



Movement vigor: consistency, conservation, and modulation

Ph.D. Dissertation

Thomas R. Reppert

Abstract

This thesis is devoted to a better understanding of the vigor with which primates perform everyday actions. Indeed, there are great differences in baseline amounts of movement vigor across individuals. However, vigor depends upon the environment, and the reward landscape. The first portion of this text focuses on the effects that the reward landscape has on movement vigor. Chapter 2 demonstrates the effect of rate of reward on vigor, in particular reaction time and peak velocity. In Chapter 3, we analyze the effects of monetary reward offerings on saccade vigor. We show that choice preference for one reward over another is manifest in saccade vigor at the time of decision. We also show that saccade vigor tracks the decision making process, with a steep drop in vigor once a decision is made. Finally, in Chapter 4, we investigate changes in movement vigor during combined head-free gaze shifts and reaching movements. The objective is threefold: 1. Analyze *consistency* of head and reaching movements within an individual, across days; 2. Analyze *conservation* of movement vigor within an individual, across movement modalities; and 3. Assess *modulation* of movement vigor as a function of time, to determine if changes in one modality predict simultaneous changes in other modalities.

Contents

1	Introduction.....	1
1.1	Saccadic eye movements	1
1.2	The neural basis of saccades	1
1.2.1	Brainstem nuclei	1
1.2.2	Superior colliculus	2
1.2.3	Frontal cortex	2
1.3	The main sequence of saccades	2
1.4	Neural basis of vigor	4
1.5	Scope of this thesis.....	5
2	Evidence for hyperbolic temporal discounting of reward in control of movements.....	7
2.1	Introduction.....	7
2.2	Materials and methods	8
2.2.1	Model of the eye plant	8
2.2.2	Making saccades to maximize probability of success.....	10
2.2.3	Discounting probability of success	11
2.2.4	Experimental methods.....	11
2.3	Results.....	13
2.3.1	Temporal discounting of reward	15
2.3.2	Temporal discounting as reward rate optimization.....	16
2.3.3	Predictions of the rate of reward theory	17
2.3.4	Change in ITI alters saccade vigor.....	18
2.3.5	Control experiment	21
2.3.6	Characteristics of the temporal discount function.....	23
2.4	Discussion	26
3	Modulation of saccade vigor during value-based decision making	30
3.1	Introduction.....	30
3.2	Materials and methods	31
3.2.1	Participants.....	31
3.2.2	Behavioral task.....	31
3.2.3	Data analysis: saccade vigor	34
3.2.4	Data analysis: decision-making	34

3.3	Results.....	36
3.3.1	Saccade vigor was higher during the deliberation period	36
3.3.2	Saccade vigor encoded preference	39
3.3.3	Between subject differences in saccade vigor.....	42
3.4	Discussion	44
3.4.1	Link between the neural basis of vigor and encoding of reward	44
3.4.2	Subjective value vs. motivational salience.....	44
3.4.3	Modulation of vigor during decision-making	45
3.4.4	Between-subject differences in vigor.....	46
3.4.5	Conclusions.....	46
4	A shared mechanism for control of vigor in saccades, reaching, and head movements	48
4.1	Materials and methods	48
4.1.1	Experiment 1	48
4.1.2	Experiment 2.....	49
4.1.3	Data analysis	50
4.2	Results.....	50
4.2.1	Conservation of vigor of saccades	50
4.2.2	Reaction times were lower for horizontal saccades	53
4.2.3	Speed-accuracy tradeoff in saccades.....	54
4.2.4	Conservation of vigor across movement modalities	56
4.2.5	Temporal modulation of movement vigor	60
5	Discussion.....	62
6	References.....	63

Figure 2-1: Relationship between saccade duration, endpoint accuracy, and expected value of reward. ..	14
Figure 2-2: Changes in saccade vigor associated with changes in inter-movement interval, as predicted by a rate of reward cost function.....	18
Figure 2-3: Experimental protocol, model predictions, and results.	19
Figure 2-4: Control experiment.....	22
Figure 2-5: Effect of inter-movement interval on movement duration under various temporal discounting regimes.....	24
Figure 3-1: Experimental design.....	32
Figure 3-2: Saccade vigor was higher during deliberation than following decision-time.	37
Figure 3-3: Saccade vigor encoded choice preference just before time of decision.	39
Figure 3-4: The stimulus that was target of gaze just before the key-press was often the option that was eventually chosen.....	41
Figure 3-5: Between-subject differences in vigor.....	43
Figure 4-1: Conservation of vigor across saccade directions.....	52
Figure 4-2: Subjects responded to horizontal target jumps quicker than to vertical jumps.	53
Figure 4-3: We found a small yet significant relationship between speed and accuracy of saccades.	55
Figure 4-4: Reaction time and peak velocity of movements (eye, head, and hand) vs. target displacement for all subjects that completed a total of 16 blocks (n=16).	57
Figure 4-5: Consistency of vigor across movement modalities.	59
Figure 4-6: Within-subject modulation of movement vigor for those subjects that completed 4-6 blocks of Experiment 2 (n=20).	61

1 Introduction

1.1 Saccadic eye movements

The term “saccade” has French origin, and refers to either the jerking of a horse’s reigns, or the flicking of a sail in the wind [1]. In the biological sciences, saccade refers to a rapid eye movement, and the term was originally used in this context to describe movement of the eye during reading or free viewing. The next several chapters will focus on two primary types of saccades: voluntary and reflexive. Voluntary saccades are purposeful movements made to orient the viewer to a new location with the intent of acquiring new information. Reflexive saccades, on the other hand, are generated in response to an unexpected stimulus.

The velocity profile of saccades has a stereotypical bell shape. This shape results from the combination of acceleration and deceleration phases of the movement. For larger saccades (greater than roughly 15 deg in size), the velocity profile becomes skewed to the right. That is, the deceleration phase is longer than the acceleration phase. This skew of the velocity profile reflects the biological limits on the maximum velocity with which primates can move their eyes.

1.2 The neural basis of saccades

1.2.1 Brainstem nuclei

Primates use saccades to place the fovea on an intended location in space. In order to do so, the brain must perform three steps: 1. Collect information regarding the intended location relative to other locations in space; 2. Transform that location into a motor command used to orient the primate to the location; and 3. Execute the motor command [1]. The brainstem houses several nuclei which are vital for this final step. In particular, the paramedian pontine reticular formation (PPRF) and rostral interstitial nucleus of the medial longitudinal fasciculus (riMLF) contain excitatory burst cells for horizontal and vertical saccades, respectively. Indeed, a unilateral lesion to the PPRF preempts the ability to make an ipsilateral eye movement. The brainstem also holds inhibitory burst neurons, which are vital for suppression of antagonist ocular motor neurons. The brainstem is also home to omnipause neurons, located in the nucleus raphe interpositus, at the level of the midline of the pons. These neurons exhibit tonic firing, except for during a saccade, for which they are silenced. Stimulation of these neurons during a saccade leads to rapid deceleration of the eye movement. Intracellular recordings have revealed an inhibitory input to the omnipause neurons, the origin of which acts as a switch to enable the execution of saccades [1].

1.2.2 Superior colliculus

Cortical eye-movement related areas project heavily to the superior colliculus (SC), located at the level of the midbrain of the brainstem. The SC can be thought of as having two independent layers, a dorsal layer and a ventral layer. The dorsal layer contains a map of the visual field, whereas the ventral layer contains a “motor map” of intended saccade displacement. This ventral layer can further be specified as having a rostral pole, which dictates fixation and small saccades, and a caudal end, which responds to target selection and the initiation of gaze shifts. Indeed, stimulation of the rostral pole of ventral SC suppresses saccades [1]. Activity in SC is often referred to as “all or none,” a scenario in which once a particular threshold of neural activity is reached, a saccade is then generated.

Ventral SC has two types of neurons: build-up neurons, which exhibit activity well before the saccade-related burst of activity, and burst neurons, which are intermixed with the former. The site of maximum activity on the collicular motor map strongly correlates with the intended saccade direction and size. The motor map can be thought of as having two axes: one of which dictates size, and the other of which dictates direction.

1.2.3 Frontal cortex

The area of frontal cortex which has received the most attention regarding saccade generation is the frontal eye field (FEF). Stimulation of FEF typically elicits a saccade; however, it may also suppress saccade generation, depending upon the location in SC to which the stimulated neurons project. Previous work has found FEF to be strongly modulated by the presence of reward [1]. Stimulation of neurons in supplementary eye field (SEF) also elicits a saccade, although at a longer latency than in FEF. In particular, neurons in SEF are active when the subject performs a sequence of saccades from memory. Additionally, dorsolateral prefrontal cortex (dlPFC) houses activity related to the voluntary control of saccades, such as execution of saccades to remembered target locations.

1.3 The main sequence of saccades

The main sequence describes the relationship between the amplitude, duration, and peak velocity of saccades. Both the duration and peak velocity change in nonlinear fashion as functions of saccade amplitude. In particular, duration increases linearly for saccades up to roughly 10 deg, before the slope changes and duration increases more steeply. Peak velocity increases linearly with amplitude up to amplitudes of 15 to 20 deg, beyond which peak velocities become asymptotically stable [2]. In the late 20th century, the classic view was that the main sequence completely described the variability in saccade velocities. In [2], the authors state that “*the width of the controller signal determines the saccadic*

amplitude and also, because of the main sequence relationship, the duration and peak velocity of the saccade.”

Since then, many works have tried to explain the reason why such a main sequence relationship exists between the primary saccade parameters. In particular, Harris and Wolpert [3] showed that a model which placed a cost on the endpoint variability of the saccade could account for the nonlinear main sequence relationship of speed and size. They showed that signal-dependent noise in the motor command for saccade generation imposed a tradeoff between movement duration and accuracy. This tradeoff explained the shape of the main sequence of saccades. Other works, such as that of van Beers [4], studied the sources of variability of saccades in more detail. He showed that in addition to signal-dependent noise, both constant motor noise and sensory noise contributed to the variability inherent in saccade execution. He also showed that variability in size, duration, and peak velocity was relatively constant for saccades in different directions.

It wasn't until the early 21st century that more works began to consider factors other than sensorimotor noise which contributed to variability in saccade execution. One of the first works to do so was that of Takikawa and colleagues [5]. They showed that expectation of reward at the completion of the saccade also influenced the kinematics of the eye movement. In particular, they presented a series of trials to monkeys (*macaca fuscata*) during which most trials went unrewarded, but roughly 25% of trials were followed with juice reward. The monkeys' saccades were both faster and quicker (i.e. lower reaction time) to the rewarded targets. Curiously, these faster movements were executed with less variability in saccade amplitude, and a lower rate of movement execution errors.

Another more recent work by Xu-Wilson and colleagues showed a similar effect of expectation of reward on saccade kinematics in humans [6]. In that work, they instructed participants to make reflexive saccades to targets placed 15 deg apart. For a brief period of 500 ms, the target changed to a picture, one of four type: object, face, inverted face, and noise. After the saccade was completed, the participants were shown the image which they had been promised. Whereas saccade amplitude was not affected by the image type, saccade peak velocity was higher to the faces than to the objects or visual noise. There was no effect of image type on reaction time. In that work the authors also noted a steady decline in saccade peak velocity during each block. This decline in velocities was consistent with the notion that the rewarding value of the stimulus decreased with repetition.

As mentioned previously, Takikawa et al. [5] found that faster saccades were also more accurate than slower saccades. How can this be? A more recent work investigated the speed-accuracy tradeoff in saccades. Manohar and colleagues [7] devised a task in which human participants were asked to make a

saccade to a target, after having avoided a distractor target. The maximum reward to be received was broadcast at the beginning of the trial. Participants were rewarded as a function of reaction time, with longer reaction times yielding lower reward (monetary, in pence). The authors found that, after grouping trials post-hoc according to maximum reward amount, participants exhibited a speed-accuracy tradeoff; that is, longer reaction times were accompanied by higher success rates. However, the entire speed-accuracy tradeoff curve was shifted based on the maximum reward to be obtained. On trials with higher reward amount, participants were both faster and more accurate.

1.4 Neural basis of vigor

Both the speed and reaction time of saccades change with reward context. What is the neural basis of these changes in saccade execution? Kawagoe and colleagues studied the effect of reward expectation on saccade execution and neural activity in the caudate nucleus (CN) [8]. They found, as in [5], that expectation of juice reward modulated the speed and reaction times of saccades performed by monkeys (*macaca fuscata*). In addition, they found that expectation of reward modulated firing rates of neurons in CN. Most of the neurons that the authors analyzed exhibited firing rates that were facilitated by the expectation of reward. When the rewarded target was switched from one location to another at a set break, cell activity adapted to the new reward context very quickly. To further establish the relationship between saccade speed and CN firing rate, Itoh and colleagues [9] calculated a correlation coefficient for trials in which expectation of reward was constant. They found that, for constant reward, some neurons exhibited an increased firing rate with increased saccade velocity. This lent further validity to the notion that CN firing rate was related to the motivational state of the monkey.

The vigor with which a saccade is performed is associated with activity of ‘build-up’ cells in the intermediate layers of the superior colliculus (SC) [10]. When a saccade is planned toward a location that falls within the receptive field of a superior colliculus cell, the upcoming saccade displays greater vigor if that cell fires more strongly during the period before the saccade. This build-up activity is partly under the control of cells in an output nucleus of the basal ganglia, the substantia nigra pars reticulata (SNr). SNr cells constantly inhibit SC, but generally pause before a movement [11] [12]. More vigorous saccades are associated with a deeper pause in the firing rates of SNr cells [13], and reward itself is known to modulate the depth of this pause [11]. Indeed, saccadic vigor is increased by blocking the SNr-SC inhibition [12]. Therefore, control of vigor is partly a function of the basal ganglia.

Within the basal ganglia, some cells in the caudate influence the discharge of SNr neurons directly, while other cells do so indirectly via their projections to the external segment of globus pallidus (GPe). Caudate cells receive dopamine projections and generally fire more before a rewarding saccade [8]. Onset of a stimulus that promises reward results in a burst of dopamine [14], which is followed by a

more vigorous saccade [15]. Indeed, chronic reduction in the concentration of dopamine in the caudate reduces saccade vigor by around 30% [16]. GPe cells inhibit SNr and fire more strongly preceding a more vigorous saccade, and bilateral lesion of this region eliminates the ability of the animal to modulate saccade vigor in response to changes in reward [15]. Therefore, control of vigor is partly associated with the amount of dopamine in the basal ganglia, modulating activity of caudate, affecting the depth of pause in the SNr.

One of the primary symptoms of Parkinson's disease is a compromise of the ability to make fast, accurate movements. As mentioned previously, Manohar et al. [7] found that the presence of reward shifted the speed-accuracy tradeoff curve of healthy participants. They also analyzed this tradeoff in PD patients. They found slower mean saccade speeds and longer reaction times. However, importantly, the amount of reward offered did not affect the patients' motor performance as it did in the healthy age-matched controls. The authors state that "*these findings are consistent with reduced reward sensitivity in PD.*"

In a more recent work, Panigrahi and colleagues used a mouse model of Parkinson's disease to study the relationship between the presence of dopamine and movement vigor [17]. They found that the neural representation of forelimb movements in the dorsal striatum was altered by the progressive depletion of dopamine neurons (PDD). The number of neurons with a significant modulation of activity at movement onset was much greater in wild type mice than in dopamine compromised mice. Administration of oral Levodopa to the compromised mice improved their rate of movements, and increased the vigor of those movements. In summary, dopaminergic input the dorsal striatum appeared to play a role in the ability of the striatum to mediate changes in movement vigor.

1.5 Scope of this thesis

This thesis is devoted to a better understanding of movement execution, and in particular movement vigor. As a student in the Laboratory for Computational Motor Control the past seven years, I have worked next to many colleagues, both students and postdocs. One thing that I noted along the way was the vast differences in how each individual typed at the keys of their computer. Some students pressed the keys quite rapidly, and seemingly with great haste, whereas others typed softly. Most interestingly, these differences persisted across the entire tenure of those students' stay at Johns Hopkins. Why were there such differences in vigor of typing?

This work investigates the effects of movement frequency and reward on movement vigor, and it finishes with an analysis of natural movement vigor. Indeed, as I have found in the lab, there are great differences in baseline levels of movement vigor across individuals. However, vigor depends upon the

environment, and the reward landscape. The first portion of this text focuses on the effects that the reward landscape has on the vigor of the individual. In particular, Chapter 2 investigates the effect of inter-trial interval (ITI) on movement vigor, in particular reaction time and peak velocity. We show that movement vigor changes sharply in response to a change in rate of reward. Reaction time becomes shorter, and peak velocity higher. In Chapter 3, we analyze the effects of monetary reward offerings on saccade vigor. We show that choice preference for one reward over another is manifest in saccade vigor at the time of decision. We also show that saccade vigor tracks the decision making process, with a steep drop in vigor once a decision is made.

Finally, in Chapter 4, we investigate changes in movement vigor in a head-free gaze shift and reaching scenario. The objective in designing this last task was threefold: 1. Analyze *consistency* of head and reaching movements within an individual, across days; 2. Analyze *conservation* of movement vigor within an individual, across movement modalities; and 3. Assess *modulation* of movement vigor as a function of time, to determine if changes in one modality predict simultaneous changes in other modalities.

The questions that I attempt to address with this thesis are all founded in basic science; however, the implications of this work extend beyond basic science. In particular, this work can be used to supplement and extend current methods of diagnosing personality traits and disorders, such as impulsivity. One of the most widely used methods for assessing such traits is the questionnaire; for instance, the Barratt Impulsiveness Scale in the case of impulsivity. While such surveys are beneficial, patients' responses may be susceptible to the environment, such as time of day and mood. A behavioral assessment of movement vigor as a trait provides an additional survey which is based upon an exposition of personality, as opposed to a communication. The behavioral measure is likely to be less susceptible to factors like time of day and mood.

2 Evidence for hyperbolic temporal discounting of reward in control of movements

2.1 Introduction

Temporal discounting of reward is a ubiquitous phenomenon in decision making. Across many types and magnitudes of reward, multiple timescales, and various species, small, immediate rewards are often preferred over larger, delayed rewards. Mathematically, temporal discounting of reward may be described in terms of a multiplicative discount function:

$$V(t_o + t) = V(t_o)F(t). \quad (2.1)$$

In Eq. (2.1), reward value at current time t_o is discounted by a function $F(t)$ to produce value at time $t_o + t$, with $F(0) = 1$. The two most common forms of $F(t)$ that have been used to describe discounting are exponential

$$F(t) = \exp(-kt) \quad (2.2)$$

and hyperbolic

$$F(t) = 1 / (1 + \beta t). \quad (2.3)$$

For example, exponential temporal discounting is routinely used in a form of reinforcement learning known as temporal difference learning [18], which provides a prominent theory of learning in the basal ganglia [19]. Exponential discounting has also been suggested in models of human decision making [20]. Hyperbolic discounting, however, is more consistent with behavioral data in humans [21] and monkeys [22]. While it is clear that the brain temporally discounts reward, the exact shape of this function is not entirely clear. Perhaps more significantly, the reason why temporal discounting occurs at all is poorly understood.

Recently we proposed that the way the brain discounts reward may have implications for control of movements [23]. Suppose that a movement is made with the purpose of acquiring some rewarding state that has value $V(t_o) = a$. In this framework, the duration of the movement acts as a delay in acquiring reward. Performing a movement slowly diminishes the value of reward upon its acquisition (movement end), making it preferable to move quickly. Fast movements, however, are more variable [24] [25], reducing the probability of success for the movement. Therefore, the expected value of a stimulus that is acquired after some movement duration τ is affected by two factors: probability of

successfully acquiring the stimulus $P[\text{success}|\tau]$, which increases with duration τ , and temporal discounting of reward value $F(\tau)$, which decreases with duration τ :

$$E[\text{reward}|\tau] = aP[\text{success}|\tau]F(\tau). \quad (2.4)$$

Thus, if the objective for the brain is to produce movements that maximize the expected value of reward, then movement speed and duration or, collectively, vigor should be a balance between the competing concerns of time and variability.

However, our proposed link between temporal discounting in motor control and decision making is tenuous: the movements that we are considering (saccades) are tens of milliseconds in duration. Why should a few milliseconds make a meaningful difference in the value of reward? Here, we show that one interpretation of hyperbolic temporal discounting is that the brain selects actions so as to maximize the rate of reward. This idea leads to a novel prediction about how the brain should select vigor in response to changes in the inter-trial interval between movements. We propose that rate of reward provides a unifying principle that governs control of movements in timescale of milliseconds, as well as decision making in timescales of seconds to years.

2.2 Materials and methods

Our concern is the general question of why movements have their specific kinematic properties, i.e., why movements of a given amplitude have a particular duration and velocity. Here, we present a novel framework for considering the influence of temporal discounting of reward on movement vigor. Our focus is on saccades, as numerous theories have been proposed to explain the kinematic patterns of these simple movements, enabling us to focus on the question of how temporal discounting influences choice of movement vigor. Our principal new theoretical result, presented in the Results section, is that the shape of the discount function should leave its signature in how the brain alters saccadic vigor in response to changes in inter-trial intervals between saccades. We will first present the computational methods that we used to study the theoretical relationship between movement vigor and temporal discount functions, and then the experiments that we performed to test some of the predictions.

2.2.1 Model of the eye plant

We modeled the oculomotor plant as a second order dynamical system:

$$m\ddot{x} = -kx - b\dot{x} + f. \quad (2.5)$$

In this equation, x is the lateral deviation from the equilibrium point of the eye, m is the inertia of the eye, k is stiffness, b is viscosity, and f is the instantaneous force generated by the extra-ocular muscles, which act as a first-order linear filter of the motor command u :

$$\gamma \dot{f} = -f + u. \quad (2.6)$$

Here, γ is a time-constant that determines how quickly motor commands are transmitted into forces f .

If we represent the full state of the plant by the vector $\mathbf{x} = [x, \dot{x}, f]^T$, the dynamics can be more compactly expressed in continuous time as:

$$\dot{\mathbf{x}} = A_c \mathbf{x} + \mathbf{b}_c u \quad (2.7)$$

with $A_c = \begin{pmatrix} 0 & 1 & 0 \\ -k/m & -b/m & 1/m \\ 0 & 0 & -1/\gamma \end{pmatrix}$ and $\mathbf{b}_c = \begin{pmatrix} 0 \\ 0 \\ 1/\gamma \end{pmatrix}$. As with our previous work [23], we set the

parameters of the eye plant to match the three timescales described by Robinson et al. [26]: $\tau_1 = 0.224$, $\tau_2 = 0.013$, $\tau_3 = 0.004$ sec. This can be achieved by setting $k = 1$, $b = \tau_1 + \tau_2$, $m = \tau_1 \tau_2$, and $\gamma = \tau_3$. These equations were converted into discrete-time using matrix exponentials for a time step Δ of 0.1ms:

$$\mathbf{x}_{t+\Delta} = A \mathbf{x}_t + \mathbf{b} u_t.$$

Next, we added signal dependent $\varepsilon_t \sim N(0, \kappa^2 u_t^2 \Delta)$ and non-signal dependent $\chi_t \sim N(0, \lambda^2 \Delta)$ noise to the model:

$$\mathbf{x}_{t+\Delta} = A \mathbf{x}_t + \mathbf{b}(u_t + \varepsilon_t + \chi_t) \quad (2.8)$$

as described by van Beers [4]. It is difficult to reliably estimate the magnitude of signal-dependent noise from empirical data due to the many potential sources of variability in eye movements, although constant noise is more reliably inferred [4]. We therefore set $\lambda = 0.0075 \text{ kg m s}^{-2}$ to match the horizontal endpoint variability reported in that work, and left the magnitude of signal-dependent noise κ as an open parameter.

