R=0.96, p=0.009$ T3/T4: $R=0.96, p=0.008$

Table S2. Simulation Parameters and Goodness of Fit for Figure 3A-3C

		Parameters								Goodness of Fit
Simulation	Fig	\dot{e}_w (J/s)	\dot{e}_f (J/s)	\dot{e}_p (J/s)	t_w (sec)	t_f (sec)	t_p (sec)	α (kJ)	γ	
Walk/Run Indifference Points: Temporal discounting of reward and effort	3A	2 J/s	31.7	1.09	0.6	1.1	1.25	1.3	0.03	R=0.99, p<0.0001, SSE=74.35
Walk/Run Indifference Points: No temporal discounting	3B	2 J/s	31.7	1.09	0.6	1.1	1.25	1.3	0	R=0.99, p<0.0001, SSE=757.84
Walk/Run Indifference Points: Temporal discounting of reward only	3C	2 J/s	31.7	1.09	0.6	1.1	1.25	1.3	0.007	R=0.99, p<0.0001, SSE=109.28

Table S3. Simulation Parameters and Goodness of Fit for Figure 3E-3G

		Parameters				Goodness of Fit
Simulation	Fig	α	a_1	a_2	γ	
Temporal discounting of reward and effort: Force-Time Integral	3E	1	1	1	25	R=0.92, p<10 ⁻¹¹
No temporal discounting: Force ² -Time Integral	3F	1	1	n/a	n/a	R=0.67, p<10 ⁻⁴
No temporal discounting: Force-Time Integral	3G	1	1	n/a	0.001	R=0.62, p<10 ⁻³

Supplemental Experimental Procedures

EXPERIMENT: Energetic cost of reaching

We measured rate of metabolic energy expenditure via expired gas analysis as subjects made reaching movement of different distances and durations. Fifteen healthy adults (mean \pm SD age 24.2 ± 4.4 yrs, 8 females, 7 males) participated in this study. All subjects were right-handed and recruited from the University of Colorado Boulder student body. The University of Colorado Institutional Review Board approved the study protocol and all subjects gave informed consent.

Seated subjects grasped the handle of a robotic arm (Interactive Motion Technologies, Shoulder-Elbow Robot 2) to move a circular cursor from a home circle to a target circle at five pre-determined reaching speeds. The cursor, home circle, and target circle were displayed on a vertically mounted computer screen at the subject's eye-level. The five speeds are referred to as Very Slow, Slow, Medium, Fast, and Very Fast. We tested three reach distances of 10, 20, and 30cm. On odd numbered trials, reaches started ~ 15 cm in front of the chest area with the arm in a flexed position. On even numbered trials, the reach started at the previous target location with the arm in an extended position and involved flexing the arm back towards the center target. Subjects wore bilateral shoulder straps and a lap belt to limit torso movement. A cradle attached to the robot handle supported the right forearm against gravity and restricted wrist movement.

A training bar that moved with a velocity that corresponded to the minimum jerk trajectory was used to illustrate the desired reaching speed during a familiarization period at the beginning of each reaching block. Additionally, the target turned gray if the reach was too slow, green if the reach was too fast, and "exploded" as a flashing yellow ring if the reach was within ± 50 ms of the desired movement duration. A pleasant auditory tone was also used to signal that the subjects successfully hit the target within the desired time window. After reaching the target, subjects had 800ms to settle in the center ring of the home circle before the next target circle was displayed. Thus, the inter-trial time was fixed at 800ms for all speeds at each reach distance.

We measured metabolic rate using expired gas analysis (ParvoMedics, TrueOne 2400). Subjects wore a nose clip and breathed in and out of a mouthpiece during data collection. We sampled the rates of oxygen consumption (\dot{V}_{O_2}) and carbon dioxide production (\dot{V}_{CO_2}) at 5 second intervals as subjects made reaching movements at the desired speeds. Data collections occurred early in the morning, after subjects had fasted overnight. We calibrated the system prior to each data collection using certified gas mixtures and with a range of flow rates using a 3 liter calibration syringe. All metabolic data were corrected with standard temperature and pressure, dry (STPD).