To extend this 1-dimensional model to a more realistic 2-dimensional one, we assumed independent vertical and horizontal components of the eye, with independent sources of noise. Following [4], we scaled variability in the vertical direction by a factor of 1.14 to reflect the sparser innervation of

muscles in that direction relative to horizontal. Our experiments presented targets approximately along the horizontal axis. As a result, the magnitude of signal dependent noise introduced for the vertical component of movement was found to be negligible for all movements we considered and we therefore included only signal-independent variability along this axis. We used this 2-dimensional plant model in all subsequent simulations.

2.2.2 Making saccades to maximize probability of success

Suppose that success or failure of a point-to-point movement such as a saccade is determined by whether or not the location of the effector at the end of the movement falls within a specified goal region. The probability of success depends on the distribution of the effector endpoints. For our linear system with Gaussian additive and multiplicative noise (Eq. 2.8), for a movement of duration τ , the endpoint distribution for any sequence of motor commands $u_0, \dots, u_{\tau-1}$ is Gaussian. Suppose that, in 1 dimension, for a particular sequence of motor commands the end position x_τ of the saccade has a distribution with mean μ and variance σ^2 . If the target of the movement is at location a with respect to the fovea, and the fovea has width w , then the probability of success (i.e., acquiring reward) is defined as:

$$P[\text{success}] = \int_{a-\frac{1}{2}w}^{a+\frac{1}{2}w} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_\tau - \mu)^2}{2\sigma^2}\right) dx_\tau. \quad (2.9)$$

We assumed that $w = 1^\circ$ (approximate width of the fovea). If we assume that the mean of the endpoint distribution is aligned with the target center, the probability of success becomes:

$$P[\text{success}] = \text{erf}\left(\frac{w}{2\sqrt{2\sigma^2}}\right). \quad (2.10)$$

In 2-dimensions, we made the simplifying assumption that the overall probability of success was given by a product of the corresponding probabilities of success along the horizontal and vertical axes (effectively assuming that the target was a $1^\circ \times 1^\circ$ square).

For a given target position, we needed to compute the probability of success as a function of duration τ . To do so, for a given τ we found the motor commands that maximized the probability of success. This was achieved by finding the motor commands that minimized endpoint variance with the constraint that the mean of the endpoint distribution was at the target position at time τ and remained

there for a further 50ms hold period. We solved this constrained optimization problem analytically using Lagrange multipliers.

2.2.3 Discounting probability of success

How does one decide the vigor with which to perform a movement? A simple hypothesis is that we choose the motor commands that maximize probability of success (Eq. 2.10) [27]. However, as we will see, generating motor commands that maximize probability of success produces movement durations that agree with observed data on small amplitude saccades, but fails for large amplitude saccades. Our theory proposes that this failure is because reward does not have a constant value as a function of time. Rather, reward is discounted as a function of movement duration τ by a temporal discount function $F(\tau)$. We consider different forms for this temporal discount function (as described in Results), compute the discounted reward value, and then numerically find the movement duration that maximizes the expected value of the discounted reward function.

The value of the signal dependent noise parameter κ is unknown and difficult to estimate. Furthermore, the duration-amplitude relationships of all models we considered were highly sensitive to the exact value of this parameter. To generate the relationship between saccade amplitude and duration, we fit this single open parameter separately for the discounted reward model and for the maximum probability of reward model by finding the values that yielded a predicted duration of 105ms for a 30° saccade. This yielded estimates for κ of .0066 and .0055 for the discounted reward and maximum probability of reward models, respectively. The estimate of κ for the discounted reward model was then used directly to predict the influence of changes in inter-trial interval on movement vigor. There, we also assumed that subjects maintained an estimate of the inter-trial interval on each trial \hat{p}_i , and then updated this estimate based on the observed inter-trial interval: $\hat{p}_{i+1} = \hat{p}_i + \eta(p_i - \hat{p}_i)$. We set $\eta = 0.7$. We then set the inter-saccade interval in the discount function (δ in Eq. 2.15) equal to the estimated inter-trial interval plus a reaction time of 150 ms and determined the movement duration that theoretically maximized the discounted reward. Peak velocity was computed by simulating the saccade at the optimal movement duration using the motor commands that minimize the endpoint variance.

2.2.4 Experimental methods

A critical prediction of our theoretical work is that, if movement duration is determined by maximizing reward rate, then the brain should alter movement vigor in response to changes in the inter-trial interval

(ITI) between movements in a particular way. Our theory suggests that alternative forms of discounting may have different characteristic patterns by which changes in ITI lead to changes in movement vigor. We performed experiments to test our predictions. All experimental procedures were approved by the Johns Hopkins Institutional Review Board.

For our main experiment, we recruited $n=6$ healthy volunteers (mean age 28, range 23-47, five females). Subjects sat in a dark room in front of a CRT monitor (36.5 x 27.5 cm, 1024 x 768 pixel, light grey background, frame rate 120Hz) with head restrained using either a dental bite bar or chin and forehead rests and their left eye covered. Targets (blue, diameter = 1 deg) were presented with Matlab 7.4 (Mathworks) using the Psychophysics Toolbox. The screen was placed at a distance of 31 cm from the subject's face, and an EyeLink 1000 (SR Research) infrared camera recording system (sampling rate = 1000 Hz) was used to record movement of the right eye.

Subjects were asked to make saccades between targets having a horizontal separation of 40° , and positioned symmetrically about the center of the screen. Each saccade was cued by the appearance of one of two possible targets, having a vertical separation of 5° between them. Including two potential targets discouraged subjects from generating predictive saccades in advance of the actual target presentation. After an initial training period, subjects completed 12 blocks of 80 trials. Each trial consisted of three parts: inter-trial interval, reaction time, and movement time. The inter-trial interval began when the subject's gaze was within 3° of the target and ended with extinction of the current target and presentation of the next target. There were three different block types in which the inter-trial interval was varied in different ways. In the constant ITI blocks, the ITI was fixed at 1s throughout the block. In the increasing ITI blocks, the ITI was set to 1s for the first 10 trials, then was abruptly decreased to 0.4s, before slowly increasing to 1.6s over the next 60 trials, then was restored to 1s for the final 10 trials of the block. In the decreasing ITI blocks, the opposite sequence of ITIs was used, with an initial abrupt increase to 1.6s preceding a slow decrease to 0.4s. Blocks were presented in a pseudorandom sequence that was different for each individual subject.

An additional $n=5$ subjects (mean age 26, age range 21-47, 4 females) participated in a control experiment in which the previous target did not disappear when the new target was presented, but remained on the screen until the gaze reached the new target. Instead of two potential targets for each saccade, there was only one possible target in this control experiment.

All data analysis was completed using Matlab R2011a (Mathworks). The gaze position data were filtered using a 2nd-order Savitzky-Golay filter with a half-width of 27 ms. Saccade beginning and end were marked using a 20-deg/sec velocity threshold. Five criteria were used to assess saccades: (1)

Amplitude between 35 and 45 deg. (2) Duration between 50 and 350 ms. (3) Reaction time between 100 and 500 ms. (4) No blinking during the saccade. (5) Saccade velocity profile exhibits only one maximum. Any saccade that did not meet all 5 criteria was excluded from the analysis (about 10% of all saccades).

Peak velocities in a given block were normalized separately for nasal (leftward) and temporal (rightward) saccades by dividing by the mean peak velocity during the first 10 trials across all blocks. The normalized peak velocity was then averaged across all four repeats per block type for each subject. Saccade duration, amplitude, and reaction time were normalized using the same method. We assessed the effect of changing ITI on the variables of interest using an analysis of covariance on data from all subjects for the middle 60 trials of each block (during which the ITI changed), with trial number serving as a continuous predictor and block type as a categorical predictor. Our hypothesis that changes in ITI should lead to changes in the kinematic properties of saccades then corresponds to a predicted interaction effect between trial number and block type.

2.3 Results

An early model of saccades [3] showed that saccade trajectories for a given duration are well predicted by a model in which motor commands are selected to minimize endpoint variance in the presence of signal-dependent noise (term ε_t in Eq. 2.8). That idea is equivalent to finding motor commands that maximize the probability of success for a given duration. However, the question of how movement durations are selected was not addressed. A more recent work [4] empirically demonstrated that in addition to signal dependent noise, the oculomotor plant suffers from non-signal dependent noise (term χ_t in Eq. 2.8). The non-signal dependent noise acts as a natural cost that penalizes movement durations: the longer the duration of the movement, the greater the endpoint variance due to accumulation of this kind of noise. The impact of non-signal dependent noise grows monotonically with saccade duration, producing a greater penalty for longer durations of movement. Therefore, if one assumes that the objective is to maximize probability of success (or, equivalently minimize endpoint variance), then one can compute the optimal movement duration for any given saccade amplitude [27].

We adopted a model of the oculomotor plant based on previous publications [26] [4]. For reasons which we explain below, we used a value of $\kappa = 0.0066$ for the signal-dependent noise magnitude in our simulations. Using this model, we computed the probability that at the end of a 10° saccade the target would be placed on the fovea (Figure 2-1A, top subplot). Increasing movement duration initially increases the probability of success (this is because of the diminishing impact of signal-dependent noise). However, as movement durations become long, the probability of success tends to decline (this is because

of the increasing influence of non-signal dependent noise). Therefore, if the objective is to maximize probability of success, a 10° saccade should last approximately 56ms, a prediction that falls within the range of durations measured for actual saccades of that amplitude [28].

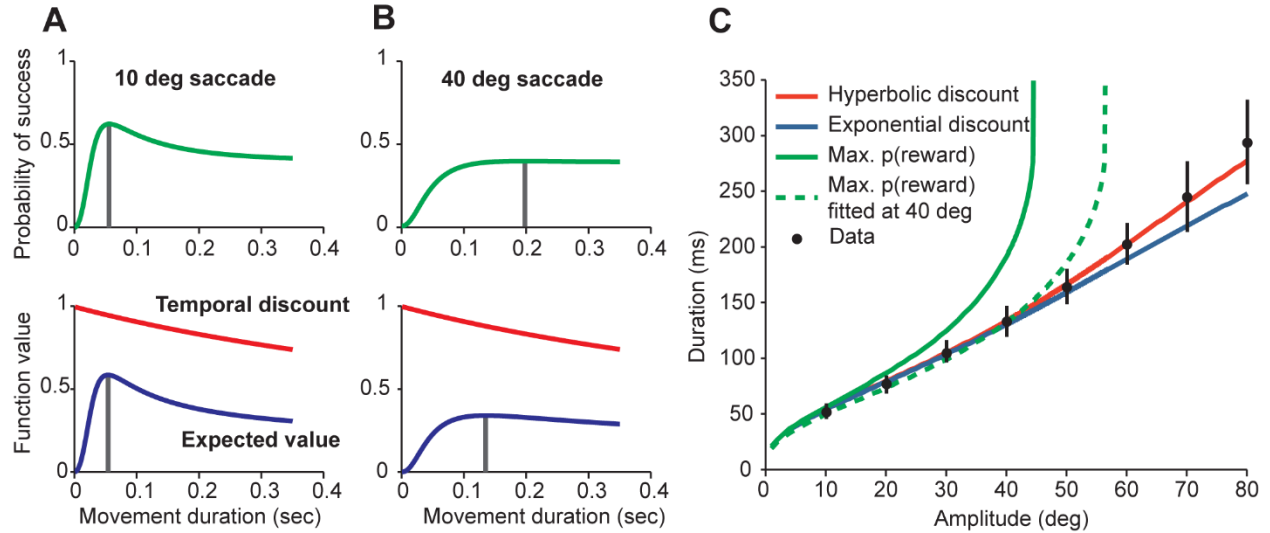


Figure 2-1: Relationship between saccade duration, endpoint accuracy, and expected value of reward. (A) Upper panel: we computed probability of success (probability that the motor commands will place the target on the fovea) for a 10° saccade as a function of movement duration. Signal-dependent noise magnitude was set to $\kappa = .0066$. Movement duration that maximizes probability of success is indicated by the vertical line. Lower panel: expected discounted value of the reward attained at the completion of the saccade (blue line) under the assumption that reward is discounted hyperbolically in time (red line). The temporal discount function is $F(\tau) = 1 / (1 + \tau)$. For a 10° saccade, the movement duration that maximizes the probability of success is similar to that which maximizes the expected value of reward. (B) Same as part A but for a 40° saccade. For this saccade amplitude, the movement duration that maximizes the expected value of reward is much shorter than one that maximizes the probability of success. (C) Relationship between saccade amplitude and saccade duration predicted by maximum probability of success hypothesis (green), and maximum expected value of reward hypothesis (red and blue). For the expected value of reward hypothesis, durations that maximize hyperbolically discounted expected rewards are shown in red, and durations that maximize exponentially discounted expected rewards are shown in blue. For hyperbolic discounting, $F(\tau) = 1 / (1 + \tau)$. For exponential discounting, $F(\tau) = \exp(-\tau)$. Also plotted are experimental data from [28] (filled circles; vertical bars indicate ± 1 std.). The dashed green line indicates predictions of the maximum probability of success model under a noise model fitted to generate accurate predictions for a 40° saccade ($\kappa = .0055$).

Figure 2-1B (top subplot) shows the probability of success for a saccade of 40° amplitude. Under the maximum probability of success hypothesis, a 40° saccade has duration of about 200ms, a value that far exceeds that of observed data (around 135ms). As the saccade amplitude increases further, the optimal duration begins to increase at an increasing rate and rapidly becomes unrealistic. Therefore, whereas maximizing probability of success produces durations that match observed data for small amplitudes, this policy grossly over-estimates saccade durations for large amplitudes (Figure 2-1C). Increasing saccade duration beyond 200-300 ms carries relatively little additional cost from signal-independent noise, but continues to reduce the impact of signal-dependent noise – particularly for large

amplitude saccades. This characteristic is independent of the particular value of κ used in our simulations. The dashed line in Figure 2-1C shows the predicted durations given a value of $\kappa = 0.0055$, which was chosen so that maximizing the probability of reward of a 40° saccade yielded a correct prediction of 135ms. Even for this optimized parameter, large amplitude saccades are predicted to have unrealistically long durations. If instead we were to fix κ based on the duration of larger amplitude saccades, the predictions for lower amplitude saccades become unrealistically short. As a result, we find that saccade durations are inconsistent with a policy that maximizes probability of success.

2.3.1 Temporal discounting of reward

Consider the possibility that the value of the stimulus is not constant as a function of time, but is discounted. As a result, the expected value of the stimulus at movement completion depends on two factors (Eq. 2.4): probability of success, and a temporal discount function $F(\tau)$ which describes change in stimulus value during the movement. Let us show that a policy that maximizes the expected discounted value of reward reproduces the observed amplitude-duration relationship.

Suppose that stimulus value is discounted hyperbolically $F(\tau) = 1/(\beta\tau + 1)$. We set $\beta = 1s^{-1}$ (for reasons we explain below). Under our assumed noise characteristics, the optimum duration for a 10° saccade is around 54ms and for a 40° saccade, the same parameters produce an optimum saccade duration of 135ms – both offering a good agreement with observed saccade durations. Indeed, a single hyperbolic discount function can accurately reproduce saccade durations for the entire range of amplitudes of recorded data (Figure 2-1). We also considered an exponential discount function, which has general form $F(\tau) = \exp(-\lambda\tau)$. We set $\lambda = 1s^{-1}$ so that the hyperbolic and exponential discount functions would share the same gradient at $\tau = 0s$. This exponential discount model also produces realistic duration predictions for the entire range of saccade amplitudes (Figure 2-1). Therefore, saccade durations are consistent with a policy that maximizes the expected discounted value of the stimulus (Eq. 2.4).

The specific value of $\kappa = 0.0066$ for these simulations was chosen such that a 30° saccade would have an optimal duration of 105ms under a hyperbolic discounting of reward model. This value of κ is approximately consistent with the magnitude of signal-dependent variability reported by [4]. However, given the sensitivity of our predicted durations to the details of our underlying model of saccade generation (plant properties and selection of motor commands for a given duration), we cannot be certain about the precise relationship between the discount function and saccade durations. Furthermore, this data does not allow us to dissociate between hyperbolic and exponential forms of discounting. We can, however, reject the possibility that saccade durations are selected to maximize expected

undiscounted reward, since no single value of κ could account for saccade durations across all amplitudes. Therefore, at this point we can only conclude that temporal discounting of reward plays a role in determining movement durations. The exact shape of the discount function remains unclear.

2.3.2 Temporal discounting as reward rate optimization

The results that we have presented thus far are similar to those that we saw in an earlier set of simulations [23]. In that work we assumed a cost in which movement endpoint errors were penalized with a quadratic function and movement duration was penalized through an added hyperbolic time cost. Here, we instead adopt a more natural framework in which successful movements (endpoint falling within a specified goal region) earn a positive reward and unsuccessful movements earn zero reward, regardless of the magnitude of the error. The value of reward associated with a successful movement is then discounted multiplicatively as a function of time. In addition, whereas our previous work assumed an oculomotor plant with only signal-dependent noise, we now consider a more accurate model of the oculomotor plant with both signal-dependent and signal-independent noise sources. One may argue over the relative merits of each model, but the fact is that both the current and the previous work suffer from two fundamental concerns: 1) We have merely shown that observed data on saccade durations are consistent with our temporal discounting framework. However, there may be many kinds of costs that are also consistent with this data. 2) In decision making, reward is temporally devalued over timescales of minutes, days, or years. In our model of saccades, reward is discounted over a timescale of milliseconds. It seems improbable that a few milliseconds should produce any meaningful change in the perceived value of reward. To address these concerns, we must first understand the deeper question of why the brain should discount reward at all.

A common interpretation of temporal discounting is that the risk of not getting a predicted reward increases with delay. If reward remains available for a duration that follows an exponential distribution (and reward disappearance behaves like a Poisson process), then exponential discounting maximizes the total expected reward. Such a framework can also account for hyperbolic discounting if the exponential distribution is replaced with a mixture of exponential distributions with different time constants [29]. An alternative interpretation is that temporal discounting reflects a desire to maximize the rate of reward acquisition, rather than the absolute value of each acquired reward [30] [29] [31]. To explain this, consider a choice between an immediate reward with magnitude α_1 , and a larger reward α_2 at some time in the future τ . If we assume that the next such decision will not occur immediately after we receive the reward, but after some average period of time δ (due to reaction time, inter-trial interval, etc.), then the reward rates associated with each choice are

$$R_1 = \frac{\alpha_1}{\delta} \quad R_2 = \frac{\alpha_2}{\tau + \delta}. \quad (2.12)$$

Suppose we vary α_2 and find the value for which we select the immediate but smaller reward α_1 at 50% probability. According to the rate of reward theory, this indifference between the immediate but smaller reward and the delayed but larger reward is occurring because $R_1 = R_2$. This condition occurs when

$$\alpha_1 = \frac{\alpha_2}{1 + \delta^{-1}\tau}. \quad (2.13)$$

We see that if we make choices in such a way as to maximize the rate of reward, then effectively we discount the value of the delayed reward α_2 hyperbolically with a rate that depends on the average duration δ between opportunities to earn reward. The key new idea that emerges is that hyperbolic temporal discounting arises because the underlying objective of the brain is to optimize the rate of reward R (i.e. reward per unit of time):

$$E[R] = aP[\text{success}|\tau] \frac{1}{\tau + \delta} \quad (2.14)$$

where δ is the inter-movement interval, and τ is movement duration. This is proportional to the expected hyperbolically discounted reward (Eqs. 2.3 and 2.4) when $\beta = \delta^{-1}$. In the simulations presented in Figure 2-1, we deliberately set β equal to $1s^{-1}$, corresponding to an inter-movement interval of 1s, roughly consistent with the experimental paradigm for the study [28] that collected the data in Figure 2-1C (although exact ITI data were not reported in that paper). That is, the specific hyperbolic temporal discount used to produce the simulations in Figure 2-1C is equivalent to Eq. (2.14) in which ITI is around 1sec. This establishes the plausibility that the rate of reward hypothesis could in principle account for vigor of saccades.

2.3.3 Predictions of the rate of reward theory

While the above findings establish the plausibility of the rate of reward hypothesis, a far stronger prediction of this theory is that if we change the inter-saccade interval δ , the brain will change the vigor of saccades. In a typical experiment one gives a sequence of targets, and the subject makes a sequence of movements to these targets. Eq. (2.14) predicts that the expected reward rate will depend on the average duration of each movement τ plus the average inter-movement interval δ . If we change δ , for example by increasing the time between the end of one movement and presentation of the target for the next movement, then the vigor with which that movement is performed should change. Here is the critical prediction of Eq. (2.14): an increase in inter-movement intervals should reduce movement vigor (produce

slower movements), whereas a decrease in inter-movement intervals should increase movement vigor (see also Results, Sensitivity to characteristics of the discount function).

To illustrate the predictions of Eq. (2.14) we performed a simulation to determine how much saccade peak velocities and durations should change as we alter the inter-trial interval (ITI). We found that with respect to ITI of 1 sec, reducing the ITI predicted increased peak velocities and increasing the ITI predicted decreased peak velocities (Figure 2-2A). Importantly, the effect was asymmetric: a 0.5sec decrease in ITI predicted a much greater change in peak velocities than a 0.5sec increase in ITI. Furthermore, the effect of ITI on peak velocities grew with saccade amplitude, but tended to saturate at around 40°. Similarly, reducing the ITI predicted decreased saccade durations and increasing ITI predicted increased saccade durations (Figure 2-2B). Put simply, if the brain is producing motor commands to maximize rate of reward (Eq. 2.14), then reducing the inter-movement intervals should increase movement vigor.

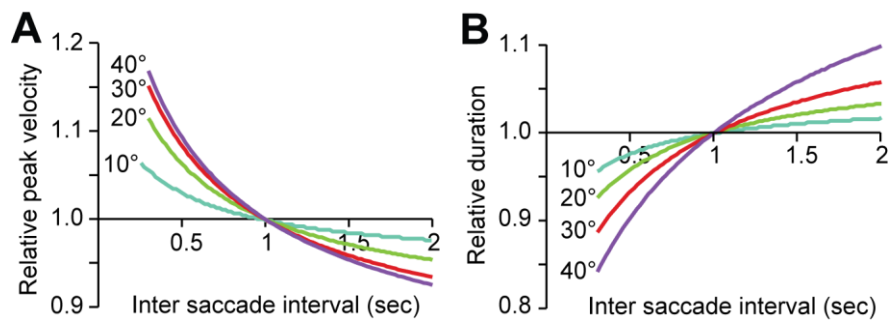


Figure 2-2: Changes in saccade vigor associated with changes in inter-movement interval, as predicted by a rate of reward cost function. (A) Changes in peak velocity. For each ITI, and each amplitude, we computed the saccade duration that maximized rate of reward. We then computed the peak velocity of that saccade and normalized it with respect to the peak velocity at ITI of 1 second. The simulations show how much the peak velocity should increase (or decrease) as a function of ITI between saccades. The effect is greatest for saccade size of around 40°. (B) Changes in saccade duration as a function of changes in ITI.

2.3.4 Change in ITI alters saccade vigor

We performed an experiment to test the prediction that changes in ITI should produce changes in saccade vigor. Subjects made alternate leftward and rightward saccades of 40° amplitude (Figure 2-3A). We employed three block types, each consisting of 80 trials in which ITI increased, decreased, or remained constant (Figure 2-3B). Each block began and ended with 10 trials having an ITI of 1sec. Figure 2-3C and E show our theoretical predictions regarding saccadic vigor. As ITI is reduced from 1sec to 0.4sec, peak velocities (for a 40° saccade) should increase by about 10%. As ITI is increased from 1sec to 1.6sec, peak velocities should decrease by about 5%. The experimental results are shown in Figure 2-3D and F. Over the first 10 saccades, the ITI was the same (1sec) in all block types, and the saccade peak

velocities did not differ significantly across blocks. For the constant ITI block (black line) there was no clear change in peak velocity other than a trend for the peak velocity to decrease - a ‘fatigue-like’ effect thought to be associated with stimulus devaluation due to repetition of the stimulus [32]. For the increasing ITI block (blue line), the abrupt decrease in ITI from 1 to 0.4sec on trial 10 (bin 5) was accompanied by a sharp increase in peak velocity. Over the next 60 trials the ITI was varied linearly from 0.4sec up to 1.6sec. This was associated with a steady decrease in peak velocity. In the last 10 trials of the block, after the ITI was decreased abruptly from 1.6sec back to 1sec, the peak velocity began to increase again, becoming similar to the constant ITI block. Saccade peak velocities in the decreasing ITI block (red line) showed the opposite trend.

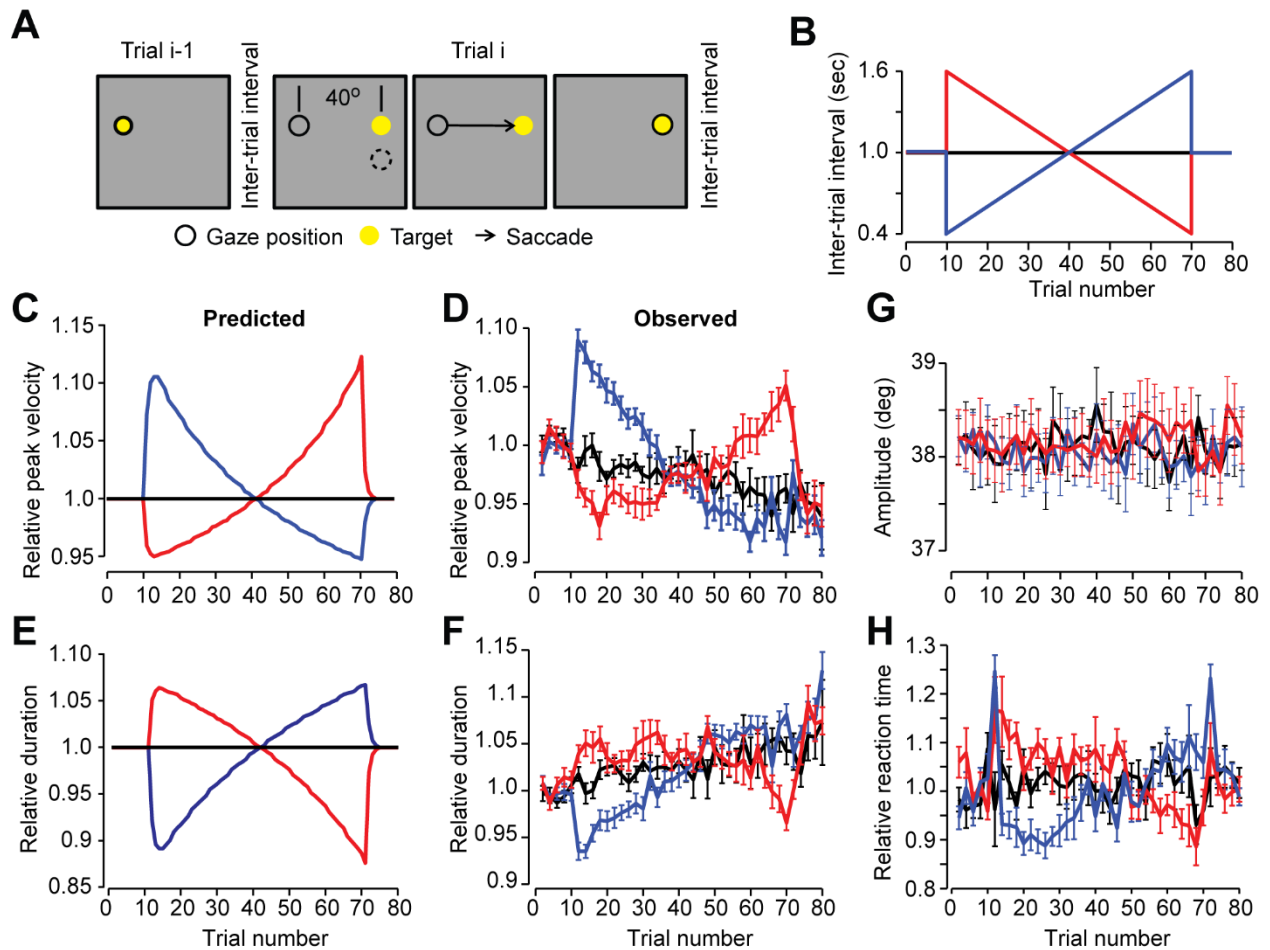


Figure 2-3: Experimental protocol, model predictions, and results. (A) Experimental protocol. Subjects were asked to make alternate leftward and rightward saccades to one of 2 possible targets. (B) The duration between the end of one saccade and displaying the ‘go’ cue for the next saccade (the inter trial interval, ITI) was varied through the course of each 80 trial block. Three possible block types: increasing ITI (blue), decreasing ITI (red), and constant ITI (black). (C) Predictions of the rate of reward model regarding changes in peak velocity of saccades. (D) Changes in peak velocity with respect to the first 10 saccades for each block type. (E) Predictions of the rate of reward model regarding changes in saccade durations. (F) Changes in saccade duration. (G) Saccade amplitudes. (H) Changes in reaction times. Error bars indicate SEM.

Analysis of covariance on each block confirmed that the changes in peak velocity followed significantly different trends across block types (ANCOVA, BLOCK×TRIAL interaction, $F(2,1074)=296.6$, $p<10^{-10}$). Similarly, we saw a significant effect of ITI on movement duration (BLOCK×TRIAL interaction $F(2,1074)=111.2$, $p<10^{-10}$) (Figure 2-3F). Post-hoc comparisons of the saccades following the initial abrupt ITI change and control block showed that the effects were significant (paired t-test, velocity: ITI increase vs. control, $p=0.005$; ITI decrease vs. control, $p<0.01$; duration: ITI increase vs. control, $p=0.007$, ITI decrease vs. control, $p=0.018$). However, manipulation of ITI did not affect saccade amplitude (BLOCK×TRIAL interaction $F(2,1074)=2.71$, $p>0.05$), as shown in Figure 2-3G. In addition, we observed significant changes in reaction time (BLOCK×TRIAL interaction $F(2,1074)=50.9$, $p<10^{-10}$), with reaction times becoming shorter as ITI decreased (Figure 2-3H). Therefore, reductions in ITI generally produced saccades with faster velocities, shorter reaction times, and shorter durations.

Increases in the value of a visual stimulus results in saccades with shorter reaction time in both monkeys [33] [34] and humans [35]. We noted that whereas reaction time generally followed the same trends as velocity and duration, in one instance these measures could be dissociated. At the onset of the 11th trial (and the 71st trial) the ITI sharply changed, either increasing or decreasing. We observed an increase in velocity for ITI decrease, and a decrease in velocity for ITI increase (Figure 2-3D and F). In contrast, both the sudden increase and the sudden decrease in ITI produced an increased reaction time. If we view reaction time as a period in which the upcoming movement is planned, this result suggests that the sudden change in ITI resulted in significantly longer time to plan the upcoming movement. Following this increased planning period, there was either a sharp decline (in case of increased ITI) or a sharp increase (in case of reduce ITI) in the vigor of the upcoming saccade. This dissociation allows us to rule out the possibility that changes in movement vigor were directly caused by changes to the reaction time that affected the process of movement planning.

It is noteworthy that for the block type with an initial decrease in ITI, saccadic vigor sharply changed within two trials of this decrease (e.g., maximum saccade velocity was reached within two trials). However, the rate of change in saccade vigor following an increase in ITI was much less, with subjects reaching minimum velocity after 7-8 trials. If we view changes in ITI as change in reward rate, then an unexpected change in ITI is equivalent to a reward rate prediction error. The fact that it takes longer for vigor to decrease than increase may be attributable to differences in learning from positive and negative reward rate prediction errors, suggesting that in this task the brain learns more from positive prediction errors than negative prediction errors.

2.3.5 Control experiment

An alternate interpretation of our experimental results is that when we reduced ITI, we are reducing the time that we are allowing the subject to view the target (in the experimental setup of Figure 2-3A, the current target disappears when the new target is shown). Perhaps this reduced viewing time is influencing vigor by encouraging the subject to get to the new target earlier so that they can view it for a longer period of time before it disappears. To test for this, we performed a new experiment (Figure 2-4A). In this version of the task the target that the subject was viewing did not disappear when a new target was presented. Rather, the subject could choose to continue viewing the current target for as long as they wanted. In this way, the viewing time of the current target was chosen by the subject, and not by the experimenter. Remarkably, we still observed that changing ITI produced robust changes in saccade vigor (Figure 2-4B and C). The patterns of change in velocities and durations were essentially identical to that which we had recorded in the main experiment (velocity, ANCOVA, BLOCK×TRIAL interaction, $F(2,874)=301.1$, $p<10^{-10}$; duration, BLOCK×TRIAL interaction, $F(2,874)=106.7$, $p<10^{-10}$). We did also observe a significant effect of amplitude (BLOCK×TRIAL interaction, $F(2,894)=19.9$, $p<10^{-8}$). However the changes in amplitude were of the order of 1% and are not sufficient to account for the changes in peak velocity and duration we observe, which were an order of magnitude larger than would be predicted on the basis of observed changes in amplitude alone.

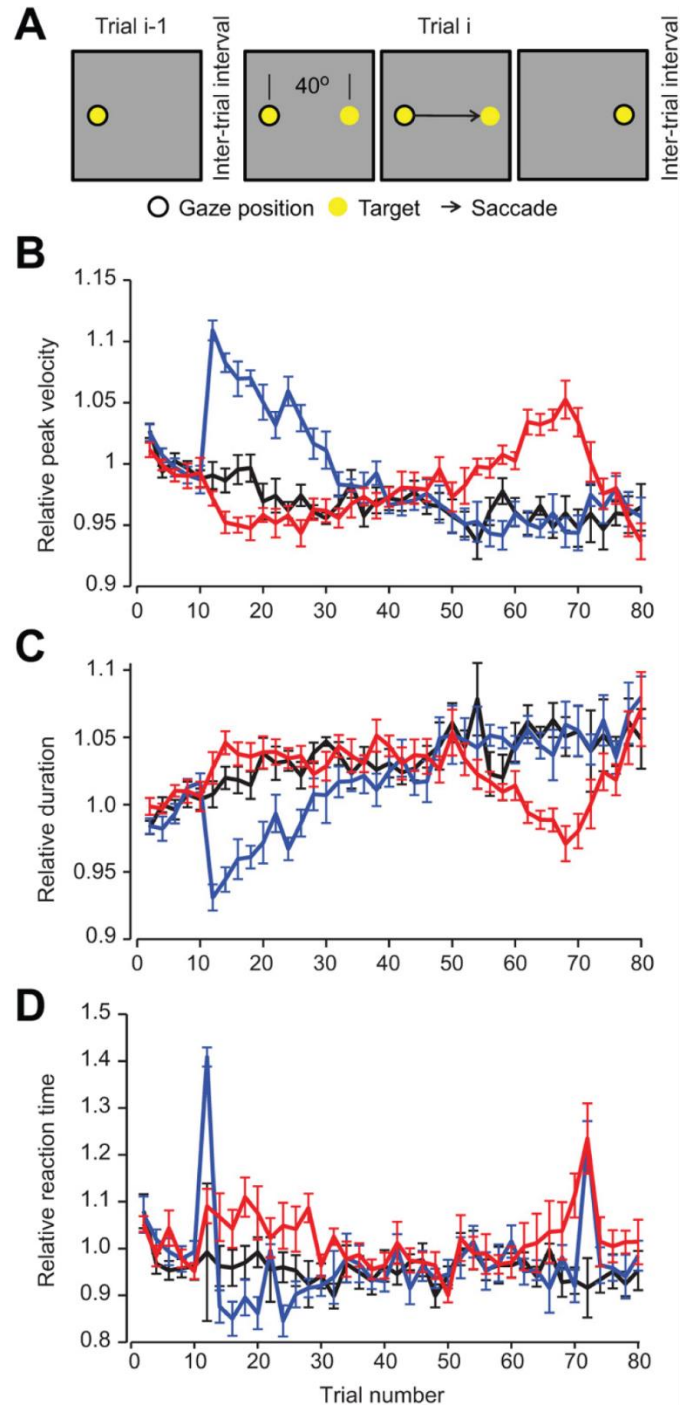


Figure 2-4: Control experiment. (A) Experimental protocol. This experiment was similar to that shown in Figure 2-3A, except that the previous target (current point of fixation) was not extinguished until after the saccade to the next target had begun. In this way, subjects could linger on the current target as long as they wanted. (B) Changes in peak velocity with respect to the first 10 saccades for each ITI block type (as in Figure 2-3B). (C) Changes in saccade duration. (D) Changes in reaction time. Error bars are SEM.

Interestingly, reaction time in this task was markedly higher than in our main experiment (mean RT = 213ms for the control experiment versus 158ms for the main experiment). Despite this marked difference in mean reaction time, the patterns of change (Figure 2-4D) due to changes in ITI were qualitatively similar to the main experiment: the sudden change in ITI on the 11th trial produced an increase in reaction time, regardless of whether ITI was decreased or increased. Following sudden change, gradually reducing the ITI produced longer reaction times and increasing the ITI produced shorter reaction times. Although there was no statistically significant effect of ITI on reaction time (BLOCK×TRIAL interaction, $F(2,874)=1.24$, $p=0.3$), this was largely caused by a single trial early in the block when an abrupt change in ITI caused unusually high reaction times. Overall, these observations are very similar to those in our main experiment. Therefore, the changes in vigor were unlikely to be due to subjects feeling rushed by the increased pace of the experiment.

2.3.6 Characteristics of the temporal discount function

We noted earlier that durations of saccades of different amplitudes could be accounted for by both hyperbolic and exponential temporal discount functions (Figure 2-1C). However, we found that saccade durations not only depend on saccade amplitude, but also on the time since the last saccade, i.e., ITI (Figure 2-3F). This experimental result confirms a prediction that we derived based on the premise of rate of reward, providing a rationale for hyperbolic temporal discounting (as in Eq. 2.14). However, let us now ask a more general question: in principle, what kinds of temporal discount functions could account for the data in Fig. 3? For example, could exponential discounting account for this data?

In general we can imagine discount functions in which ITI combines additively with movement duration:

$$E[\text{reward} | \tau, \delta] = aP[\text{success} | \tau] F(\tau + \delta). \quad (2.15)$$

Figure 2-5 illustrates the influence of ITI on movement duration under a variety of temporal discount functions. Suppose that for some class of movements the probability of success increases with movement duration (i.e., the slower the movement, the more accurate), as displayed in Figure 2-5A. If the objective is to maximize rate of reward, then time carries a specific cost in which the probability of success is multiplicatively penalized by a hyperbolic temporal discount function: $F(\tau + \delta) = \frac{1}{\tau + \delta}$. Increasing the inter-movement interval δ shifts the temporal discount function to the left, altering its slope and shifting the peak of the discounted reward function to long duration movements. As a result, for hyperbolic discounting, an increase in ITI reduces the vigor of movements.

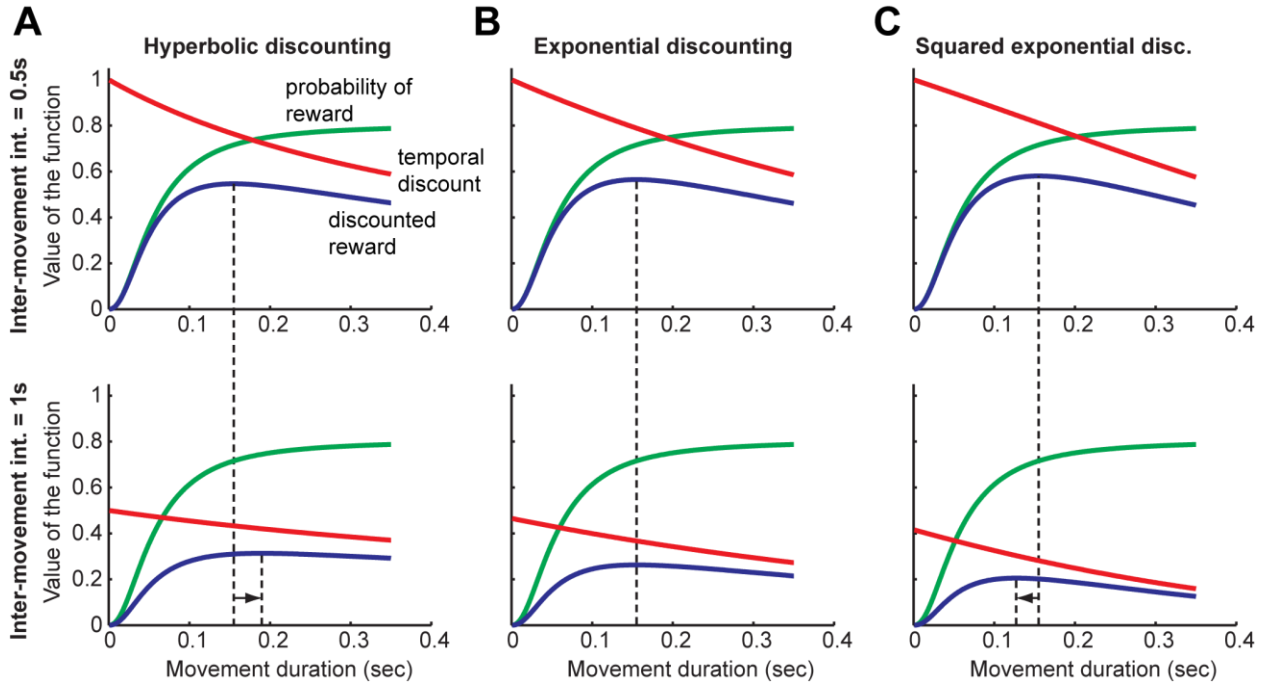


Figure 2-5: Effect of inter-movement interval on movement duration under various temporal discounting regimes. (A) Upper panel: We consider an arbitrary class of movements for which probability of success (acquisition of reward) increases with movement duration (green line). Hyperbolic temporal discounting, plotted here by the red line, is the function $\frac{1}{\tau + \delta}$, τ is movement duration, and δ is inter-movement interval (here assumed to be 0.5 sec). The blue line is the multiplication of probability of reward with the temporal discount function (Eq. 2.1). The movement duration that maximizes the discounted reward is noted by the dashed line. In this case, the discounted reward corresponds to reward rate. Upper panels of (B) and (C) show the corresponding plots for an exponential discount function with linear exponents $\exp(-k(\tau + \delta))$, and an exponential discount function with squared exponents $\exp(-k(\tau + \delta)^2)$. All discount functions are scaled to be equal 1 at 0.5s and parameters for the exponential discount functions were adjusted to predict the same optimal movement duration as rate of reward for $\delta = 0.5$ sec. Lower panels show the effect of increasing the inter-movement interval δ to 1s. For hyperbolic discounting, as this delay is increased the optimum movement duration becomes longer, i.e. the movement vigor decreases. For an exponential temporal discount function with linear exponents there is no change in the optimum movement duration as inter-movement intervals are changed. For an exponential discount function with quadratic exponents, movement duration decreases as the inter-movement interval increases.

Now instead consider exponential discounting: $F(\tau + \delta) = \exp(-k(\tau + \delta))$. If reward is discounted exponentially, changing δ has no effect on the optimal duration because it simply leads to an overall scaling of the expected discounted reward (Figure 2-5B). Therefore, the fact that we observed changes in saccadic vigor due to changes in ITI rejects the hypothesis that reward is discounted exponentially.

There are of course other plausible forms of temporal discounting, such as exponentials with squared exponents $F(\tau + \delta) = \exp(-k(\tau + \delta)^2)$. These forms imply that the cost of time is fairly

constant for short durations, but durations that are longer carry increasingly greater cost. Such forms can also be dissociated from hyperbolic discounting, as they predict a sensitivity to ITI opposite that of hyperbolic discounting. In this case, an increase in ITI leads to an increase in movement vigor (Figure 2-5C and F), which is inconsistent with our experimental results.

Therefore, the fact that we observed reduced saccadic vigor with increased ITI implies that temporal discounting has a specific shape. What is the class of functions that in principle could account for our data? The objective function to be maximized is

$$J(\tau) = P(\text{success}|\tau)F(\tau + \delta). \quad (2.16)$$

Suppose that, for a given value of δ , the optimal movement duration is τ_1 . This implies that the gradient of J at τ_1 is zero, i.e.

$$\left. \frac{\partial J}{\partial \tau} \right|_{\tau_1, \delta} = J' = P'F + PF' = 0. \quad (2.17)$$

Suppose we now increase δ to some new value. The resulting change in the gradient J' at τ_1 reveals how the optimal duration will change. If the gradient becomes positive, this means that the peak of J must have shifted to a larger value of τ . Likewise, a negative gradient implies a decrease in the optimal duration. In other words, the optimum movement duration will increase with inter-movement interval if the gradient of J' with respect to δ is positive:

$$\frac{\partial J'}{\partial \delta} = P'F' + PF'' > 0 \quad (2.18)$$

We arrived at Eq. (2.18) by noting that $\frac{\partial^2 F}{\partial \tau \partial \delta} \equiv \frac{\partial^2 F}{\partial \tau^2}$ because τ and δ appear additively in F , and that

$\frac{\partial P}{\partial \delta} = 0$ (because the probability of success for a given movement duration is independent of the inter-trial interval). We can eliminate P from Eq. (2.18) by dividing through by P , which is strictly positive and substituting $\frac{P'}{P} = -\frac{F'}{F}$ (which follows from Eq. (2.17)). If we further multiply through by F (which is also strictly positive), we obtain the following condition:

$$FF'' > (F')^2. \quad (19)$$

Importantly, this condition on the discount function is independent of the probability of success P and thus independent of the particular plant model and control policy we assume (since these will affect P but not F). Any temporal discount function F that satisfies Eq. (2.19) will lead to the prediction that

movement duration will increase (i.e., movement vigor will decrease) as inter-movement interval increases. For example, hyperbolic temporal discount functions (Eq. 2.14) have the property of Eq. (2.19), as do sums of exponentials. For exponential discounting with linear exponents (Eq. 2.15) the left and right hand sides of Eq. (2.19) are equal. For exponential discounting with quadratic exponents (Figure 2-5C), the inequality is reversed.