Subjects performed five 5-minute reaching blocks at each of the five fixed speeds. The speeds for these five reaching blocks were randomized for each subject. Each 5-minute reaching block began with 20 practice trials during which no energetic data was recorded. After the practice trials, subjects placed the clip on their nose, inserted the mouthpiece, and breathed for ~ 1 minute while sitting quietly. After this 1-minute breathing period, subjects performed N number of reaches, where N was chosen to last ~ 5 minutes. Thus, all subjects performed the same number of reaches for a given reaching speed and reach distance. In between reaching blocks, subjects rested for at least 5 minutes during which no energetic data was recorded. If subjects were naïve to reaching with the robotic arm, we asked them to come in for a brief ~ 15 minute familiarization session the day prior to the data collection. The familiarization session involved short reaching blocks of 50 trials at relatively slow and fast speeds until the subject appeared to be comfortable with the robotic arm and the task. Only the trials performed during the last 3 minutes of each reaching block, corresponding to the steady-state metabolic data, were used in the calculations.

We only analyzed metabolic data with respiratory exchange ratio, $RER = \dot{V}_{CO_2} / \dot{V}_{O_2}$, values less than 1.0 and generally below 0.85, suggesting that oxidative metabolism was primarily involved [S21]. Normal resting RER values range from 0.74 to 0.87, partly depending on diet and other factors [S22, S23].

We calculated the average gross metabolic rate to perform the task in terms of Joules per second using the Brockway equation [S24]. As we were interested only in the cost of moving the arm, we calculated the net metabolic rate by subtracting the bias representing the cost of sitting quietly with the hand resting in the cradle. To obtain a movement's net metabolic cost in units of Joules, we multiplied the measured net metabolic rate in J/s for each movement duration and distance, by trial duration. Trial duration is the time between the start of consecutive movements, which represents the sum of the movement duration, the inter-trial time, and reaction time. The net metabolic rate for a movement of a fixed distance and duration was then calculated as the net metabolic cost of the movement divided by the movement duration. Note that \dot{e}_r represents the average not the instantaneous metabolic rate over the course of a movement. We next parameterized the relation between a movement's net energetic rate

and the duration and distance of the movement. We fit metabolic rate to Eq. (1), to determine the free parameters a , b , i , and j . The distance d was known. The effective mass of the arm and robot was estimated as $m = 2$ kg, based on an inertial model of the arm using standard anthropometric values (see section on Mass of the Arm and Movement Vigor for more details regarding the arm model).

SIMULATIONS

Goodness of fit for each model was determined by calculating the correlation coefficient and the sum of squared errors, SSE. Parameter values and goodness of fit for each simulation are provided in Supplemental Tables S1-S3. Simulation code is available upon request.

Mass of the arm and movement vigor

We tested the predictions of the theory in conditions where the mass of the limb was varied via the direction of the reach (Figure 2). We considered an inertial model of the human arm that was composed of two segments, with the following properties:

$$\begin{aligned} d_1 &= 0.33 & d_2 &= 0.43 & \text{meters} \\ m_1 &= 1.93 & m_2 &= 1.52 & \text{kg} \\ \lambda_1 &= \frac{d_1}{2} & \lambda_2 &= \frac{2d_2}{3} & \text{meters} \\ I_1 &= 0.014 & I_2 &= 0.019 & \text{kg m}^2 \end{aligned}$$

In the above expressions, d_i is length of each segment, m is mass, λ is length from point of rotation of the segment to its center of mass, and I is the inertial of the segment, with the subscript 1 referring to the upper arm, and subscript 2 referring to the forearm and hand. To predict what the movement duration and velocity should be for each direction θ , we first computed the effective mass along that direction $m(\theta)$ by computing the length of the vector that resulted when an acceleration of 1 m/s^2 in the direction of movement was multiplied by the mass matrix M . The result was scalar value function $m(\theta)$, which was then used to compute the predicted duration for a reach in that direction, with an amplitude of 10cm, that is, $d = 0.1$ (as shown in Figure 2B). We then computed the peak velocity of the resulting movement using a minimum-jerk trajectory.

To compute effort of each movement, we used the energetic costs that we had measured in Figure 1A, and set $\gamma = 1$, with the only free parameter α . We found that a value that provided a good fit to the data of Gordon et al. [S25] was $\alpha = 310$.