In summary, our experimental observations imply that control of saccades relies on a temporal discount function that satisfies Eq. (2.19). The dependence of saccade vigor on inter-trial interval emerges naturally in this mathematical framework and is unexplained by any previous model of which we are aware.

2.4 Discussion

We have proposed a new framework for control of movements. Our theory is founded on the principle of rate of reward – an idea that was previously invoked to explain aspects of decision-making in primates [36] and response intensities of rodents in free-operant tasks [31]. We have shown that a rate of reward principle not only provides an explanation for vigor of saccades of varying amplitudes but also generates the novel prediction that vigor should be modulated by changes in ITI. This is in contrast to previous models which assume that motor commands that guide a movement are independent of movements that occurred previously. Our experiments confirm our predictions with remarkable precision: as the ITI changed, so did saccadic vigor. For example, we observed an increase in saccade peak velocity of $9.0 \pm 0.9\%$ (mean \pm SD), compared to our theoretical prediction of 10%, and a decrease of $-7.9 \pm 1.1\%$, compared to our theoretical prediction of -5%. Changes in durations were in the predicted direction but somewhat smaller in magnitude, particularly for decreases in ITI. We are unaware of any previous motor control model that can explain such changes.

Our idea that time carries a cost in control of movements may explain a number of curious findings: 1) when a stimulus moves toward the fovea, saccades take longer to initiate than when the stimulus is moving away from the fovea [37]. Thus, subjects are not willing to wait for the stimulus to reach the fovea, highlight the idea that waiting even a few hundred milliseconds carries a cost. 2) An effective way to train monkeys to slow down their reaching movements is to impose a time penalty for overly fast movements [38]. The effectiveness of such a training protocol clearly illustrates the importance of the cost of time and is easily explained through a rate of reward framework, but is difficult to reconcile with models in which only the duration of the current movement is important. 3) The idea that changes in the available rate of reward can affect movement vigor is supported by a study by Ljungberg et al. [39] who reported that in non-human primates, reaching movements made to collect a

food reward were significantly slower when such rewards were available infrequently compared to when they were available frequently. This result could be considered analogous to our finding that increasing ITI decreases the vigor of saccades.

A fundamental question in neuroscience is why movements have characteristic kinematics. Why not move faster or slower? To approach this problem, previous works have sought to minimize weighted sums of rather ad hoc penalties for accuracy (quadratic distance to target) or time (linear or hyperbolic). We have suggested a different approach here: the decision regarding vigor of a movement depends on the probability of success of that movement multiplied by a function that represents temporal discounting of the value of the reward associated with that movement. These two approaches can in fact be related by applying a logarithmic transformation to the discounted reward (Eq. 2.14). This gives rise to an additive cost function that closely resembles those used in previous theories [23], but with the quadratic accuracy term replaced by $-\log(P[\text{success}|\tau])$ and the time penalty term replaced by the logarithm of the discount function. Notably, if the discount function is exponential, the logarithmic transformation gives rise to a linear time cost. Thus we can interpret previous models that have employed linear time costs [40] as tacitly assuming an exponential discounting of reward.

One aspect of the data which we could not explain through our model is the fact that saccade vigor tends to decline over the course of a block, even if the inter-trial interval remains constant. This decline in vigor is not due to muscular fatigue [41], but instead is likely due to the fact that repetition devalues the stimulus [6]. In monkeys, saccades to targets that predict a juice reward are significantly more vigorous than unrewarded targets [5]. Saccades that accompany reaching to a stimulus are also more vigorous than saccades without the reach [42] [43]. An increased probability of reward also increases the vigor of wrist movements made by monkeys to acquire reward [44]. Our previous formulation [23] was able to explain this dependence of movement vigor on reward value by including an effort penalty that was independent of reward value and a time penalty that scaled with reward value. In our present formulation, we did not include such an effort cost as we found that it was not necessary to generate strong predictions about behavior in the tasks we considered. Expanding our framework to include an effort penalty in the net reward, prior to applying the temporal discount factor, could enable us to explain these aspects of behavior through the rate of reward framework.

In addition to changes in movement speed, we also observed changes in reaction time. Although a low reaction time can be associated with a general increase in vigor, a quantitative prediction of this effect is beyond the scope of our model. One way in which we might be able to account for such an effect would be to view reaction time as a period in which a decision must be made about the goal of the

upcoming saccade. Computational models have described such decision-making in terms of a stochastic accumulation of evidence until a pre-determined threshold is reached, at which point the decision is made and an action triggered. Modulation of reaction time could therefore be interpreted as a change in the height of this threshold. It has been suggested that the level of the threshold might itself be set based on a rate of reward principle [36]. Indeed, experiments directly analogous to our own that vary ITI during perceptual discrimination tasks find that decreases in ITI lead to faster, less accurate decisions [45] [46]. Similarly, decreasing ITI causes monkeys to adopt a more risky policy when gambling for a juice reward [47].

In reinforcement learning, future rewards are typically discounted exponentially. Infinite horizon control problems, however, are commonly formulated in terms of minimizing the average-cost per stage, exactly analogous to the rate of reward cost function we proposed here. The average cost per stage framework has been invoked to model behavior when animals face a sequence of choices between discrete actions [29]. This idea has even been extended to include a basic notion of movement vigor, thereby offering an explanation for differences in the intensity of a rat's free-operant responses (such as lever press frequency) across different motivational states [31]. That work posited a conceptual link between dopamine, rate of reward, and response vigor through the idea that tonic activity of dopamine neurons encodes the background average rate of reward. When rewards are plentiful, it is worthwhile choosing the more costly (either energetically or in terms of risk) policy and acting more vigorously. In Parkinson's disease, reduced tonic dopamine leads to less vigorous actions [48], consistent with the idea that tonic dopamine encodes a rate of reward. Phasic activity of dopamine neurons may also be linked to rate of reward. Stimuli that predict reward at various delays elicit phasic responses in dopamine neurons that hyperbolically decline with the delay duration [22]. Phasic dopamine activity may thus reflect a reward rate prediction error, rather than an error in total predicted reward.

If we view ITI as a factor that alters rate of reward, then a sudden change in ITI introduces a reward rate prediction error. When ITI is reduced, the prediction error is positive, implying that the brain is receiving greater amount of reward than anticipated. Similarly, when ITI is increased, the prediction error is negative. We consistently observed that the change in vigor was faster when the prediction error was positive as compared to negative. This implies that there may be a differential sensitivity to positive and negative reward prediction errors in the population we sampled – an idea that is consistent with basal ganglia neurophysiology [49]. Viewed in this way, our experiment may provide a way to assess reward dependent learning – namely, monitoring changes in behavior in response to altered rate of reward.

The fact that a single optimization principle seems to be shared by such a broad variety of tasks suggests that it may offer a unifying normative view of temporal discounting in decision-making and

motor control. In effect, the long term behavioral goal of reward rate optimization is achieved through the short-term mechanism of temporal discounting. Hyperbolic discounting of reward may therefore be an obligatory phenomenon that has evolved because it tends to optimize reward rate in most ecologically relevant scenarios.

3 Modulation of saccade vigor during value-based decision making

3.1 Introduction

There is some evidence that the vigor with which a movement is performed (i.e., its peak speed as a function of amplitude) is affected by the subjective value that the brain assigns to the goal of the movement. For example, [50] asked participants to rank order in terms of preference a number of different kinds of candy bars. When asked to reach for a single candy bar, participants reached faster and with a shorter duration for the more preferred bar than the less preferred bar. Similarly, monkeys reached with a greater speed toward stimuli that promised higher probability of reward [44]. These observations raise the possibility that movement vigor may be modulated by the subjective value that the brain assigns to the goal of the movement. Humans and other primates use saccadic eye movements to examine their available options. During deliberation, as one makes saccades to accumulate information about the available options, does saccadic vigor reflect the subjective value that the brain currently assigns to each option?

Previous work has shown that stimulus value can grossly affect peak velocity of saccades. Monkeys exhibit a greater saccade peak velocity when the visual target is paired with food reward [5]. Humans perform their saccades with greater peak velocity if the target is a valued stimulus like a face [6]. Here, we considered a decision-making task where participants were offered monetary rewards. We asked whether the vigor with which a saccade was performed was affected by the subjective value that the brain assigned to the potential reward. A critical component of our design was that the movement that we considered (saccade) had no bearing on the reward itself: that is, people were not rewarded for making saccades. Rather, the saccades were a mechanism with which the motor system acquired information for the purpose of decision-making.

Our participants completed a temporal discounting task in which they chose between a small, immediate reward (in dollars), and a large reward to be received in 30 days. People prefer rewards sooner rather than later, but individuals vary widely in how much they are willing to wait for delayed reward. We measured participants' eye movements as they considered their two choices. The protocol allowed us to examine the relationship between saccade vigor and several aspects of value-based decision-making, including timing of decision, choice, and subjective valuation of each option. We found that saccade vigor remained high during the deliberation process, but dropped right after the decision was made. Furthermore, as the deliberation process proceeded, saccade vigor dissociated the option that the participant would eventually choose over the one that they did not prefer.

3.2 Materials and methods

3.2.1 Participants

We recruited $n=60$ healthy participants from the NYU community with no known neurologic deficits (21.75 ± 3.01 years old, mean \pm sd, 35 females). All were naïve to the paradigm and purpose of the experiment. Each participant signed a written consent form approved by the New York University Committee on Activities Involving Human Subjects. Each was paid \$10/hr in cash for participating in the study, as well as additional compensation based on their decisions in the task, as described below.

3.2.2 Behavioral task

Subjects sat in a darkened room in front of a CRT monitor (36.5 x 27.5 cm, 1024 x 768 pixel, light grey background, frame rate = 120 Hz), head stabilized with the use of a chinrest. The screen was placed at a distance of 55 cm from the subject's eyes. An EyeLink 1000 (SR Research) infrared camera recording system recorded movements and pupil diameter of the right eye. Gaze position and pupil diameter were recorded at 250 Hz for all subjects, with the exception of two subjects recorded at 1000 Hz and one recorded at 500 Hz. A superset of the data from this study was also examined with regard to changes in pupil diameter. A report of those findings appeared in [51].

We measured eye movements during a temporal discounting task. The time-course of a typical trial is displayed in Figure 3-1A. The trial began with a one-second fixation period (dot displayed at center of screen). Right after the fixation period, written description of the two possible rewards appeared simultaneously on the screen: a text that described a small immediate monetary reward (for example "\$10 today"), and another text that described a larger delayed monetary reward (for example, "\$11 30 days"). Each text was centered at 10° to the left or right of center, and was 4.7° to 7.9° wide, and 5.4° tall (as shown in Figure 3-1C). The placement of the text on left or right was chosen at random. One option was always for an immediate reward, while the other option was always for a reward to be attained in 30 days.

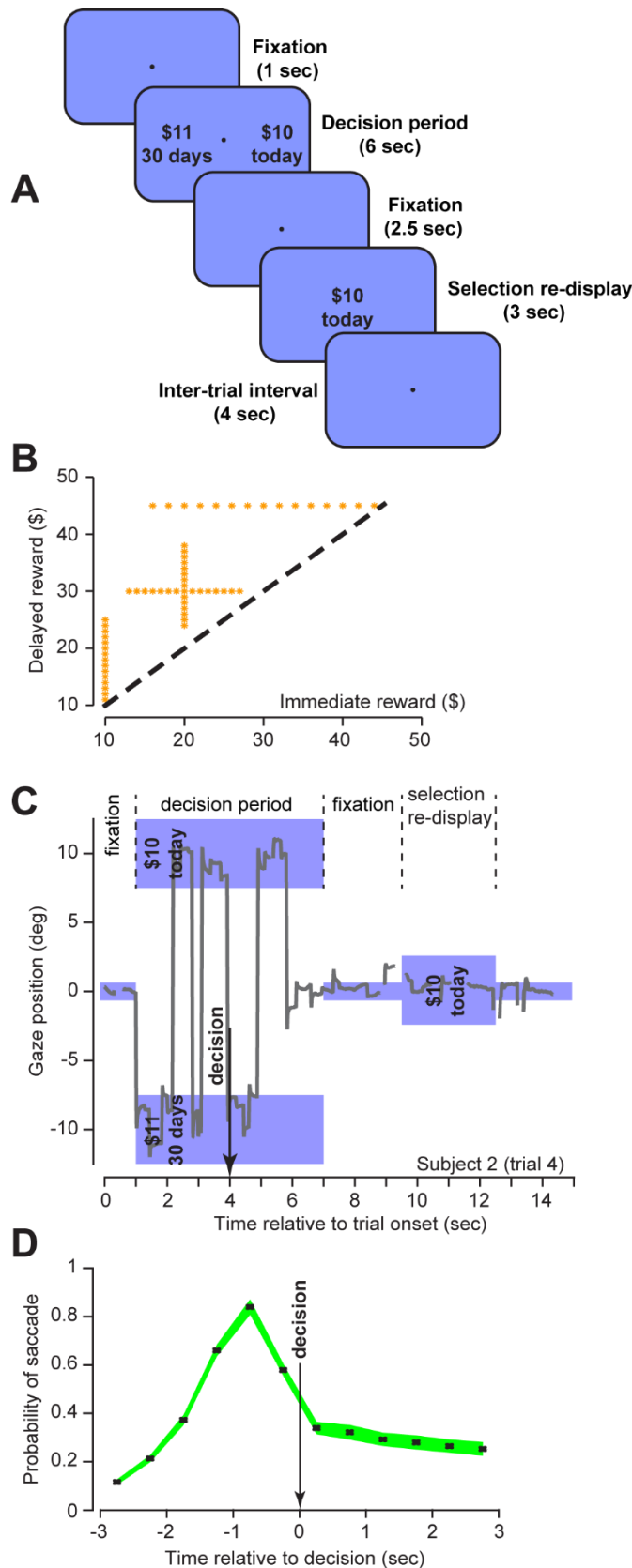


Figure 3-1: Experimental design. Participants completed a decision making task in which they made a choice between a small immediate monetary reward, and a larger delayed monetary reward, to be attained at 30 days. **A.** Each trial began with a fixation dot displayed for 1 sec, followed by display of the two reward options at either side of the fixation dot. Subjects were instructed to press a key designating their choice within 6 sec. Regardless of when the participants indicated their choice, the options remained on the screen for the full six seconds. **B.** The reward pairings. Each dot represents one of the 60 reward pairings that was presented to each participant. The delayed reward was always greater than the immediate reward. Each reward pairing was presented twice. The order of presentation was random for each participant as well as across participants. The position of presentation of each option was randomly selected to be centered at 10° to the left or right side of fixation. **C.** Exemplar trace of gaze position during a single trial. During the decision-period, the arrow indicates when the participant pressed the key designating her decision. Note that the participant continued to make saccades after decision-time, while both options remained on the screen. In this particular trial the participant chose the immediate reward (\$10) over the delayed option (\$11), and thus the \$10 option was re-displayed after the fixation period. **D.** Probability of saccade with respect to decision-time, computed for bins of 0.5sec in duration. Shaded region is SEM.

During this decision-period participants made saccades to the stimuli. The stimuli remained on the screen for exactly 6 seconds, during which time the subjects indicated their decision by pressing a key (typically at around 2-3 seconds into the trial). Subjects were instructed to place the left hand over the 1 key and right hand over the 0 key, in order to respond to leftward and rightward rewards, respectively. Regardless of when the subjects pressed the key to designate their decision, the two options remained on the screen for the full six seconds. This was critical as this allowed the subjects to make

saccades to the stimuli both before and after they made their decision.

After the completion of the 6 second decision-period, the fixation-dot reappeared for 2.5 seconds. Finally, the fixation-dot was removed and the participants were presented with the option that they had chosen for 3 seconds. A new trial commenced following an inter-trial interval of 4 sec. There were 120 trials in the experiment. Our analysis focused solely on the saccades made during the 6 second decision-period.

The participants were presented with 60 distinct monetary reward pairings, as shown in Figure 3-1B, with each pairing presented twice. The reward pairings were selected in random order such that no two subjects saw the same ordering of stimuli. On every trial the delayed monetary reward was of greater magnitude than the immediate reward. The delayed and immediate rewards could switch sides from one trial to the next.

To increase the relevance of their choices, participants were instructed that one trial would be selected at random and they would receive the amount that they chose on that trial. That is, if they chose the immediate reward on the randomly selected trial, they would receive the money in cash after completion of the session. If they chose the larger, delayed reward, they would receive a debit card that would be activated after the delay (30 days) had elapsed.

One participant did not complete all 120 trials and was excluded from analysis. In addition, we were unable to achieve good eye calibration in six participants, which prevented measurement of saccades for those subjects. In particular, these subjects did not fixate all targets during the calibration phase, precluding the possibility of a reliable regression of raw data to gaze position in degrees. As a result, we analyzed the data from a total of $n=53$ participants.

During each trial we continuously recorded gaze position. Raw gaze position signals were smoothed and differentiated with the use of a Savitzky-Golay filter (second-order). The filter width was chosen as a function of the sampling rate such that each filter window encompassed 20 ms of data. We used the gaze velocity trace to determine onset and offset of saccades, with a $30^\circ/\text{sec}$ threshold. We used the following five criteria to identify task-relevant saccades: (a) horizontal amplitude greater than 2° and less than 25° , (b) vertical amplitude less than 6° , with the ratio of vertical amplitude to horizontal amplitude less than 0.7, (c) peak horizontal acceleration less than $35,000^\circ/\text{sec}^2$, (d) skew (defined as the ratio of time from saccade start to peak velocity to saccade duration) less than 0.7, (e) duration greater than 20ms and less than 120ms. To identify an outlier saccade, we used the median absolute deviation technique [52]. The first criterion removed $45 \pm 10\%$ (mean \pm sd) of saccades (as many of the saccades were associated with the act of reading the text on the screen, a series of micro-saccades). The remaining criteria together excluded $29 \pm 9\%$ of the remaining saccades. To identify an outlier saccade, we used the

median absolute deviation technique (on the parameter saccade vigor), which excluded $2.7 \pm 1.5\%$ of the remaining saccades.

3.2.3 Data analysis: saccade vigor

During the decision-period, subjects made saccades that terminated at either at one of the stimuli, or at the center fixation point (as illustrated in Figure 3-1C). These saccades had a participant-specific velocity-amplitude relationship: some participants exhibited fast saccades, whereas others exhibited slow saccades [53]. Our hypothesis was that in a given individual, for a given saccade amplitude, the brain modulated saccade velocity as a function of reward or context [6]. To dissociate amplitude-dependent changes in velocity from reward-dependent changes in each individual, we first modeled the amplitude-dependent effects of saccade velocity for that individual, and then compared changes in velocity that were present when amplitude was kept constant but reward or context changed. The result was a within-subject measure of saccade vigor, as described below.

For each participant n we measured the amplitude of the saccade (represented by x), and its peak velocity (represented by v) in all trials. Previous work had shown that a hyperbolic function is generally a good fit to human saccade data [53]. We therefore fitted the data to the following function:

$$v_n = \alpha_n \left(1 - \frac{1}{1 + \beta_n x} \right) \quad (3.1)$$

We quantified the goodness of fit of the function for each participant using correlation coefficients. This fit produced parameter values $\hat{\alpha}_n$ and $\hat{\beta}_n$.

Given saccade amplitude x , the expected saccade velocity in subject n was represented by $\hat{v}_n(x)$. For each saccade we computed the ratio between the measured velocity and the expected velocity: v_n/\hat{v}_n . This ratio defined a within-subject measure of saccade vigor. When this ratio was greater than one, the saccade had a velocity that was larger than expected, reflecting a greater than average vigor for that subject. We used this within-subject measure of vigor to quantify changes in saccade peak velocity as a function of time during the decision-making period, and as a function of the preference that the subjects exhibited toward the available options in each trial.

3.2.4 Data analysis: decision-making

We analyzed the decisions that each participant made by finding the value of the delayed reward that made that option equivalent to the immediate reward. For each participant we represented the probability

of choosing the delayed reward r_d as a function of the difference in the value of the delayed and immediate rewards $r_d - r_i$:

$$\Pr(\text{choice} = r_d) = \frac{1}{1 + \exp(-b(r_d - r_i - a))} \quad (3.2)$$

In the above expression a represents the point of subjective equivalence between the delayed and immediate options. We fitted the above equation to the choices that the participant had made across all trials. To do so, we analyzed the trials based on the difference between the delayed and immediate rewards, and then measured the probability of choosing the delayed reward in each trial. Therefore, in a trial in which $r_d - r_i = \hat{a}$, the participant was equally likely to pick the delayed or the immediate option. Participants who preferred the immediate reward more often, and thus were more impulsive in their decision-making, had larger values of \hat{a} .

To estimate the subjective value of an option for participant n , we considered a hyperbolic model of temporal discounting [54] [55] [56]. In this model, one assumes that people evaluate a future reward (promised to arrive after time delay t) by discounting it hyperbolically to produce a subjective value at present:

$$r_i = \frac{r_d}{1 + k_n t} \quad (3.3)$$

In our experiment the time delay t was always 1 month. For each participant n we estimated discount factor k_n as a function of the mean ratio $\frac{r_d}{r_i}$ for all trials in which the absolute difference $|r_d - r_i|$ was within \$5 of equivalence point \hat{a} . To confirm this estimate, we also divided up the trials into four subsets (each vertical and horizontal line in Figure 3-1B), and then re-estimated k_n independently for each subset of trials in each participant. This way of estimating k_n kept either the immediate or the delayed reward constant for each subset of trials. We compared the two methods and found that the two estimates correlated very well ($r^2 = 0.96$, slope of 1.174, bias of -0.14). In our results we report the estimate arrived at using the entire data set.

All statistical analyses were performed using SPSS (IBM, v22), or Matlab R2014b. All t-tests presented are two-sided, unless otherwise specified.

3.3 Results

3.3.1 Saccade vigor was higher during the deliberation period

On each trial, the participants were presented with two options: a monetary reward to be acquired on the day of the study, and a larger reward to be acquired in 30 days. As the participants evaluated the two options and made their decision, they made saccades from one stimulus to another. On average, participants made 6.2 saccades per trial (4.6, 25th percentile; 7.7, 75th percentile), and on average they announced their decision at 1.93 ± 0.46 sec into the decision-period by pressing a key. However, regardless of when the decision was made (indicated by the key-press), the stimuli remained on the screen for 6 seconds. As a result, the participants made saccades to the stimuli both before and after their decision, as illustrated in Figure 3-1C. To compute probability of saccades during a trial, we aligned the data to decision-time and then counted number of saccades performed by a given subject in bins of 0.5sec in duration across all trials. For each bin of 0.5 sec duration we computed probability of saccade for that subject, and then computed the across-subject mean and SEM of that probability, as shown in Figure 3-1D. We found that probability of saccade reached its peak about 1 sec before decision-time, but was always significantly greater than zero during the entire decision-period (all p-values less than 10^{-9}).

For each subject we considered each saccade that they made during the decision-period and measured its amplitude and velocity (data for a typical subject are shown in Figure 3-2A). Inspection of the data suggested that saccades made before decision-time, i.e., during the period of deliberation before key-press, may have had a higher velocity than saccades made after (right panel, Figure 3-2A). To explore this question, we examined probability of saccade as a function of amplitude and found it to have four modes (Figure 3-2B), with peaks at $\pm 9^\circ$ and $\pm 18^\circ$ (the options were displayed at $\pm 10^\circ$ with respect to central dot). We focused our analysis on those saccades in which one of the stimuli was the goal of the saccade (i.e., center-out or stimulus-to-stimulus saccades), or the fixation spot (stimulus to center saccades). For each saccade, we computed peak velocity as a function of amplitude. The result for a typical participant is shown in Figure 3-2C, and the population average is shown in Figure 3-2D. A within-subject comparison demonstrated that peak velocity and amplitude were significantly higher before decision-time than after (Figure 3-2E, within-subject comparison, peak velocity, $t_{(52)}=9.49$, $p < 10^{-12}$, amplitude, $t_{(52)}=9.06$, $p < 10^{-11}$). Indeed, for 96% of the participants the average peak velocity of saccades was smaller in the post-decision period (Figure 3-2F).

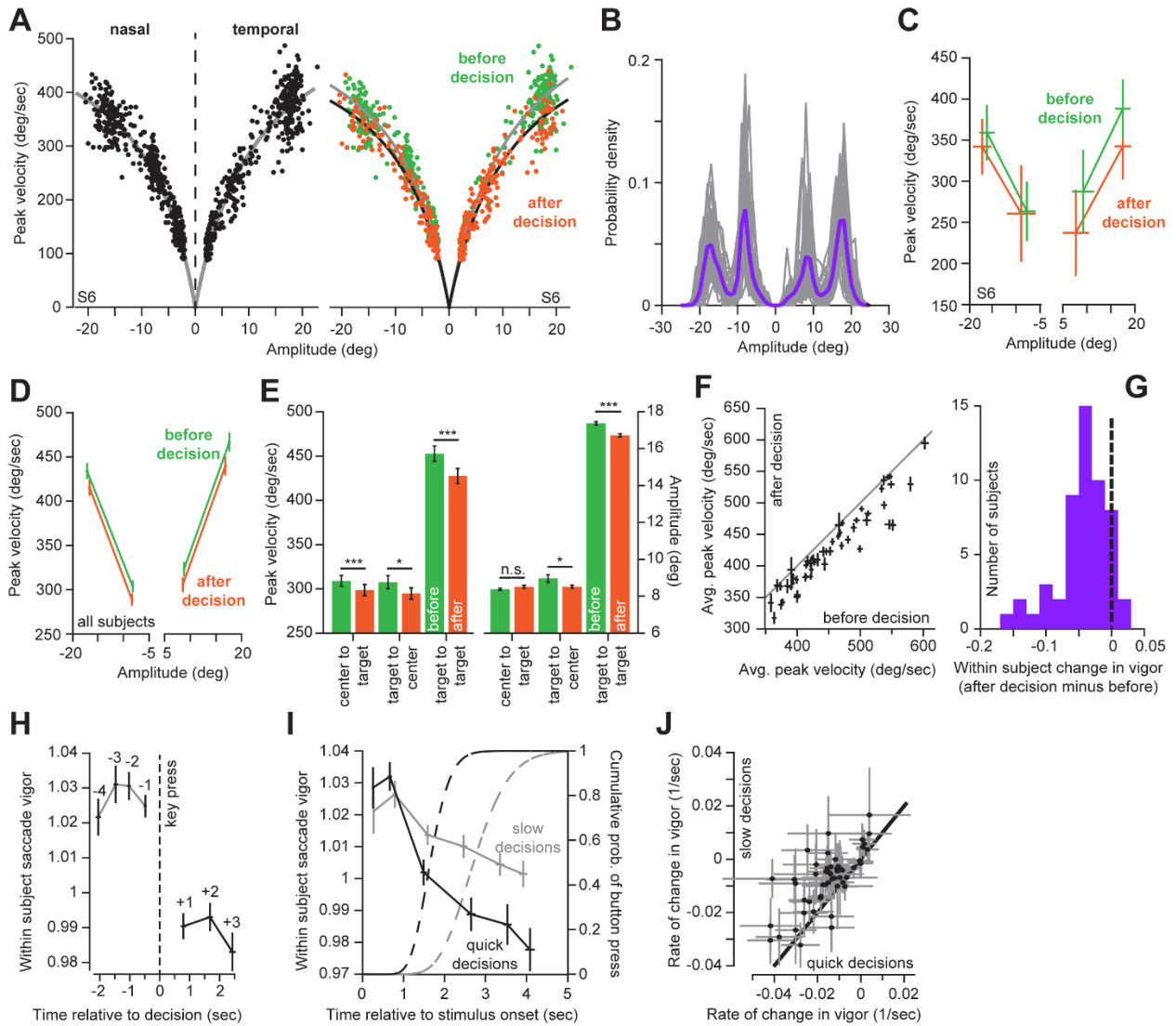


Figure 3-2: Saccade vigor was higher during deliberation than following decision-time. **A.** Velocity-amplitude relationship for a representative participant for saccades made during the decision-period. Data were fit to a hyperbolic function, separately for nasal and temporal saccades. Saccades made before decision-time (gray dots) appeared to have a greater velocity than those made after decision-time (black dots). **B.** Distribution of saccade amplitudes suggested that there were four groups of saccades made during the decision-period: from one stimulus to another ($\pm 18^\circ$ saccades), and from center to one stimulus or back ($\pm 9^\circ$ saccades). Gray lines represent probability density for each participant. Black line represents the across-subject values. Data were binned with step size of 1° . **C.** Velocity-amplitude data from an exemplar subject split by timing of saccade relative to key-press. The saccade amplitudes were binned with bin centers located at $\pm 9^\circ$ and $\pm 18^\circ$, with bin width of 9° . Error bars represent standard deviation. **D.** Across subject data. The error bars represent SEM, and are plotted for both amplitude and velocity. **E.** Across subject values of amplitude and velocity for saccades made before and after decision-time. Statistics refer to within-subject changes ($* = p < 0.05$; $*** = p < 0.001$). Error bars are SEM. **F.** Average saccade peak velocities for each subject before and after decision-time. **G.** Distribution of within subject change in vigor with respect to decision-time. **H.** Within subject measure of saccade vigor as a function of timing of saccade with respect to the key-press. The number on each data point represents saccade number with respect to key-press. Error bars along x- and y-axis are SEM. **I.** Within-subject measure of saccade vigor (solid lines) as a function of saccade timing with respect to start of decision-period (stimulus onset). The plot also shows cumulative probability of key-press (dashed lines) for slow-decision and quick-decision trials. Error bars are SEM. **J.** Rate of change in vigor for each participant in the quick-decision and slow-decision trials. Error bars are standard deviation.

Because saccade velocity is a function of amplitude, the critical question was whether the higher velocities observed during the deliberation period were due to greater vigor, or simply due to increased

amplitude. To answer this question, we accounted for the effect of amplitude on velocity by fitting a hyperbolic function (Eq. 3.1) to the velocity-amplitude data of all saccades made by each participant (left panel, Figure 3-2A), and then used the resulting fit to predict the expected saccade velocity at a given amplitude. The average r values of fits to nasal and temporal saccades were 0.94 ± 0.03 and 0.95 ± 0.03 (mean \pm sd).

For each saccade during the decision-period we measured its amplitude and computed the ratio of the measured velocity vs. the expected velocity. This ratio, our proxy for a within-subject measure of saccade vigor, indicated whether the peak velocity of a given saccade was higher or lower than the expected velocity for that amplitude. For each saccade we computed its vigor and then computed the average within-subject change in vigor from before decision-time to after. We found that a significant number of subjects showed a drop in vigor after decision-time (Figure 3-2G, $t_{(52)}=8.23$, $p < 10^{-10}$). Saccade vigor as a function of time relative to decision is plotted in Figure 3-2H, where we have numbered each saccade and plotted its timing with respect to key-press. There was an approximately 4% reduction in saccade vigor following decision-time (within-subject comparison, $t_{(52)}=5.97$, $p < 10^{-6}$).

In some trials the participants took a relatively long time to make a decision, whereas in other trials the decision was made quickly. For each participant we computed the median decision-time, and then labeled each trial for that participant as quick-decision or slow-decision (decision-times for quick and slow trials were 1.40 ± 0.35 , and 2.31 ± 0.62 sec, mean \pm sd). Figure 3-2H plots saccade vigor with respect to stimulus onset for the two types of trials. When we plotted saccade vigor with respect to the onset of the decision-period we found that in quick-decision trials saccade vigor declined rapidly, whereas in slow-decision trials saccade vigor declined gradually (Figure 3-2I). We tested this difference in vigor as a function of saccade index with a repeated measures ANOVA and found a significant group by saccade index interaction (Wilks' Lambda = 0.604, $F_{(5,46)} = 6.041$, $p < 10^{-3}$, as there were several subjects who never made more than 5 saccades to stimuli during deliberation, the number of degrees of freedom corresponding to saccade index was 4). Indeed, a within-subject analysis revealed that the rate of decline in vigor was significantly steeper in quick-decision trials than slow-decision trials (Figure 3-2J, within subject t-test, $t_{(52)}=6.41$, $p < 10^{-7}$).

In summary, we found that saccade vigor (as measured via velocity of saccades normalized by amplitude for each subject) was greater during the deliberation period (before the decision was made) than immediately after. Vigor dropped quickly in trials in which participants made a quick decision, but dropped slowly in trials in which they took longer to make their decision.

3.3.2 Saccade vigor encoded preference

On each trial the participants pressed a key to indicate which of the two options they preferred. We asked whether saccade vigor predicted this preference. We separated the saccades based on whether they were directed toward the preferred or the non-preferred stimulus, where the preferred stimulus was the option that was eventually chosen by the participant on that trial. Figure 3-3A plots vigor as a function of time of saccade, indexed with respect to key-press. It appeared that saccades made before decision-time did not differentiate between the preferred and non-preferred options, except for the last saccade just before key-press (Figure 3-3A). This final saccade took place at 0.520 ± 0.16 sec before decision-time (mean \pm sd), and had a higher vigor if it was directed to the preferred stimulus (within-subject difference in vigor between the preferred and non-preferred options, $t_{(52)}=3.31$, $p=0.0017$). Following the decision, the subsequent saccade also exhibited a greater vigor when it was directed to the preferred stimulus (within-subject difference in vigor, $t_{(52)}=2.40$, $p=0.020$). There was no difference in the vigor of saccades to preferred and non-preferred options outside of this window, suggesting that the encoding of choice preference was a phenomenon that affected vigor only near time of decision. In summary, whereas average vigor of saccades always dropped within-trial, there was an additional effect of choice preference around the time of decision.

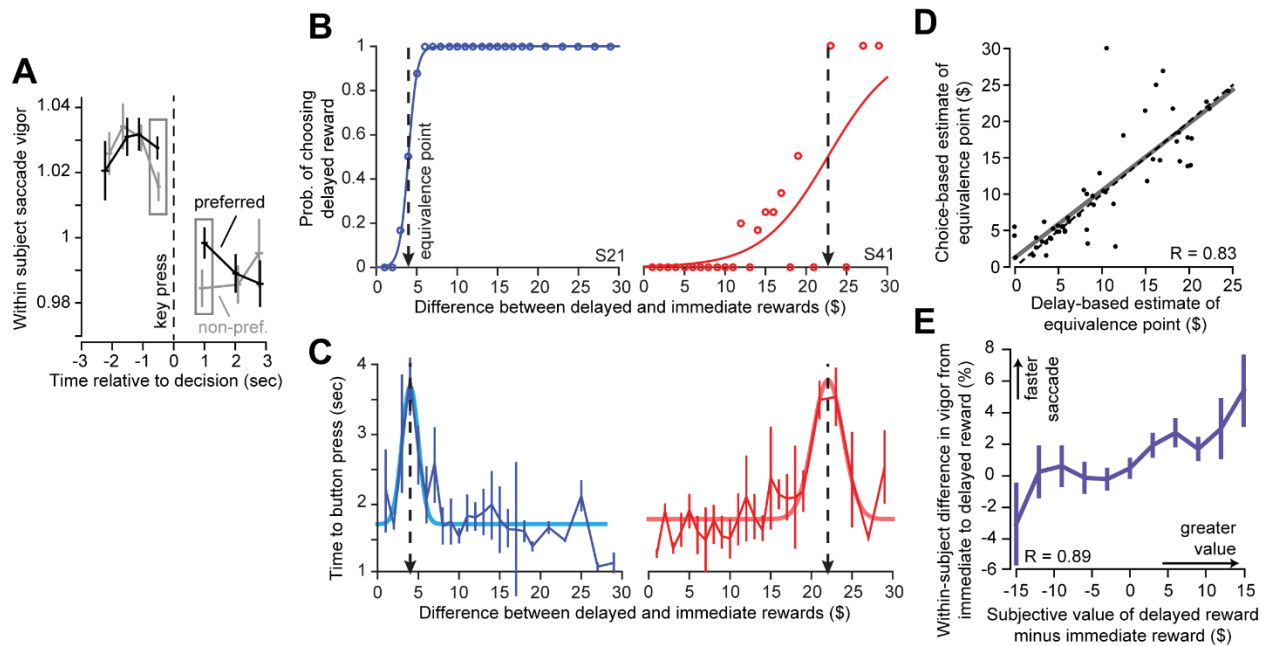


Figure 3-3: Saccade vigor encoded choice preference just before time of decision. **A.** For each participant we labeled each saccade based on whether gaze was directed toward the preferred or non-preferred option. We then indexed saccades relative to timing of key-press. Around the time of decision, vigor of saccade made to the preferred target was higher than vigor of saccade made to the non-preferred target. Gray boxes denote saccades immediately before and after key-press. **B.** Quantifying point of subjective equivalence via explicit decisions. The left and right columns show the probability of choosing the delayed reward option as a function of the difference in the rewards offered for Subjects 21 and 41. Subject 41 required a larger amount of delayed reward in order to switch preference from the immediate to delayed reward, and thus tended to favor the immediate

reward more than Subject 21. We fitted a two-parameter sigmoid function to the subjects' choice data and estimated the point at which subjects switched preference from the immediate to delayed reward, which we labeled the point of subjective equivalence (\$3.97 and \$22.71 for the two subjects). **C.** Delay to decision reflected difficulty of decision making. The plots show the distributions of time to key-press for the same two subjects. We computed the average time to key-press at each value of difference in reward (solid black vertical lines represent 25th and 75th percentiles), and then fitted a Gaussian function to the data (gray curve). The location of the mean of the Gaussian represents the difference in reward offerings for which the subjects took the longest time to make a decision (\$4.01 and \$22.02 for the two subjects). **D.** Quantifying robustness of estimate of equivalence point. For each participant we compared the equivalence point estimated from their explicit decisions (as in part B), to the equivalence point estimated from their time to key-press (as in part C). Each data point is one participant. The dashed line represents equality between the two measures. Gray line is the best linear fit. **E.** For saccades made immediately before and after decision, within-subject difference in saccade vigor was related to within-subject difference in subjective value of the delayed and immediate rewards. Error bars are SEM.

One may estimate the degree of preference for one option over the other via the difference in their subjective value. Is the difference in subjective value reflected in the difference in saccade vigor? To compute subjective value of a given option we analyzed the choices that the participants made. Figure 3-3B illustrates the choices made by two participants. Participant S21 (left subplot, Figure 3-3B) often picked the delayed reward when the dollar amount of that option exceeded that of the immediate option by more than \$5. In contrast, participant S41 picked the delayed reward only when the dollar amount of that option exceeded that of the immediate option by more than \$20. We fitted these data to Eq. 3.2, resulting in an estimate of the point of subjective equivalence for each participant (dashed line, Figure 3-3B). For participant S21, a difference of \$4 made the delayed reward equivalent to the immediate reward. For participant S41, a difference of \$23 was required to make the delayed reward equivalent to the immediate reward.

How robust was this estimate of subjective equivalence? To answer this question, we imagined that for each participant the decision should be most difficult when the two options differed in value by the amount specified by the point of subjective equivalence. For example, for participant S21 the most difficult choice should be in trials where the delayed reward was \$4 greater than the immediate reward. A proxy for this difficulty is the time that the participants needed to make their choice. We measured the time from stimulus display to key-press and have plotted the results in Figure 3-3C. For each participant we fitted their time to key-press with a Gaussian and estimated its center, resulting in the difference between delayed and immediate reward that produced the longest deliberation time. As a result, the explicit choices that participants made provided one measure of subjective equivalence (Figure 3-3B), and the time they took to make that choice provided a second measure (Figure 3-3C). The two measures were well correlated (Figure 3-3D, $r^2 = 0.68$, $p < 10^{-12}$). This result indicated that the point of subjective equivalence derived from the explicit choices was reasonable and robust.

We next used the decision-based estimate of subjective equivalence to compute the rate of temporal discounting (parameter k in Eq. 3.3), which then allowed us to compute the (relative) subjective

value of the delayed reward for each participant (assuming a linear utility function). Focusing on the two saccades made immediately before and after decision-time, we measured vigor when the participants looked at the immediate reward and compared it to vigor when they looked at the delayed reward. The difference in vigor is plotted on the y-axis in Figure 3-3E. Vigor increased from the immediate to the delayed reward as a function of the difference in the subjective value of the delayed reward vs. the immediate reward ($r = 0.89$, $p = 0.0002$). The magnitude of this increase was roughly 6% (Figure 3-3E). That is, around the time of decision, vigor of the saccade that placed a stimulus on the fovea was correlated with the subjective value that the brain assigned to that stimulus.

We found that the saccade made just before decision-time tended to be to the preferred option. In Figure 3-4A we have plotted probability that the saccade was to the preferred option, given that the participant made a saccade, computed over time bins of 0.5 seconds in duration. This conditional probability became significantly greater than chance around 1 second before decision time, and reached its peak at the final time-bin before decision-time (within-subject comparison, $p < 10^{-11}$). Furthermore, it appeared that as time passed following the decision the participants were more likely to saccade to the chosen option than the non-chosen option (post-decision region of Figure 3-4A).

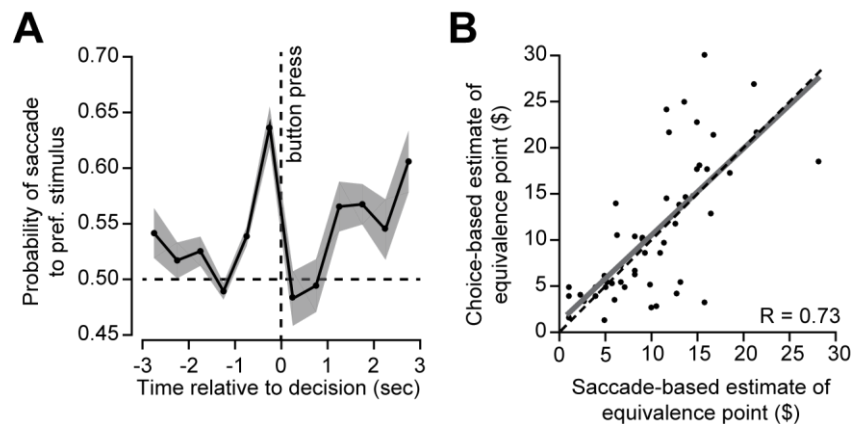


Figure 3-4: The stimulus that was target of gaze just before the key-press was often the option that was eventually chosen. **A.** Probability of the saccade target being the chosen option, given that a saccade was made to one of the two stimuli at that time bin. Time bins are 0.5 seconds in length. **B.** Quantifying robustness of estimate of equivalence point. For each participant we used the last saccade before decision-time as the predictor of the preferred option and used that result to compute an equivalence point (labeled as saccade-based estimate). Error bars are SEM.

These observations suggested that saccade patterns may be used as an implicit measure of preference. How well does this implicit measure predict the eventual choice? To check for this, we compared the choices that subjects made to the choices that would be expected if the saccade just before decision-time was used as a marker of preference. We computed an implicit equivalence point based on

the option that was the target of the last saccade before decision-time, and found that this implicit equivalence point matched well with the explicit equivalence point as estimated from the actual choices that the subjects made (Figure 3-4B, $r = 0.73$, $p < 10^{-9}$) [57] [58]. Thus the target of the final saccade before decision was an excellent predictor of the explicit choices that participants made.

In summary, during the deliberation period vigor of the saccades that placed each of the two stimuli on the fovea was similar, but diverged at around 0.5 seconds before decision-time, becoming larger for the preferred stimulus. As the difference between the subjective values of the delayed and immediate rewards increased, so did the difference in vigor in the movements made toward the two options.

3.3.3 Between subject differences in saccade vigor

In addition to within-subject changes in saccade vigor during the decision-period, there were also between-subject differences in the saccadic eye movements: for a given saccade amplitude, some participants consistently moved their eyes with high velocity, whereas others consistently moved their eyes with low velocity. That is, there were between-subject differences in saccade vigor. We quantified this difference and asked whether it was related to differences in decision-making.

We began by fitting Eq. (3.1) to the velocity-amplitude data of each participant. For participant n , this produced parameter values $\hat{\alpha}_n$ and $\hat{\beta}_n$. We found the median of the $\hat{\alpha}$ and $\hat{\beta}$ distributions across all participants, producing $\bar{\alpha}$ and $\bar{\beta}$. The values of $\bar{\alpha}$ and $\bar{\beta}$ were 690.4 and 0.089, and 764.3 and 0.082, for nasal and temporal saccades, respectively. We used this estimate to produce a canonical relationship between amplitude and velocity across the entire population:

$$\bar{v} = \bar{\alpha} \left(1 - \frac{1}{1 + \bar{\beta}x} \right) \quad (4)$$

We used the above relationship to quantify the relative vigor of saccades in one participant as compared to another. We followed the procedure described in [53]: we re-fitted each participant's saccade velocity-amplitude data to a one-parameter scaling function of the canonical function:

$$\hat{v}_n(x) = \lambda_n \bar{v}(x) \quad (5)$$

Parameter λ_n is the between-subject measure of vigor for subject n . When we have $\lambda_n > 1$, it indicates that the saccades of participant n are generally faster than the population median.

Figure 3-5A (left panel) illustrates saccade peak velocity as a function of saccade amplitude for two participants. Participant S14 had consistently faster saccades than participant S6. The right panel of Figure 3-5A shows the canonical function (dashed line, representing the population median), and the function representing the data for each participants (derived from Eq. 3.5). To quantify goodness of fit,

we computed correlation coefficients, reflecting the ability of the one parameter model of Eq. (3.5) to account for the saccade velocity/amplitude data of each subject. The results are illustrated in Figure 3-5B. For every subject the fit between the model used to estimate between-subject saccade vigor and actual velocities was significant at a level of $p < 0.00001$.