Mass of the arm and choice of the movement

We next tested the predictions of the theory by considering the choices that people made when they were free to choose their own movement direction. In this experiment, the right and left arms were placed in a given configuration and the subjects were asked to make an out-and-back reaching movement to a circle of 15cm radius, but to a direction of their choice [S26]. The resulting probability distribution of the directions that they chose is shown in Figure 2C (gray region). To see whether our theory could account for the data, we fit the data, with only one free parameter, α , keeping all other parameters unchanged from the simulations shown in Figure 2B. We found that a value that provided a good fit to the data was $\alpha = 100$.

We first computed the effective mass for the left and right arms for the out-and-back movement by using the mass at the start point and each possible turn-around point about a circle of 15cm radius. We then used this effective mass to predict the duration of each 30cm movement, and then to compute the utility of that movement. For each possible movement direction we computed its utility, and then computed the ratio of this utility to the sum of utilities across all movement directions, producing the following probability:

$$\Pr(\theta_i) = \frac{J(\theta_i)}{\sum_i J(\theta_i)} \quad (\text{S30})$$

The results are shown with the black curve in Figure 2C, with the effective mass distribution for an out-and-back movement plotted in Figure 2D. In Figure 2E, the sum of utilities for quadrants 1 and 3 is compared to the fractions of trials that the subjects chose to reach to those quadrants.

Our formulation of utility function was further tested by considering the choices that people made when they were given the option of reaching to one of two possible targets (Figure 2F). The idea was that for each target, the effective mass of the movement described the utility for that movement, and the difference in the utilities associated with the two targets should describe the probability of choosing one target over another. We kept the parameters that we had found in Figure 2C unchanged. This produced a utility function with nothing to fit. To compute the effective mass for the reach to a given target, we computed the effective mass at the start and end points and averaged the two. To compute the probability of choosing a target, we used a logistic function in which the probability was a function of the difference in the utility of each target. Target T1 had a utility that was 16% larger than the utility of target T2. When the two targets were equally distant from the start point, the subjects chose target T1 on around 80% of the trials. To model the choice of targets as a function of movement distance, we used a logistic function:

$$\Pr(T1) = \left[1 + \exp \left(k \left(\frac{1}{2} - \frac{J_{T1}}{J_{T1} + J_{T2}} \right) \right) \right]^{-1} \quad (S31)$$

Here, k is a free parameter representing noise in the decision making process and was fixed to the same value in all simulations. Figure 2H illustrates the fit of the function to the data for the probability of choosing target T1 over target T2, and target T3 over target T4. As the distance to target T1 and T3 increases, the preference shifts to target T2 and T4. In the logistic function, the only free parameter was k , which we found to be 28 for the data in Figure 2I.

To walk or to fly

One of our main ideas is that representation of effort may depend on temporally discounted energetic cost of action. To test this idea, we considered an experiment in which both the caloric values of reward and the energetic cost of performing the action were objectively known.

Bautista et al. [S27] trained starlings to choose between walking a short distance or flying a long distance to acquire a piece of food. The caloric content of this reward was known, $\alpha = 1.3 \times 10^3$, as was the energetic rate associated with walking, perching, and flying: $\dot{e}_w = 2$, $\dot{e}_p = 1.09$, and $\dot{e}_f = 31.7$. Time spent in each act was also known: $t_w = 0.6$, $t_p = 1.25$, and $t_f = 1.1$ sec. According to our theory, the utility function takes the following form:

$$J = \frac{\alpha - e}{1 + \gamma T} \quad (S32)$$

Here, T is total travel time, e is energetic cost of the movement, and γ is a temporal discounting factor. The variable T represents the time the animal spends performing three different activities: moving, perching, and handling the food. The movement time t_w is the time the birds spend walking in a one-way trip (or time spent flying, represented as t_f). In addition to walking or flying, the animals spent time perching in between walking or flying one-way trips (t_p), and there is additional time spent handling the reward (t_h) before they consumed it. If the animal chose to walk n_w times to acquire reward, the travel time is:

$$T_w(n_w) = 2n_w(t_w + t_p) + t_h \quad (S33)$$

Bautista et al. [S27] estimated the energetic rate during perching \dot{e}_p from previous recordings and assumed that the energetic rate of handling \dot{e}_h was the same. Thus the total energetic cost for making n_w walking trips is:

$$e_w(n_w) = 2n_w(\dot{e}_w t_w + \dot{e}_p t_p) + \dot{e}_p t_h \quad (S34)$$

Combining the above equations, we find the utility for the choice of taking n_w walking trips:

$$J_w(n_w) = \frac{\alpha - (2n_w(\dot{e}_w t_w + \dot{e}_p t_p) + \dot{e}_p t_h)}{1 + \gamma(2n_w(t_w + t_p) + t_h)} \quad (S35)$$

We can similarly define the utility for flying, J_f .