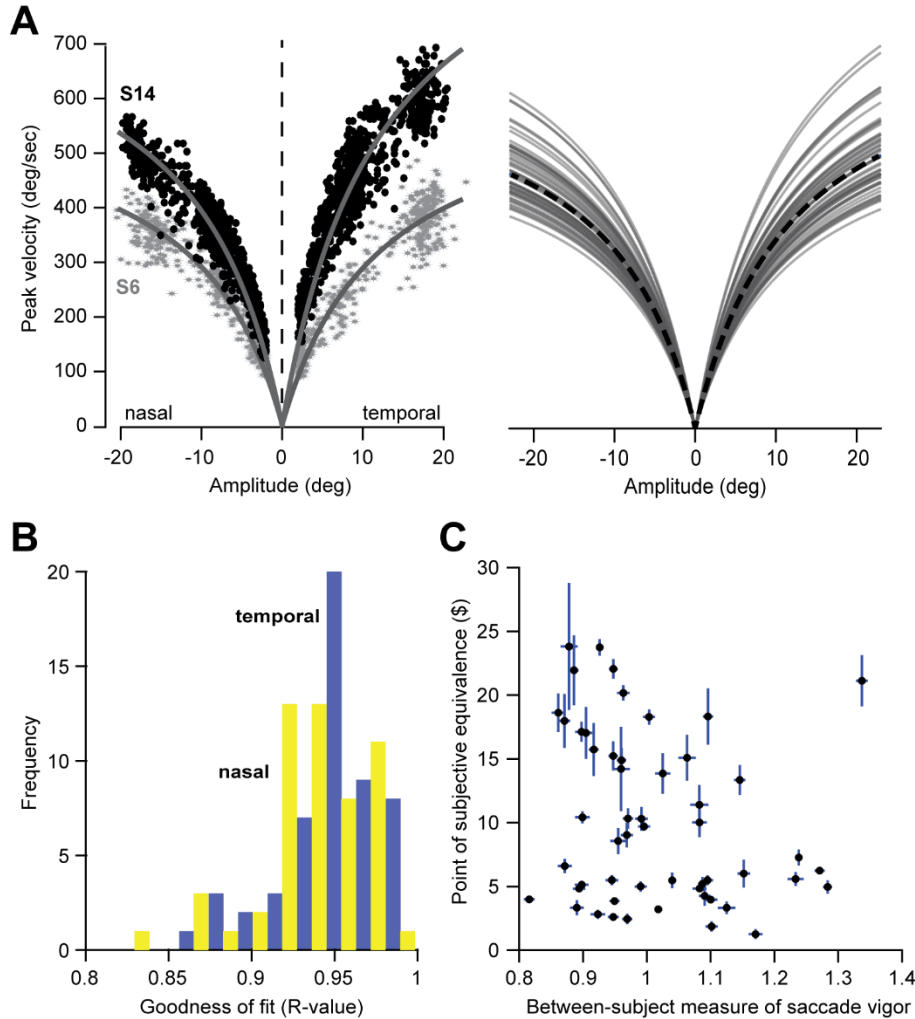


Figure 3-5: Between-subject differences in vigor. **A.** Peak velocity as a function of amplitude for two representative participants. Subject 14 made consistently faster saccades than subject 6. For each subject, we fitted the velocity-amplitude data for nasal and temporal saccades separately to a two parameter hyperbolic function. Using these parameters, we computed the average relationship between velocity and amplitude across the population (dashed lines, right). Finally, we computed the between-subject measure of vigor as the scaling of the population average velocity-amplitude relationship. **B.** Goodness of fit of the hyperbolic model to the velocity-amplitude relationship in each subject. Dark and light gray bars represent fits to nasal and temporal saccades, respectively. **C.** Between-subject measure of saccade vigor was not a predictor of subjective equivalence between immediate and delayed rewards.

Using this measure of between-subject saccade vigor, we asked whether individuals who moved with greater vigor were distinguishable in their patterns of decision making. We focused on between subject differences in impulsivity, i.e., the equivalence point between the immediate and delayed reward. For example, participant S41 has a larger equivalence point than participant S21 (Figure 3-3B). This

translates into a larger temporal discount rate, implying a greater impulsivity. However, we did not find a significant relationship between vigor and impulsivity (Figure 3-5C, $p=0.078$). Nor did we find any relationship between vigor and discount factor k ($p=0.074$). Therefore, in this task the between-subject differences in saccade vigor were not a predictor of differences in decision-making.

3.4 Discussion

We considered saccades that participants made as they considered two monetary options: a small reward to be obtained immediately vs. a larger reward to be obtained at a delay of 30 days. We found that saccade vigor, a within-subject measure of peak velocity normalized by amplitude, was greater during the deliberation period, before the decision was made, than immediately after. Vigor dropped rapidly in trials in which participants made a quick decision, but dropped slowly in trials in which they took longer. Among the saccades made just before and just after the decision, saccades to the preferred option exhibited a greater vigor than saccades to the non-preferred option. The participants signaled their decision about one half second after saccade vigor diverged between the two options. The disparity between vigor of saccades to the two options near decision-time became larger as the difference in the subjective values of the two options increased. Therefore, during decision-making, the subjective value that the brain assigned to a stimulus was shared with the motor system, influencing the vigor with which the eyes moved toward that stimulus.

3.4.1 Link between the neural basis of vigor and encoding of reward

During decision-making, temporal discounting is associated with release of dopamine. When an animal makes a decision between a small magnitude, small-delay reward, and a large magnitude, large-delay reward, dopamine cells fire in response to each stimulus by an amount that correlates with the subjective value of that stimulus [22]. It appears that some of the neural circuits that are critical for control of vigor are also influenced by a neurotransmitter that has been linked to subjective valuation of reward. This link, we speculate, may be the reason for the modulation of saccade vigor during the deliberation process. Future functional imaging studies will allow us to assess this speculation.

3.4.2 Subjective value vs. motivational salience

We found that vigor reflected the subjective value of the stimulus that acted as the goal of the movement. However, an alternate hypothesis is that vigor is a reflection of the motivational salience of the stimulus. This alternate hypothesis predicts that because motivational salience associated with loss of \$10 is greater than loss of \$5, vigor will be greater toward -\$10 than -\$5, despite the fact that the subjective value of -\$5

is greater than -\$10. The subjective value hypothesis predicts the opposite: vigor should be higher for -\$5 than -\$10.

Kobayashi et al. [59] asked monkeys to view a cue that determined whether the upcoming saccadic eye movement was to be rewarded (apple juice), punished (airpuff), or neutral (sound). Motivational requirements were highest for juice and airpuff, and lowest for neutral, as evidenced by the fact that correct performance rates were highest in the reward and airpuff trials, and lowest in the neutral trials. In contrast, the subjective value was highest for juice, and lowest for airpuff. They observed that saccade peak velocity was highest for juice trials, lowest for airpuff, and in-between for neutral. This suggests that vigor is more affected by subjective value than motivational salience. However, an experiment that more directly asks this question remains to be performed.

Activity of saccade related cells in the lateral intraparietal sulcus (LIP) rises with the subjective value of the stimulus [60]. Is this a reflection of the utility of the stimulus, or the vigor of the upcoming movement? In Louie and Glimcher [60] the authors found that the change in population response in LIP was proportional to the change in subjective value, with a gain that was close to one. In contrast, here we found that change in saccade vigor was a maximum of 7% as compared to a change in subjective value of around 35%, a gain of 0.2. Therefore, we speculate that activity of LIP is more closely related to the utility of the action, as compared to the vigor of that action.

3.4.3 Modulation of vigor during decision-making

In our task, the decision was expressed by hitting a key and reward was associated with the choice of that key. The saccades were not associated with reward, but were a means by which the brain acquired information for the purpose of making the decision. This is in contrast to many earlier experiments where the act of making a saccade was itself associated reward [8] [5] [59], but see [61] [62]. Despite this, saccade vigor was modulated by subjective value of the stimulus. Our results suggested that during decision-making, actions that acquired information relevant to the eventual decision had a subjective value associated with them, as evidenced by the vigor of that action.

This view provides a potential explanation as to why vigor dropped after the choice was indicated. We speculate that saccades that were made during the deliberation period had a greater vigor because each movement acquired information relevant to the eventual reward. Once the choice had been indicated, the same actions no longer acquired relevant information. In this sense, the subjective values of the movements performed during deliberation were higher than those performed after.

A recent experiment by Cisek and colleagues [61] noted that urgent decisions were followed by more vigorous movements. They suggested that during decision-making, a rising urgency signal combined with the process of evidence accumulation, invoking a hastier (earlier) decision in some

circumstances, a more deliberate (later) decision in other circumstances. Vigor was affected by the rate of rise of this urgency signal. These results complement our findings by demonstrating that in addition to subjective value of the stimulus, other contextual factors like rate-of-reward can affect both decision-making and movement vigor [63].

3.4.4 Between-subject differences in vigor

As in our earlier work [53], here we found that there were consistent between-subject differences in saccade vigor: some individuals moved their eyes with much greater velocity than other individuals. In Choi et al. [53] we found that individuals who had greater saccade vigor were also less willing to wait in order to increase their probability of success. In that specific task we found that the high vigor individuals appeared impulsive. However, here in a value-based decision-making task we found no relationship between temporal discounts rates (as measured by point of subjective equivalence between the delayed and immediate rewards) and saccade vigor.

There are a number of reasons that could underlie this disparity. To measure temporal discounting, in Choi et al. [53] we designed a task in which each choice had an immediate and real consequence, acting as an operant reinforcement on the next choice. In contrast, here we measured temporal discounting in a task where choices had consequences that were not experienced until after the end of the experiment. While both types of approaches produce measures of temporal discounting, they produce inconsistent results in the same person [64], and produce greatly differing discount rates [65]. Therefore, fundamental differences in how one measures temporal discounting during decision-making may underlie differences in the two studies.

Another possibility is that in a value-based decision-making task without immediate consequences participants may have more control over their explicit decisions, a phenomenon commonly referred to as impulse control [66] [67]. For example, it has been shown that Parkinson's disease patients who have been treated with a dopamine agonist have both increased saccade vigor [68] and a higher propensity for impulse control disorders [69]. Thus it seems possible that modulation of movement vigor is a measure which can be used to ascertain choice preference, even when subjects may be hiding their true preferences with their explicit decisions.

3.4.5 Conclusions

During deliberation period of a decision-making process, vigor was similar as saccades were made between the two options, but diverged about 0.5 seconds before decision-time, becoming greater for the option that was eventually chosen. Therefore, vigor of the movement that brought the gaze toward an option was affected by the value that the brain assigned to that option. Overall, our results suggest a link

between the neural mechanism that assigns value to a stimulus, and the mechanism that controls vigor of movements toward that stimulus.

4 A shared mechanism for control of vigor in saccades, reaching, and head movements

4.1 Materials and methods

4.1.1 Experiment 1

4.1.1.1 *Participants*

We asked $n=335$ healthy subjects, between the ages of 18 and 46 and with no known neurologic deficits, to participate in the first study (age 21.8 ± 4.2 mean \pm sd, 157 females). All were naïve to the paradigms and to the purpose of the experiments. Subjects slept 7.0 ± 1.5 (mean \pm sd) hours the evening before the study. All participants signed a written consent form approved by the Texas State University Institutional Review Board. Each subject was compensated with \$10/hr for participating in the study.

4.1.1.2 *Experimental protocol*

We used a View Sonic 22 inch screen (474x297 mm, 1680x1050 pixels) to display the targets. The distance from the subject to the display was 550 mm. Subjects' eyes were located approximately 419 mm from the level of the desk upon which the screen sat. Targets were displayed from 234-532 mm above the level of the desk. We used an Eyelink 1000 tracker system (SR Research, sampling rate = 1000 Hz) to track the gaze orientation of the left eye of all subjects. The spatial accuracy of the eye tracker was 0.48 ± 0.17 deg and 0.50 ± 0.17 deg for the first and second sessions, respectively. We removed 49 subjects from all analyses due to poor eye tracker calibration, leaving a total of $n=286$ subjects for full analyses.

We divided the first experiment into two sessions. For the first session, subjects were asked to make a series of 100 saccades between two target locations. The targets were located symmetrically about the vertical meridian of the screen, specifically at ± 15 deg. The target was black, with a white edge shading to aid subjects in fixation. Each of the 100 trials lasted 1 sec, for a total of 100 sec per session. In this manner, subjects, knew the upcoming target location.

For the second session, subjects were asked to make a series of 100 saccades among random target locations on the viewing screen. Targets could appear anywhere between ± 15 deg horizontally, and ± 9 deg vertically. The minimum target jump size was 4 deg. As in the first session, each of the 100 trials lasted 1 sec. Because the target location was random, subjects could not predict the upcoming location of each target.

4.1.2 Experiment 2

4.1.2.1 *Participants*

We asked n=50 healthy subjects, between the ages of 18 and 45 and with no known neurologic deficits, to participate in the second study. All were naïve to the paradigms and to the purpose of the experiments. All participants signed a written consent form approved by the Johns Hopkins University Institutional Review Board. They were compensated with \$15/hr in cash for participation.

4.1.2.2 *Experimental protocol*

Subjects sat in a darkened room in front of an LG Model 32LN5300 TV screen (70.0 x 39.5 cm, 1920 x 1080 pixels, light grey background, frame rate = 60 Hz). An Eyelink 1000 Remote recording system was used to acquire gaze position on the viewing screen at 500 Hz. Additionally, a Trakstar 3D Guidance electromagnetic tracking system (Northern Digital, Inc.) was used to record position and orientation of the head and hand at 200 Hz. All experimental procedures was approved by the Johns Hopkins Institutional Review Board.

We recorded simultaneous eye, head, and reaching movements as subjects completed a natural pointing task. Fifty subjects completed the task, 14 of which were excluded from all analyses because the proportion of primary saccades recorded for those subjects was below 65%. For instance, if a subject blinked during the primary gaze shift, then we could not analyze the saccade and therefore could not record a movement for that trial. Subjects were presented with black visual targets (diameter=1 deg), and then reached and touched each target. The movement included motion of the eye, head, and arm, with amplitudes ranging from 10-50 deg in 24 subjects, and 30-50 deg in 12 subjects. Targets always appeared in the upper quadrant of the viewing screen, and appeared at the same vertical location throughout the experiment. In this manner, all gaze shifts and reaching movements were horizontal (i.e. leftward or rightward).

Of the 36 subjects that completed the task, 20 completed a version of the task consisting of 4-6 blocks (mean length = 60 trials). These subjects performed movements following horizontal targets jumps with amplitudes ranging continuously between 5 and 50 deg. All subjects were required to complete at least 4 blocks of the task. The other 16 subjects completed a version of the task requiring completion of 16 blocks. These subjects completed four mini-sessions, each consisting of four blocks. During the first and third min-sessions, subjects were instructed to look at the targets as they appeared on the viewing screen. No instruction was given regarding head movements. During the second and fourth mini-sessions, subjects were instructed to look at and reach to the targets. Subjects were seated and strapped with seatbelt into a Corbeau Model Forza race car chair, which was mounted onto a rolling

platform. The chair provided resistance to rotation of the torso, such that subjects' gaze shifts consisted primarily of eye and head movements.

After completion of the head-free gaze shifts, we asked a sub-set of 23 subjects to complete four additional blocks of 60 saccades with head restrained using a dental bite bar. Subjects were located at a distance of 40 cm from the viewing screen, and the sampling rate of the eye tracker was 1000 Hz.

4.1.3 Data analysis

All data analyses were completed using either Matlab R2015b, or Matlab R2016a (Mathworks, Inc.). For each movement modality, we identified all primary movements for each subject. Primary movements had a displacement that was at least 66% of that of the target on a given trial. The reaction time of all primary movements was less than 550 ms. We used all movements to construct a main sequence relationship of peak velocity and displacement for head, hand, and eye movements. For the saccade data, we used our definition of the main sequence from Experiment 1 based on 286 subjects. For the reaching and head movement data, we based our definition of the main sequence on the collection of all task-relevant movements of each modality.

Given the definition of the main sequence for each modality, we fit a scaling of this function (Equation 3.1) to each subject's data. This provided an estimate of vigor of movements of the individual, which we used to assess conservation of vigor across modalities. We also used the fit of the main sequence to establish the vigor of each individual movement. This vigor was the ratio of the peak velocity to the expected peak velocity, at that movement's displacement. We used our estimate of the vigor of individual movements to assess trial-to-trial changes in movement vigor.

4.2 Results

4.2.1 Conservation of vigor of saccades

We previously showed that vigor of saccades is a consistent trait of the individual [53]. In particular, we showed that across-subject differences in saccade vigor are much more significant than day-to-day within-subject differences. Asymptotic peak velocities of subjects ranged from roughly 400 to 700 deg/sec, and those differences were consistent across multiple days. Here, we asked whether or not those across-subject differences were conserved in vertical saccades. We found a wide range of saccade vigor across the population. Importantly, subjects who exhibited high vigor for horizontal saccades also exhibited high vigor for vertical saccades.

We asked subjects ($n=286$) to make series of 100 saccades across the viewing screen at a rate of 1 Hz. Target location was constrained to lie within a rectangle ± 15 deg horizontally and ± 9 deg vertically.

The minimum target displacement was 4 deg. In this manner, we collected a series of both horizontal and vertical saccades in each subject. We repeated this 100-trial procedure twice in each subject. Average velocity profiles of vertical and horizontal saccades for subjects S141 and S270 are given in Figure 4-1A and B. These profiles correspond to saccades with total displacement between 10 and 20 deg. Subject S141 (black profiles) exhibited more vigorous saccades than Subject S270 (blue profiles) in both horizontal and vertical directions. Figure 4-1C and D show these between-subject differences for the main sequence of horizontal and vertical saccades. To construct the hyperbolic fits to the data in Figure 4-1C, we grouped all horizontal saccades recorded across the subject population, and fit Equation (3.1) to that data. We then fit each individual subject's data to a scaling of this baseline definition of the horizontal main sequence of saccades. An analogous procedure was used to construct the fits to vertical saccades shown in Figure 4-1D. We observed a high degree of conservation of saccade vigor across subjects. Figure 4-1E gives vigor of vertical saccades of each subject as a function of their horizontal saccade vigor. Those subjects that exhibited more vigorous horizontal saccades also executed more vigorous vertical saccades ($R^2=0.63$).

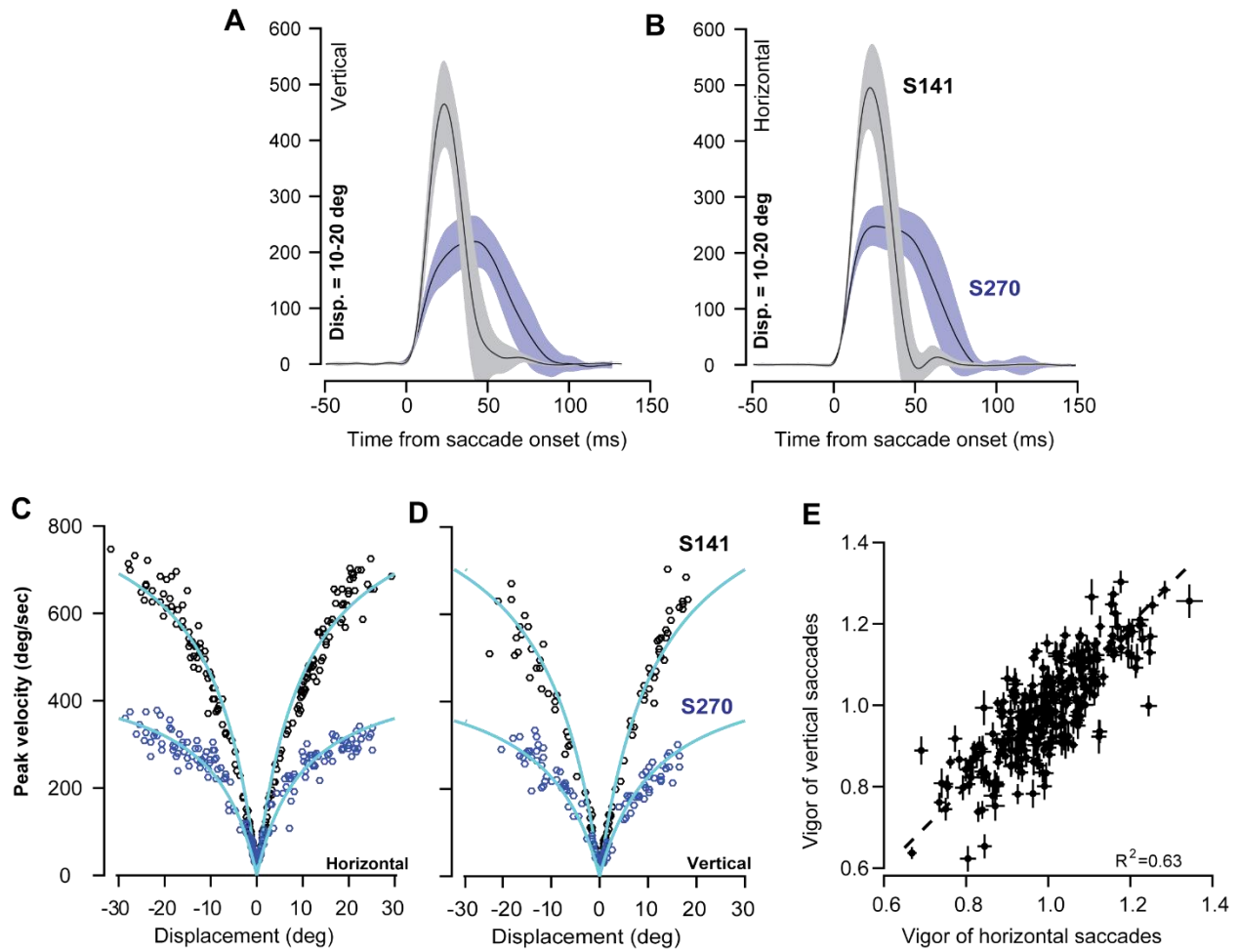


Figure 4-1: Conservation of vigor across saccade directions. **A.** Mean velocity profiles of vertical saccades for two subjects (S141 – high vigor – black, and S270 – low vigor – blue). We averaged across all saccades with total displacement between 10 and 20 deg. Error bars are SD. **B.** Mean velocity profiles of horizontal saccades for the same two subjects. Error bars are SD. **C.** Main sequence relationship of peak velocity and displacement of horizontal saccades for subjects S141 (black dots) and S270 (blue dots). Light blue lines represent the fit of the definition of the main sequence of horizontal saccades to each subject’s data. **D.** Main sequence relationship of vertical saccades for the same two subjects. Light blue lines are the fit of the vertical main sequence to each subject’s data. **E.** Relationship of vigor of horizontal and vertical saccades across subjects. Each point represents a single subject. Different definitions of the main sequence were used for horizontal and vertical saccades. Error bars are SEM.

To summarize, we asked whether or not across-subject differences in vigor of horizontal saccades were conserved in vertical saccades. We found that these differences were indeed conserved across saccade directions. Thus, vigor seems to be a general property of the saccadic system, as opposed to a phenomenon which depends on the direction in which the saccade is executed.

4.2.2 Reaction times were lower for horizontal saccades

We next asked if there were any differences in the reaction time of saccades as a function of movement direction. We first split saccades into bins according to the total displacement of the saccade. Figure 4-2A and B shows average across-subject timing and magnitude of velocity profiles for horizontal and vertical saccades of magnitude 8-12 deg and 16-20 deg, respectively. We found that reaction times of horizontal saccades (purple traces) were lower for both saccade magnitudes. Figure 4-2C shows a histogram of the difference in timing of horizontal and vertical saccades across subjects. Horizontal saccades were initiated 10 ms faster, on average, than their vertical counterparts. We then split vertical saccades into two groups according to movement direction: upward and downward (Figure 4-2D). We found that downward-directed saccades were 2-3% more vigorous than upward saccades ($p < 10^{-4}$).

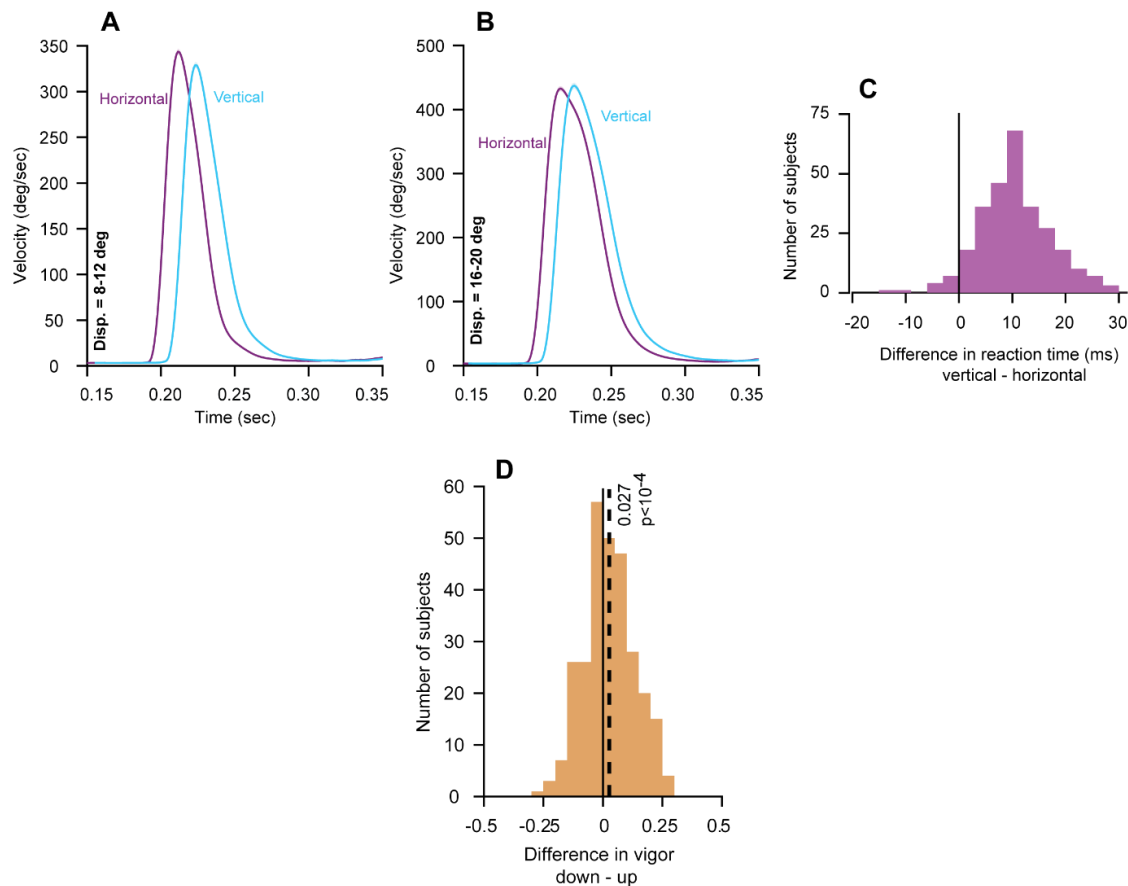


Figure 4-2: Subjects responded to horizontal target jumps quicker than to vertical jumps. In addition, downward saccades were more vigorous than upward saccades. **A.** Mean velocity profiles of horizontal (purple) and vertical (blue) saccades across all subjects, aligned to time of target onset. We found the average profile of all saccades with displacement between 8 and 12 deg for each subject, and averaged across all subjects. Error bars are SEM. **B.** Mean across-subject velocity profiles of horizontal and vertical saccades of displacement between 16 and 20 deg. Error bars are SEM. **C.** Histogram of the difference in reaction times between vertical and horizontal saccades for all subjects. Horizontal saccades were made with quicker reaction time than their vertical counterparts. **D.** Histogram of the difference in vigor of downward and upward vertical saccades for each subject. Downward saccades were significantly more vigorous than upward saccades.

To summarize, we analyzed within-subject differences in reaction time and vigor vs. saccade direction. We found that horizontal saccades were generated with quicker reaction times than vertical saccades. In addition, downward-directed saccades were more vigorous (by roughly 2-3%) than upward saccades.

4.2.3 Speed-accuracy tradeoff in saccades

A number of previous works have emphasized the possible absence of a speed-accuracy tradeoff in saccades [5] [7]. Those works focused on the effects of reward on saccade execution. In particular, with the introduction of reward, saccades became both faster and more accurate. If motivational state has a significant effect on the speed-accuracy tradeoff in saccades, then there should be differences in this tradeoff across subjects. In particular, we would expect more motivated subjects to execute faster and more accurate saccades than their less motivated counterparts.

To test this hypothesis, we asked subjects ($n=286$) to make a series of 100 30-deg horizontal saccades across the vertical midline of the viewing screen. Targets switched back and forth between ± 15 deg at a rate of 1 Hz. For each subject, we identified all primary saccades (saccades with displacement ≥ 20 deg and reaction time ≥ 55 msec). We then calculated the standard deviation of the distributions of peak velocity and saccade displacement for each subject. Figure 4-3A and C gives the histograms of standard deviation of peak velocity and displacement, respectively, across all subjects.

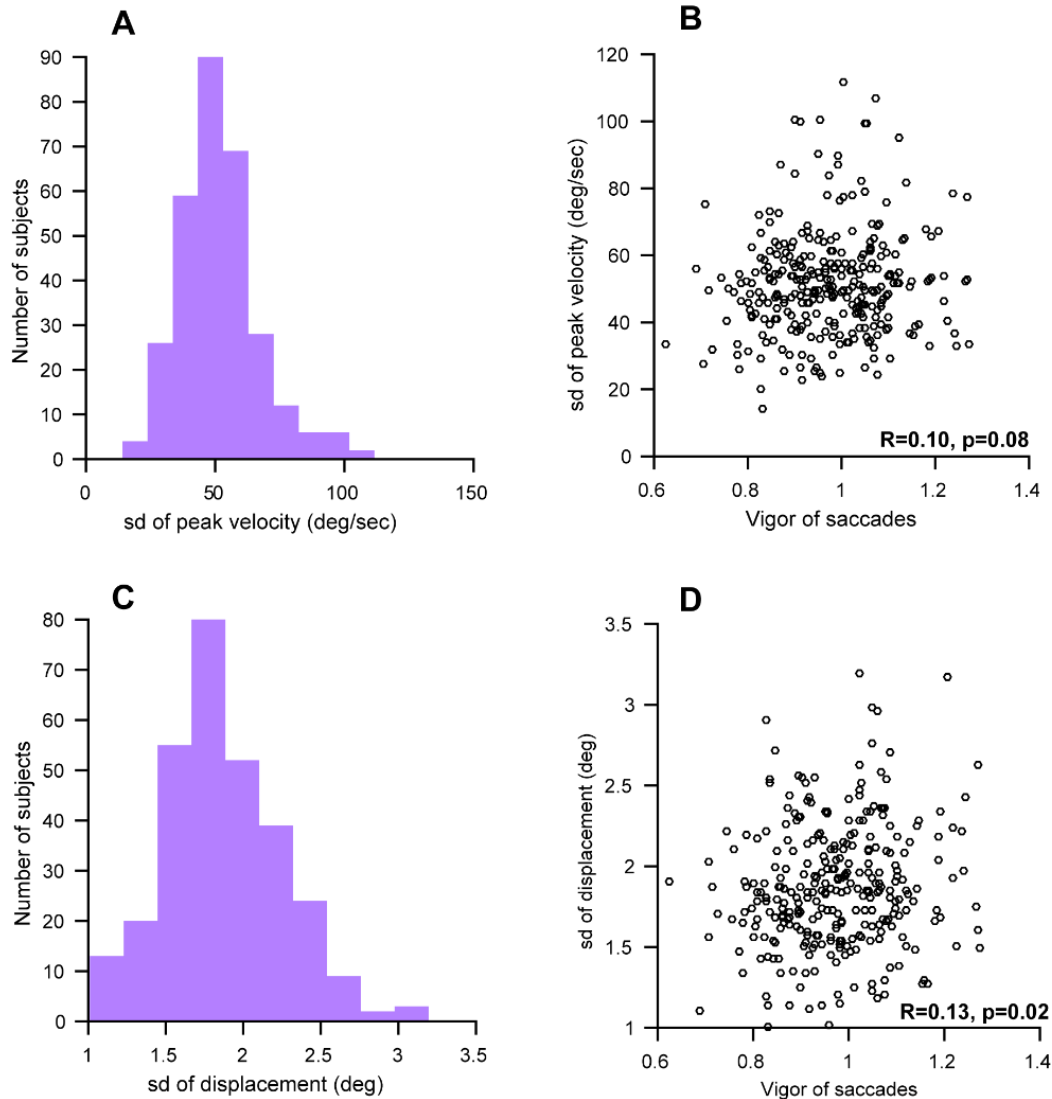


Figure 4-3: We found a small yet significant relationship between speed and accuracy of saccades. **A.** Histogram of standard deviation of peak velocity for all subjects. For each subject, we found the standard deviation of peak velocity for all saccades to targets with displacement 15 deg. **B.** Across-subject relationship between standard deviation of peak velocity and vigor of saccades. Each dot represents a single subject. Subjects with more vigorous saccades did not necessarily have higher noise in their peak velocities. **C.** Histogram of standard deviation of displacement for all subjects, again for saccades following target displacement 15 deg. **D.** Across-subject relationship between standard deviation of displacement of saccades and vigor. Subjects with more vigorous saccades had slightly higher variability in saccade displacement ($p < 0.05$).

To compare saccade accuracy to vigor, we needed an across-subject measure of vigor of saccades for each subject. We used the definition of the main sequence of horizontal saccades from Equation ## and the horizontal saccade data from the previous task (Section 4.2.1) as our baseline measure of saccade vigor. Each subject's vigor was determined as a best-fit scaling of this baseline function to the main sequence relationship of peak velocity and displacement. We found that vigor of saccades of an individual did not predict standard deviation of peak velocity, a measure of accuracy (Figure 4-3, $R=0.10$,

$p=0.08$). However, vigor of saccades was a significant predictor of standard deviation of saccade displacement (Figure 4-3D, $p<0.05$), although this relationship was also weak ($R=0.13$).

In summary, we asked whether or not the vigor with which a subject made horizontal saccades was a significant predictor of that subject's saccade execution accuracy. We found that vigor was a significant predictor of execution accuracy across subjects, although this effect was quite small. Thus, in the absence of reward, subjects with more vigorous saccades exhibited more noise in saccade execution, suggesting the existence of a speed-accuracy tradeoff in saccades.

4.2.4 Conservation of vigor across movement modalities

We then asked whether vigor was conserved across movement modalities. We recorded simultaneous eye, head, and reaching movements in subjects ($n=36$) as they completed a natural pointing task. Subjects were presented with visual targets, and then reached and touched that target. Of the 36 subjects, 20 were asked to complete 4-6 blocks with mean 60 trials, whereas the other 16 were asked to complete 16 blocks containing 70 trials. This latter group of subjects performed 8 blocks with reaching movements, and 8 blocks consisting of only head-free gaze shifts to the targets.

We first analyzed saccade peak velocity and reaction time as a function of displacement. We binned each of these parameters based on target displacement (bin width = 10 deg). Figure 4-4A and C presents saccade reaction time and peak velocity, respectively, as a function of target displacement. We found that, as target displacement increased, reaction time of saccades increased, whereas that of head movements generally decreased. This decrease in the reaction time of head movements relative to saccades was already reported [70]. We also found a slight increase in the reaction time of reaching movements with target displacement, although this effect was much less appreciable. Across subjects, we found the typical nonlinear relationship between saccade peak velocity and displacement. However, peak velocities of both reaching movements and head movements increased in a linear fashion with target displacement.

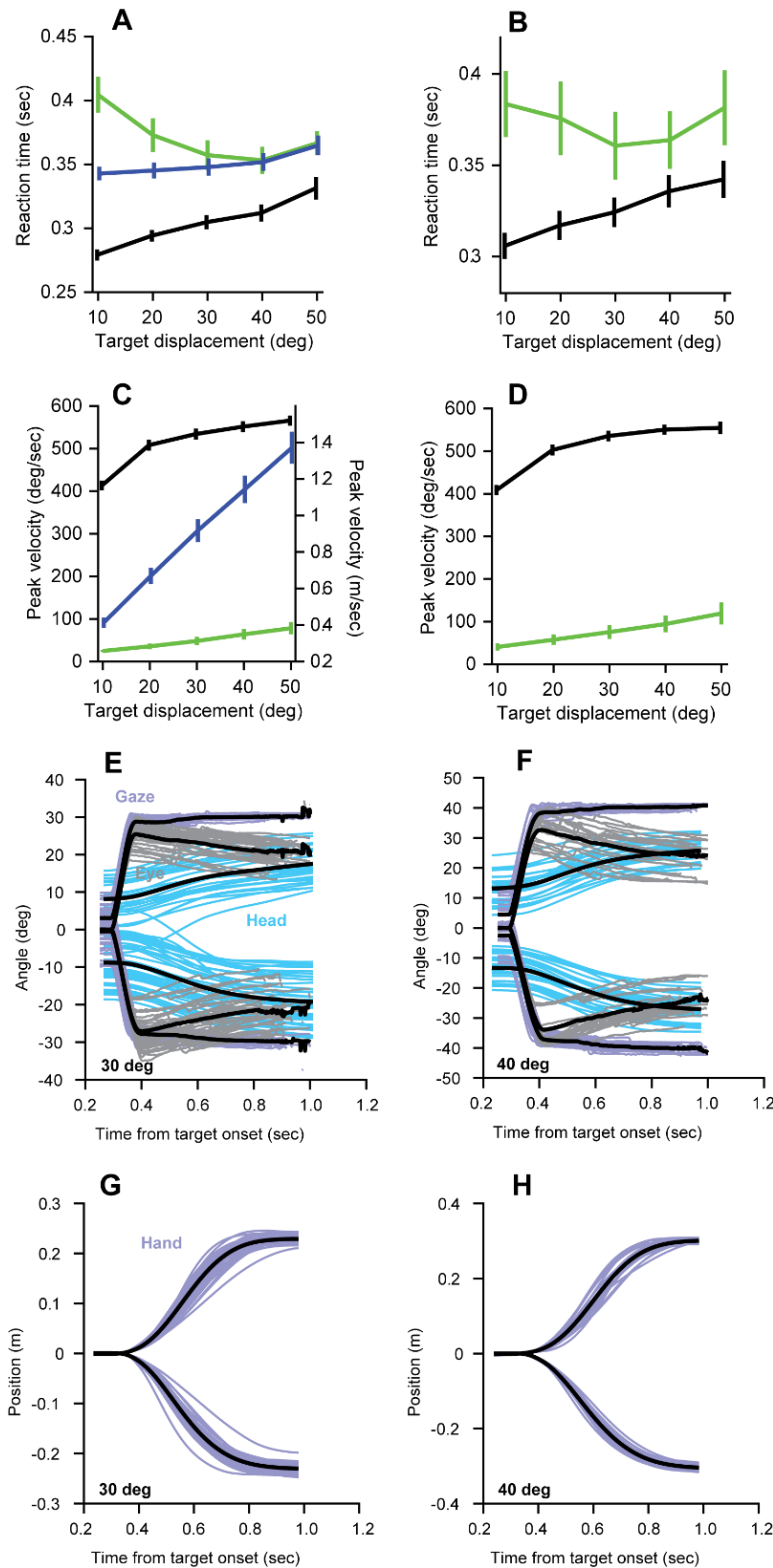


Figure 4-4: Reaction time and peak velocity of movements (eye, head, and hand) vs. target displacement for all subjects that completed a total of 16 blocks (n=16). Single-subject example of movement position traces generated during the head-free movement task. **A.** Reaction time of head (green), hand (blue), and eye (black) movements vs. target displacement with reaching. In general, reaction time of both reaching movements and saccades increased with displacement, whereas that of head movements tended to decrease. Error bars are SEM. **B.** Reaction time vs. target displacement without reaching. In the absence of reaching movements, there was no effect of target displacement on head movement reaction time. Error bars are SEM. **C.** Peak velocity vs. target displacement with reaching. We observed a linear relationship between peak velocity and displacement for head (green) and hand (blue) movements. Error bars are SEM. **D.** Peak velocity vs. target displacement without reaching. **E.** Trajectories of all gaze shifts for a single subject to targets of displacement 30 deg. Leftward trajectories have negative orientation values. Head, eye-in-head, and gaze on screen are shown in blue, gray, and purple, respectively. Black lines denote mean trajectories of each type. **F.** Trajectories of all gaze shifts for the same subject to targets with displacement 40 deg. Color key is the same as in (E). Leftward trajectories have negative orientation. Black lines denote mean trajectories of each type. **G.** Trajectories of all reaching movements to targets of displacement 30 deg. **H.** Trajectories of all reaching movements to targets of displacement 40 deg.

Figure 4-4E and F shows all gaze shifts for a single subject for target displacements of 30 and 40 deg, respectively. Head motion started consistently later than eye motion. There was a large amount of trial-to-trial variability in the starting head orientation; this was due to an effect of the previous trial's target location.

Once the saccade was complete, the vestibulo-ocular reflex (VOR) kept the subject's gaze fixated on the target while the head continued to rotate. Figure 4-4G and H shows the same subject's set of reaching

movements during the same trials for which target displacement was 30 and 40 deg, respectively. In general, the hand motion started well after the saccade (Figure 4-4A).

To address our question of conservation of vigor across movement modalities, we needed a definition of across-subject vigor of movements for reaching and head movements, in addition to saccades. To generate these definitions, we used the same two-parameter hyperbolic relationship between size and peak velocity that we used for saccades (Equation (3.1)). For reaching movements, we pooled together all movements made by all subjects in the task, and fit the two-parameter model to that data. This was our baseline definition of the velocity-displacement relationship of reaching movements. We then scaled this baseline definition either up or down to fit each individual subject's data, giving a one-parameter estimate of across-subject vigor of reaches. We used the same procedure to estimate vigor of head movements.

Figure 4-5A shows the main sequence relationship of peak velocity and displacement for reaching movements for subjects S9 and S11. Subject S9 exhibited markedly faster reaching movements than S11. Figure 4-5B and C shows the main sequence of saccades for subjects S25 and S3, and head movements for subjects S9 and S16. We observed a marked difference in the vigor of the movements generated by both pairs of subjects, highlighting the large range of vigor of movements that we observed in each modality. Given our estimates of vigor of movements in each modality, we then asked whether or not vigor was conserved across all three modalities. Figure 4-5D shows the relationship between vigor of head movements and head-free saccades. Each point represents a single subject. We observed a weak positive relationship, although it did not reach significance. Similarly, we observed no significant relationship between vigor of reaching movements and that of head-free saccades (Figure 4-5E). We did, however, observe a positive relationship between vigor of reaching and head movements (Figure 4-5F, $R^2=0.64$, $p<10^{-5}$).

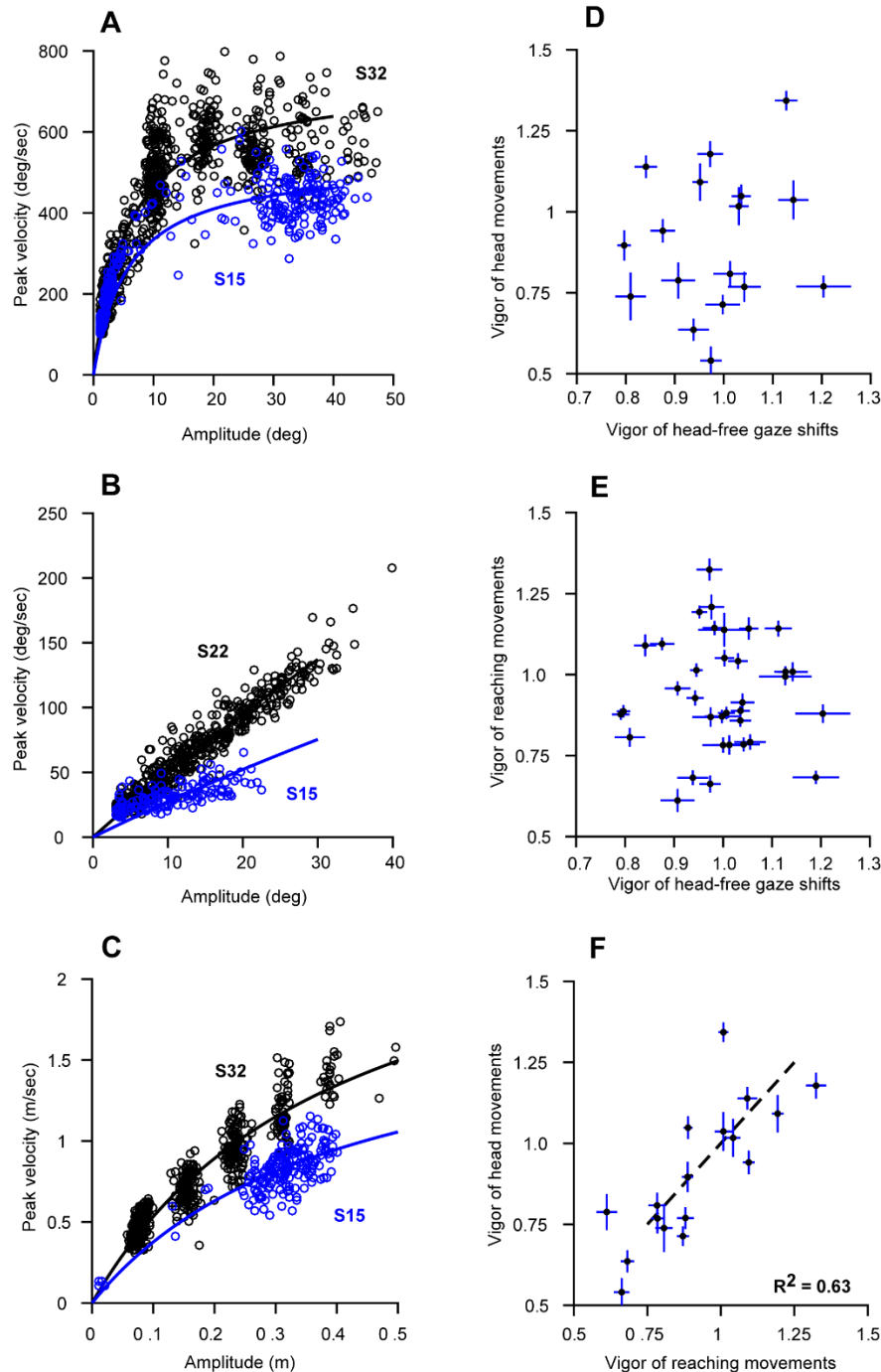


Figure 4-5: Consistency of vigor across movement modalities. **A.** Main sequence relationship of peak velocity and displacement of saccades of subjects S32 (black dots) and S15 (blue dots). We first collected all data from all subjects ($n=36$) who completed the head-free gaze shift task. We then determined the baseline definition of the main sequence of saccades with a fit to all across-subject data. For each individual subject, we scaled the baseline definition to best fit that subject's data. Best fit lines for subjects S32 and S15 are given as black and blue lines, respectively. **B.** Main sequence relationship between velocity and displacement of head movements generated by subjects S22 and S15. We pooled head movement data across all subjects and computed a baseline definition of the main sequence of head movements. We then scaled this definition to fit each subject's main sequence relationship individually. Single-subject fits to the head movement data for subjects S22 and S15 are given as black and blue lines, respectively. **C.** Main sequence relationship between peak velocity and displacement of reaching movements for subjects S32 (black dots) and S15 (blue dots). We first fit a baseline hyperbolic function to all reaching movement data from all subjects ($n=36$). We then scaled this baseline definition of the reaching main sequence up and down in order to fit each individual subject's data. Single-subject fits of the main sequence for subjects S32 and S15 are shown as black and blue lines, respectively. **D.** Across-subject plot of vigor of head movements vs. head-free gaze shifts. We found no relationship. **E.** Across-subject plot of vigor of reaching movements vs. vigor of head-free gaze shifts. We found no relationship. **F.** Across-subject relationship between vigor of reaching movements and vigor of head movements. We found a significant increase in reaching movement vigor with head movement vigor ($R^2=0.63$, $p<10^{-5}$).