The indifference point is found by setting $J_w(n_w) = J_f(n_f)$ and solving for n_w . The only free parameter is γ . We found that changing γ had a monotonic effect on the indifference curve: as Figure 3A illustrates, small γ (patient animal) led to a preference for walking, whereas large γ (impulsive animal) led to a preference for flying. The data was best fit for $\gamma = 0.03$.

If neither reward nor effort is temporally discounted, then there are no free parameters in the utility function. The number of walks is proportional to the number of flights, where the proportion is determined by the ratio of the energetic cost of flying to walking (Figure 3B).

Finally, if reward is temporally discounted but not effort, then once again there is only one free parameter, γ . However, in this case changing γ has a non-monotonic effect on the indifference curve. When γ is very small (patient animal, $\gamma = 0.0001$), there is a preference for walking (Figure 3C). When γ is very large (impulsive animal, $\gamma = 1$), there is once again a preference for walking (Figure 3C). The closest that we can come to the measured data is with $\gamma = 0.007$, which provides a poor fit. Indeed, if we assume that effort is not temporally discounted, the mean-squared error between the model and data is an order of magnitude larger than if effort is temporally discounted.

Sensitivity of results to parameter values

We chose to represent utility as a sum of effort and reward. A number of earlier approaches have used a multiplicative approach, where effort multiplicatively discounts reward [S14-S16]. An additive formulation is a better fit to the data because a multiplicative interaction cannot account for the observation that increased reward results in increased movement vigor [S10, S28].

We chose to discount the utility via a hyperbolic function of time. Earlier works have suggested an exponential discounting [S18]. Regardless of parameter values, hyperbolic temporal discounting is a better fit to the data because exponential discounting cannot account for the observation that vigor declines with increased inter-trial interval [S12]. Hyperbolic discounting performs as well or better, compared to exponential discounting in the choices birds made in walking vs. flying and the choices people made in isometric force production.

We considered the possibility that only reward but not effort may be temporally discounted. Such a scenario produced a model that had the same number of parameters as in the utility where both effort and reward were temporally discounted. However, we found that temporal discounting of reward and effort always produced a better fit to the measured data in choices that birds made in walking vs. flying, and choices that people made in isometric force production.