To summarize, we observed no significant relationship between vigor of reaching movements and saccades, or head movements and saccades. We did, however, observe a significant positive relationship between vigor of head movements and reaching movements, across subjects. That is, those subjects that moved their hand faster than average were also those who moved their head faster.

4.2.5 Temporal modulation of movement vigor

Each subject had a baseline level of vigor of saccades, head movements, and reaching movements. For 19 subjects, we were unable to ascertain the baseline vigor of head movements, as they chose not to move the head to a sufficient degree. Thus, for those subjects, we could not estimate vigor of individual head movements. We used the same procedure presented in Section 3.2.3 to estimate the vigor of a single movement. That is, for each movement, we defined vigor as the peak velocity of the movement divided by the expected peak velocity at that movement's displacement. Note that the expected movement is a function of the subject's baseline amount of vigor, and thus is different from subject to subject. Given this definition of vigor of individual movements, we asked whether or not temporal changes (i.e. changes with trial number) in one modality predicted similar changes in other modalities.

Figure 4-6A shows change in within-subject vigor of reaching movements versus time. We found that, with the exception of the first trial of the experiment, vigor of reaching movements was highest at the first trial of the block, and declined abruptly within each block. We also observed the well-established decay in saccade vigor within each block (Figure 4-6B) [6] [63]. Although we did not have the same amount of data with which to analyze trial-to-trial changes in head movement vigor, we did observe an appreciable decline in vigor within each block (Figure 4-6C). That is, within a block, as the vigor of saccades declined, so did that of both reaching movements and head movements. Vigor of all modalities was highest at the outset of the block, and declined to lowest values by block end.

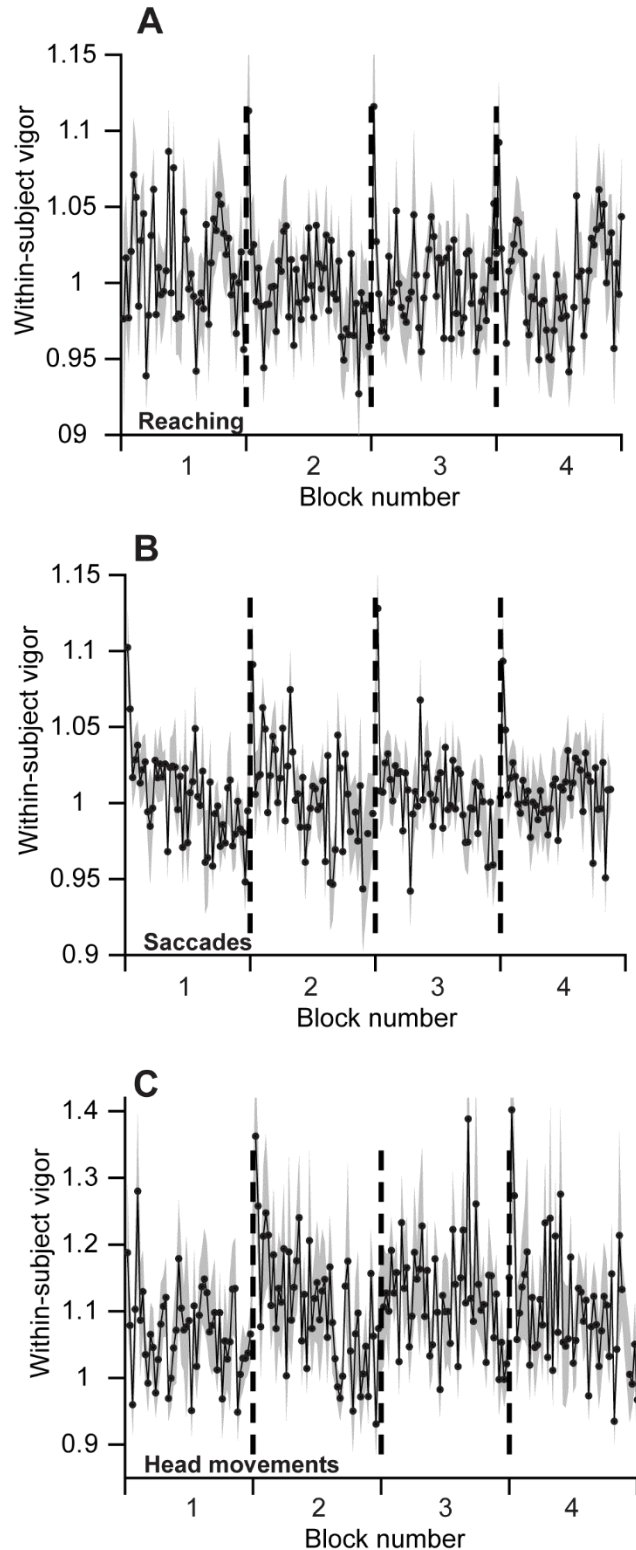


Figure 4-6: Within-subject modulation of movement vigor for those subjects that completed 4-6 blocks of Experiment 2 (n=20). For each movement modality (reaching movements, saccades, and head movements), we defined the vigor of a single movement as the peak velocity of that movement divided by the expected peak velocity given that movement's total displacement. We used this ratio of peak velocity to expected peak velocity as our definition of vigor of a single movement. This definition was the same across all modalities. **A.** Within-subject changes in vigor of reaching movements. Error bars are SEM. **B.** Within-subject changes in vigor of saccades. Error bars are SEM. **C.** Within-subject changes in vigor of head movements. Error bars are SEM.

In summary, during the task, each subject exhibited natural variability in their vigor. Given a subject's average vigor of movements for each modality, we studied within-subject changes in vigor about this mean. We found that within a subject, trial-to-trial changes in vigor of saccades were correlated with changes in vigor of both head movements and reaching movements. Specifically, vigor of all three movement modalities increased abruptly following a set break, and then decreased exponentially within-set.

5 Discussion

This thesis presents an analysis of vigor of movements in humans. Individuals tend to make movements with consistency from day to day. Moreover, there are vast differences in the vigor with which different individuals make their movements. Why do such differences exist? Moreover, do these differences relate to differences in other traits of the individual, such as impulsivity or risk sensitivity?

In Chapter 2, we showed that rate of reward has a clear effect on movement vigor. Vigor responded sharply to increases in rate of reward. In particular, saccade velocities increased and reaction times decreased. These changes in movement vigor were explained by a computational model built on the premise that the brain monitors rate of reward, and makes movements with the purpose of maximizing reward rate.

In Chapter 3, we assessed the effects of choice preference on saccade vigor. In particular, we used a task in which subjects had to make a decision between two monetary rewards: a smaller amount to be received immediately and a larger counterpart after 30 days. We found that saccade vigor served as a marker of the decision making process: in particular, vigor dropped sharply once subjects made a decision. One interpretation of the drop in vigor is a response to a drop in utility of movements: once the decision was made, movements could not acquire any new information until the next trial. We also found that, around the time of decision, saccade vigor encoded choice preference. On average, as the subjective value of the options increased, the movement vigor to those same options also increased.

Finally, in Chapter 4, we assessed movement vigor across three modalities: saccades, head movements, and reaching movements. We found that vigor was conserved across head and reaching movements. That is, those individuals who exhibited the most vigorous reaching movements also had the most vigorous head movements. Saccade vigor was not related to vigor of the other two modalities. We also demonstrated a synchronization in temporal changes of movement vigor within the individual. At set breaks, the vigor of all three movement modalities tended to increase, with a subsequent decrease during each set.

Taken as a whole, these findings suggest that movement vigor is a trait of the individual. Moreover, this work has implications for the neural circuitry that governs movement vigor in primates. This circuit must be at a level capable of affecting multiple movement modalities (i.e. basal ganglia, frontal cortex, and parietal cortex). Future neurophysiological and image-based work will further elucidate the neural underpinnings of movement vigor.

6 References

- [1] R. J. Leigh and D. S. Zee, *The neurology of eye movements*, 5 ed., New York, NY: Oxford University Press, 2015.
- [2] A. T. Bahill, M. R. Clark and L. Stark, "The main sequence, a tool for studying human eye movements," *Mathematical Biosciences*, vol. 24, pp. 191-204, 1975.
- [3] C. M. Harris and D. M. Wolpert, "Signal-dependent noise determines motor planning," *Nature*, vol. 394, pp. 780-784, 1998.
- [4] R. J. van Beers, "The sources of variability in saccadic eye movements," *Journal of Neuroscience*, vol. 27, pp. 8757-8770, 2007.
- [5] Y. Takikawa, R. Kawagoe, H. Itoh, H. Nakahara and O. Hikosaka, "Modulation of saccadic eye movements by predicted reward outcome," *Experimental Brain Research*, vol. 142, pp. 284-291, 2002.
- [6] M. Xu-Wilson, D. S. Zee and R. Shadmehr, "The intrinsic value of visual information affects saccade velocities," *Experimental Brain Research*, vol. 196, pp. 475-481, 2009.
- [7] S. G. Manohar, T. T.-J. Chong, M. A. J. Apps, A. Batla, M. Stamelou, P. R. Jarman, K. P. Bhatia and M. Husain, "Reward pays the cost of noise reduction in motor and cognitive control," *Current Biology*, vol. 25, pp. 1-10, 2015.
- [8] R. Kawagoe, Y. Takikawa and O. Hikosaka, "Expectation of reward modulates cognitive signals in the basal ganglia," *Nature Neuroscience*, vol. 1, pp. 411-416, 1998.
- [9] H. Itoh, H. Nakahara, O. Hikosaka, R. Kawagoe, Y. Takikawa and K. Aihara, "Correlation of primate caudate neural activity and saccade parameters in reward-oriented behavior," *Journal of Neurophysiology*, vol. 89, pp. 1774-1783, 2003.
- [10] T. Ikeda and O. Hikosaka, "Positive and negative modulation of motor response in primate superior colliculus by reward expectation," *Journal of Neurophysiology*, vol. 98, pp. 3163-3170, 2007.
- [11] A. Handel and P. W. Glimcher, "Quantitative analysis of substantia nigra pars reticulata activity during a visually guided saccade task," *Journal of Neurophysiology*, vol. 82, pp. 3458-3475, 1999.
- [12] O. Hikosaka and R. H. Wurtz, "Modification of saccadic eye movements by GABA-related substances. II. Effects of muscimol in monkey substantia nigra pars reticulata," *Journal of Neurophysiology*, vol. 53, pp. 292-308, 1985.
- [13] M. Sato and O. Hikosaka, "Role of primate substantia nigra pars reticulata in reward-oriented saccadic eye movement," *Journal of Neuroscience*, vol. 22, pp. 2363-2373, 2002.
- [14] M. Matsumoto and O. Hikosaka, "Lateral habenula as a source of negative reward signals in dopamine neurons," *Nature*, vol. 447, pp. 1111-1115, 2007.

- [15] Y. Tachibana and O. Hikosaka, "The primate ventral pallidum encodes expected reward value and regulates motor action," *Neuron*, vol. 76, pp. 826-837, 2012.
- [16] A. Kori, N. Miyashita, M. Kato, O. Hikosaka, S. Usui and M. Matsumura, "Eye movements in monkeys with local dopamine depletion in the caudate nucleus. II. Deficits in voluntary saccades," *Journal of Neuroscience*, vol. 15, pp. 928-941, 1995.
- [17] B. Panigrahi, K. A. Martin, Y. Li, A. R. Graves, A. Vollmer, L. Olson, B. D. Mensh, A. Y. Karpova and J. T. Dudman, "Dopamine is required for the neural representation and control of movement vigor," *Cell*, vol. 162, pp. 1418-1430, 2015.
- [18] R. S. Sutton and A. G. Barto, "Toward a modern theory of adaptive networks: expectatoin and prediction," *Psychological Review*, vol. 88, pp. 135-170, 1981.
- [19] W. Schultz, P. Dayan and P. R. Montague, "A neural substrate of prediction and reward," *Science*, vol. 275, pp. 1593-1599, 1997.
- [20] N. Schweighofer, K. Shishida, C. E. Han, Y. Okamoto, S. C. Tanaka, S. Yamawaki and K. Doya, "Humans can adopt optimal discounting strategy under real-time constraints," *PLoS Computational Biology*, vol. 2, p. e152, 2006.
- [21] J. Myerson and L. Green, "Discounting of delayed rewards: models of individual choice," *Journal of the Experimental Analysis of Behavior*, vol. 64, pp. 263-276, 1995.
- [22] S. Kobayashi and W. Schultz, "Influence of reward delays on responses of dopamine neurons," *Journal of Neuroscience*, vol. 28, pp. 7837-7846, 2008.
- [23] R. Shadmehr, J. J. Orban de Xivry, M. Xu-Wilson and T. Y. Shih, "Temporal discounting of reward and the cost of time in motor control," *Journal of Neuroscience*, vol. 30, pp. 10507-10516, 2010.
- [24] P. M. Fitts, "The information capacity of the huan motor system in controlling the amplitude of movement," *Journal of Experimental Psychology*, vol. 47, pp. 381-391, 1954.
- [25] R. A. Schmidt, H. Zelaznik, B. Hawkins, J. S. Frank and J. T. Quinn Jr., "Motor-output variability: a theory for the accuracy of rapid motor acts," *Psychological Review*, vol. 47, pp. 415-451, 1979.
- [26] D. A. Robinson, "The systems approach to the oculomotor system," *Vision Research*, vol. 26, pp. 91-99, 1986.
- [27] R. J. van Beers, "Saccadic eye movements minimize the consequences of motor noise," *PLoS ONE*, vol. 3, p. e2070, 2008.
- [28] H. Collewijn, C. J. Erkelens and R. M. Steinman, "Binocular co-ordination of human horizontal saccadic eye movements," *Journal of Physiology*, vol. 404, pp. 157-182, 1988.
- [29] N. D. Daw, "Reinforcement learning models of the dopamine system and their behavioral implications," Carnegie Mellon University, 2003.

- [30] A. Kacelnik, "Normative and descriptive models of decision making: time discounting and risk sensitivity," in *Characterizing human psychological adaptations*, Wiley, 1997, pp. 51-70.
- [31] Y. Niv, N. D. Daw, D. Joel and P. Dayan, "Tonic dopamine: opportunity costs and the control of response vigor," *Psychopharmacology*, vol. 191, pp. 507-520, 2007.
- [32] H. Chen-Harris, W. M. Joiner, V. Ethier, D. S. Zee and R. Shadmehr, "Adaptive control of saccades via internal feedback," *Journal of Neuroscience*, vol. 28, pp. 2804-2813, 2008.
- [33] K. Watanabe and O. Hikosaka, "Immediate changes in anticipatory activity of caudate neurons associated with reversal of position-reward contingency," *Journal of Neurophysiology*, vol. 94, pp. 1879-1887, 2005.
- [34] M. S. Bendiksy and M. L. Platt, "Neural correlates of reward and attention in macaque area LIP," *Neuropsychologia*, vol. 44, pp. 2411-2420, 2006.
- [35] D. M. Milstein and M. C. Dorris, "The influence of expected value on saccadic preparation," *Journal of Neuroscience*, vol. 27, pp. 4810-4818, 2007.
- [36] J. I. Gold and M. N. Shalen, "Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward," *Neuron*, vol. 36, pp. 299-308, 2002.
- [37] M. A. Segraves, M. E. Goldberg, S. Y. Deng, C. J. Bruce, L. G. Ungerleider and M. Mishkin, "The role of striate cortex in the guidance of eye movements in the monkey," *Journal of Neuroscience*, vol. 7, pp. 3040-3058, 1987.
- [38] M. M. Churchland, G. Santhanam and K. V. Shenoy, "Preparatory activity in premotor and motor cortex reflects the speed of the upcoming reach," *Journal of Neurophysiology*, vol. 96, pp. 3130-3146, 2006.
- [39] T. Ljungberg, P. Apicella and W. Schultz, "Responses of monkey dopamine neurons during learning of behavioral reactions," *Journal of Neurophysiology*, vol. 67, pp. 145-163, 1992.
- [40] C. M. Harris and D. M. Wolpert, "The main sequence of saccades optimizes speed accuracy trade-off," *Biological Cybernetics*, vol. 95, pp. 21-29, 2006.
- [41] M. Prsa, P. W. Dicke and P. Thier, "The absence of eye muscle fatigue indicates that the nervous system compensates for non-motor disturbances of oculomotor function," *Journal of Neuroscience*, vol. 30, pp. 15834-15842, 2010.
- [42] L. H. Snyder, J. L. Calton, A. R. Dickinson and B. M. Lawrence, "Eye-hand coordination: saccades are faster when accompanied by a coordinated arm movement," *Journal of Neurophysiology*, vol. 87, pp. 2279-2286, 2002.
- [43] P. van Donkelaar, K. C. Siu and J. Walterschied, "Saccadic output is influenced by limb kinetics during eye-hand coordination," *Journal of Motor Behavior*, vol. 36, pp. 245-252, 2004.
- [44] I. Opris, M. Lebedev and R. J. Nelson, "Motor planning under unpredictable reward: modulations of movement vigor and primate striatum activity," *Frontiers in Neuroscience*, vol. 5, pp. 1-12, 2011.

- [45] P. Simen, D. Contreras, C. Buck, P. Hu, P. Holmes and J. D. Cohen, "Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 35, pp. 1865-1897, 2009.
- [46] P. Cisek, G. A. Puskas and S. El-Murr, "Decisions in changing conditions: the urgency-gating model," *Journal of Neuroscience*, vol. 29, pp. 11560-11571, 2009.
- [47] B. Y. Hayden and M. L. Platt, "Temporal discounting predicts risk sensitivity in rhesus macaques," *Current Biology*, vol. 17, no. 1, pp. 49-53, 2007.
- [48] P. Mazzoni, A. Hristova and J. W. Krakauer, "Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation," *Journal of Neuroscience*, vol. 27, pp. 7105-7116, 2007.
- [49] T. V. Maia and M. J. Frank, "From reinforcement learning models to psychiatric and neurological disorders," *Nature Neuroscience*, vol. 14, pp. 154-162, 2011.
- [50] K. Sackaloo, E. Strouse and M. S. Rice, "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*, vol. 35, pp. 81-88, 2015.
- [51] K. M. Lempert, P. W. Glimcher and E. A. Phelps, "Emotional arousal and discount rate in intertemporal choice are reference dependent," *Journal of Experimental Psychology: General*, vol. 144, pp. 366-373, 2015.
- [52] P. J. Rousseeuw and C. Croux, "Alternatives to the median absolute deviation," *Journal of the American Statistical Association*, vol. 88, pp. 1273-1283, 1993.
- [53] J. E. S. Choi, P. A. Vaswani and R. Shadmehr, "Vigor of movements and the cost of time in decision making," *Journal of Neuroscience*, vol. 34, pp. 1212-1223, 2014.
- [54] J. E. Mazur, "An adjusting procedure for studying delayed reinforcement," in *The effect of delay and of intervening events on reinforcement value*, Hillsdale, New Jersey, USA, Lawrence Erlbaum Associates, Inc., 1987, pp. 55-73.
- [55] L. Green and J. Myerson, "A discounting framework for choice with delayed and probabilistic rewards," *Psychological Bulletin*, vol. 130, pp. 769-792, 2004.
- [56] J. W. Kable and P. W. Glimcher, "The neural correlates of subjective value during intertemporal choice," *Nature Neuroscience*, vol. 10, pp. 1625-1633, 2007.
- [57] A. Arieli, Y. Ben-Ami and A. Rubinstein, "Tracking decision makers under uncertainty," *American Economic Journal: Microeconomics*, vol. 3, pp. 68-76, 2011.
- [58] A. Rangel and J. Clithero, "The computation of stimulus values in simple choice," in *Neuroeconomics: Decision making and the brain*, London, UK, Elsevier Inc., 2013, pp. 125-147.
- [59] S. Kobayashi, K. Nomoto, M. Watanabe, O. Hikosaka, W. Schultz and M. Sakagami, "Influences of rewarding and aversive outcomes on activity in macaque lateral prefrontal cortex," *Neuron*, vol. 51, pp. 861-870, 2006.

- [60] K. Louie and P. W. Glimcher, "Separating value from choice: delay discounting activity in the lateral intraparietal area," *Journal of Neuroscience*, vol. 30, pp. 5498-5507, 2010.
- [61] D. Thura, I. Cos, J. Trung and P. Cisek, "Context-dependent urgency influences speed-accuracy trade-offs in decision-making and movement execution," *Journal of Neuroscience*, vol. 34, pp. 16442-16454, 2014.
- [62] D. Thura and P. Cisek, "Deliberation and commitment in the premotor and primary motor cortex during dynamic decision making," *Neuron*, vol. 81, pp. 1401-1416, 2014.
- [63] A. M. Haith, T. R. Reppert and R. Shadmehr, "Evidence for hyperbolic temporal discounting of reward in control of movements," *Journal of Neuroscience*, vol. 32, pp. 11727-11736, 2012.
- [64] C. Hyten, G. J. Madden and D. P. Field, "Exchange delays and impulsive choice in adult humans," *Journal of the Experimental Analysis of Behavior*, vol. 62, pp. 225-233, 1994.
- [65] D. Navarick, "Discounting of delayed reinforcers: measurement by questionnaires versus operant choice procedures," *Psychological Record*, vol. 54, pp. 85-94, 2010.
- [66] G. Ainslie, "Specious reward: A behavioral theory of impulsiveness and impulse control," *Psychological Bulletin*, vol. 82, pp. 463-496, 1975.
- [67] A. Bechara, "Decision making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective," *Nature Neuroscience*, vol. 8, pp. 1458-1463, 2005.
- [68] T. Nakamura, R. Kanayama, R. Sano, M. Ohki, Y. Kimura, M. Aoyagi and Y. Koike, "Quantitative analysis of ocular movements in Parkinson's Disease," *Acta Oto-Laryngologica*, vol. 111, pp. 559-562, 1991.
- [69] D. Weintraub, J. Koester, M. N. Potenza, A. D. Siderowf, M. Stacy, V. Voon, J. Whetteckey, G. R. Wunderlich and A. E. Lang, "Impulse control disorders in Parkinson disease: a cross-sectional study of 3090 patients," *Archives of Neurology*, vol. 67, no. 5, pp. 589-595, 2010.
- [70] E. G. Freedman, "Coordination of the eyes and head during visual orienting," *Experimental Brain Research*, vol. 190, pp. 369-387, 2008.