Supplemental References

- S1. Russ, D. W., Elliott, M. A., Vandenborne, K., Walter, G. A., and Binder-Macleod, S. A. (2002). Metabolic costs of isometric force generation and maintenance of human skeletal muscle. *Am. J. Physiol. Endocrinol. Metab.* 282: E448–57.
- S2. Kushmerick, M. J., and Paul, R. J. (1976). Aerobic recovery metabolism following a single isometric tetanus in frog sartorius muscle at 0 degrees C. *J Physiol-London* 254: 693–709.
- S3. Srinivasan, M. (2009). Optimal speeds for walking and running, and walking on a moving walkway. *Chaos* 19: 026112.
- S4. Ralston, H. J. (1958). Energy-speed relation and optimal speed during level walking. *Int Z Angew Physiol* 17: 277–283.
- S5. Hoyt, D. F., and Taylor, C. R. (1981). Gait and the Energetics of Locomotion in Horses. *Nature* 292: 239–240.
- S6. Kawagoe, R., Takikawa, Y., and Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nat Neurosci* 1: 411–416.
- S7. Xu-Wilson, M., Zee, D. S., and Shadmehr, R. (2009). The intrinsic value of visual information affects saccade velocities. *Exp Brain Res* 196: 475–481.
- S8. Opris, I., Lebedev, M., and Nelson, R. J. (2011). Motor Planning under Unpredictable Reward: Modulations of Movement Vigor and Primate Striatum Activity. *Front. Neurosci.* 5: 61.
- S9. Sackaloo, K., Strouse, E., and Rice, M. S. (2015). Degree of Preference and Its Influence on Motor Control When Reaching for Most Preferred, Neutrally Preferred, and Least Preferred Candy. *OTJR: Occupation, Participation and Health*, 1–8.
- S10. Reppert, T. R., Lempert, K. M., Glimcher, P. W., and Shadmehr, R. (2015). Modulation of Saccade Vigor during Value-Based Decision Making. *J Neurosci* 35: 15369–15378.
- S11. Choi, J. E. S., Vaswani, P. A., and Shadmehr, R. (2014). Vigor of Movements and the Cost of Time in Decision Making. *J Neurosci* 34: 1212–1223.
- S12. Haith, A. M., Reppert, T. R., and Shadmehr, R. (2012). Evidence for hyperbolic temporal discounting of reward in control of movements. *J Neurosci* 32: 11727–11736.
- S13. Levine, R. V., and Norenzayan, A. (1999). The Pace of Life in 31 Countries. *Journal of Cross-Cultural Psychology* 30: 178–205.
- S14. Prevost, C., Pessiglione, M., Metereau, E., Clery-Melin, M.-L., and Dreher, J.-C. (2010). Separate Valuation Subsystems for Delay and Effort Decision Costs. *J Neurosci* 30: 14080–14090.
- S15. Sugiawaka, H., and Okouchi, H. (2004). Reformative self-control and discounting of reward value by delay or effort1. *Japanese Psychological Research* 46: 1–9.
- S16. Pasquereau, B., and Turner, R. S. (2013). Limited Encoding of Effort by Dopamine Neurons in a Cost-Benefit Trade-off Task. *J Neurosci* 33: 8288–8300.
- S17. Sutton, R., and Barto, A. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychol. Rev.* 88: 135-170.

- S18. Schweighofer, N., Shishida, K., Han, C. E., Okamoto, Y., Tanaka, S. C., Yamawaki, S., and Doya, K. (2006). Humans Can Adopt Optimal Discounting Strategy under Real-Time Constraints. *PLoS Comput Biol* 2: e152–8.
- S19. Myerson, J., and Green, L. (1995). Discounting of delayed rewards: Models of individual choice. *J Exp Anal Behav* 64: 263–276.
- S20. Kobayashi, S., and Schultz, W. (2008). Influence of Reward Delays on Responses of Dopamine Neurons. *J Neurosci* 28: 7837–7846.
- S21. Brooks, G. A., Fahey, T. D., and White, T. P. (1996). Exercise physiology: Human bioenergetics and its applications. (Mountain View: Mayfield Publishing).
- S22. Siedell, J. C., Muller, D. C., Sorkin, J. D., and Andres, R. (1992). Fasting Respiratory Exchange Ratio and Resting Metabolic-Rate as Predictors of Weight-Gain - the Baltimore Longitudinal-Study on Aging. *Int. J. Obes. Relat. Metab. Disord.* 16: 667–674.
- S23. Short, K. R., and Sedlock, D. A. (1997). Excess postexercise oxygen consumption and recovery rate in trained and untrained subjects. *J Appl Physiol* 83: 153–159.
- S24. Brockway, J. M. (1987). Derivation of formulae used to calculate energy expenditure in man. *Hum Nutr Clin Nutr* 41: 463–471.
- S25. Gordon, J., Ghilardi, M. F., Cooper, S. E., and Ghez, C. (1994). Accuracy of Planar Reaching Movements. II. Systematic Extent Errors Resulting From Inertial Anisotropy. *Exp Brain Res* 99: 112–130.
- S26. Wang, W., and Dounskaia, N. (2012). Load emphasizes muscle effort minimization during selection of arm movement direction. *J Neuroeng Rehabil* 9: 70.
- S27. Bautista, L. M., Tinbergen, J., and Kacelnik, A. (2001). To walk or to fly? How birds choose among foraging modes. *P Natl Acad Sci Usa* 98, 1089–1094.
- S28. Takikawa, Y., Kawagoe, R., Itoh, H., Nakahara, H., and Hikosaka, O. (2002). Modulation of saccadic eye movements by predicted reward outcome. *Exp Brain Res* 142, 284–291.