

Control of movements and temporal discounting of reward

Reza Shadmehr

Department of Biomedical Engineering

Johns Hopkins School of Medicine

To appear in: *Current Opinion in Neurobiology*

Abstract

Let us assume that the purpose of any movement is to position our body in a more rewarding state. People and other animals discount future reward as a function of time. Recent results suggest that there is a correlation between changes in this reward temporal discount function and changes in saccadic velocity and duration. These results suggest that each movement carries a cost because its duration delays acquisition of reward. The value that the brain assigns to the stimulus, and the rate at which it discounts this value in time, form a cost that appears to influence the motor commands that move our body.

Introduction

When you view a work of art, the face of a friend, or read this text, your brain shifts your gaze from one point to another, rapidly moving your eyes. Each movement is a saccade that positions your eyes so that the fovea can sample the visual scene. In making these saccades, your brain solves two kinds of problems: first, it selects where to look, and next, it programs the motor commands so your eyes move to that location. Until recently, these two problems were thought to be independent. However, recent results suggest that the goal of the saccade affects how the brain programs the motor commands to achieve the goal. For example, saccade velocities are higher when the goal of a saccade is a face vs. other images [1], velocities are higher when the goal is to touch the target vs. only look at it [2-4], and velocities are higher when one is explicitly rewarded to look at a target [5]. A recent theory [6] considered these observations and proposed a link between the reward system and the motor system. The principal idea of this new theory is that the motor commands that move our body are a reflection of an economic decision regarding reward and effort. This link has the potential to reveal why diseases that affect the reward system, such as Parkinson's disease or schizophrenia, alter control of saccades, why young people move their eyes differently than old people, and why there are inter-species differences in how animals move their eyes during a saccade.

Why should we saccade?

The fovea is like a very high resolution camera, with densely packed neurons. The rest of our retina, however, does not have the resolution of the fovea: as we move away from the fovea on the retina, neuron density drops exponentially [7], and as a result visual acuity drops exponentially [8;9]. Therefore, if we are interested in a particular image in the visual scene, the retinal distance of that image to the fovea produces an exponential reduction in acuity. A saccade

is a movement that maximizes the acuity with which we can view an image. In fact, in vertebrates voluntary saccades occur in animals which have a fovea [10].

Where should we saccade?

The scan sequence, i.e., the locations that people fixate when they are given a visual scene, is not random. The pioneering work on this was performed by Yarbus [11], a Russian physicist turned psychologist who invented new methods to record eye movements. In thinking about how people direct movements of their eyes to examine a picture, he wrote: “It may be seen to some people that when we examine an object we must trace its outlines with our eye and, by analogy with tactile sensation, ‘palpate’ the object. Others may consider that, when looking at a picture, we scan the whole of its surface more or less uniformly with our eyes.” However, Yarbus found that motion of the eyes was neither like the motion of the hand in examining a surface, nor uniform like a scanning beam that you find on a copy machine. For example, he presented his subjects a painting by Shishkin, *Morning in the Pine Forest*, in which four black bears are playing on a fallen tree. He imagined that people might look more at parts of the image which had a lot of detail (like the intricate branches on the trees). Instead, he found that people looked at the bears. He speculated that our brain continuously assigns a value to every part of the visible space, forming a priority or salience map.

The regularity in where people tend to look would suggest consistency in how the brain assigns value to images in the visual scene: animate objects are more valuable than inanimate objects. There is now strong evidence for this idea. For example, in viewing a scene consisting of face and non-face objects, people are drawn to the face region first [12]. Hayden et al. [13] found that the opportunity to look at a person of opposite sex is a valued commodity, and this is especially true for men, for whom physical attractiveness of faces of women is a dimension along which value increases rapidly. Furthermore, evidence for a salience map has been found in the parietal cortex [14]. Taken together, it appears that each saccade is a movement with which the brain directs the fovea to a region where currently, the value is highest. The stimulus that becomes the goal of the saccade carries an intrinsic value, and because of the exponential nature of retinal acuity, the reward is attained only when the stimulus falls on the fovea.

How should the eyes move during a saccade?

Yarbus noted that during a saccade, the kinematics of the eyes were consistent across people. For example, duration of a saccade grew roughly linearly with amplitude, while peak velocity tended

to saturate as amplitude increased. One can imagine that the saturation of peak velocities arise from some limitation of the firing rates of the ocular motoneurons, or the upstream neurons that drive them. However, Harris and Wolpert [15] suggested that this feature of saccades was due to a desire to produce accurate movements (minimize endpoint variance), in a scenario in which neurons are noisy. To make this link, they introduced a cost associated with endpoint accuracy. Suppose that at time t , the state of eye is described by vector $\mathbf{x}(t)$ (representing position, velocity, etc.), motor commands are $u(t)$, and the target is a stimulus at position g (with respect to the fovea). When the saccade ends at time $t = p$, eye position $x(p)$ should coincide with where reward is, i.e., target position g . This constitutes an accuracy cost:

$$J_x = E \left[(x(p) - g)^2 \right] \quad (1)$$

In Eq. (1), $E[\]$ is the expected value operator. The cost in Eq. (1) penalizes both the endpoint bias and the endpoint variance of the movement. Because motor commands are noisy, $x(p)$ is a random variable whose variance depends on the noise in the motor commands $u(t)$. If this noise is signal dependent, e.g., its standard deviation grows with the size of the motor commands, the smallest motor commands produce the smallest endpoint variance. In particular, because of the passive dynamics of the eye, it is less costly to produce a given motor command early rather than late in the saccade (because a given noise early in the saccade can be naturally damped out by the end of the movement, but the same noise at the end of the saccade cannot be damped out). Signal dependent noise and the cost of endpoint accuracy suggested that the saturating peak velocity of saccades is probably due to a desire to produce movements that are accurate.

However, if accuracy is our only concern, then saccades should be as slow as possible. This is because the movements with the smallest motor commands will have the least bias and variance. So endpoint accuracy cannot be our only cost. Indeed, Harris and Wolpert (1998) were able to account for saccade velocities by assuming an a priori saccade duration. The critical question, however, is regarding saccade duration: why are saccades of given amplitude a particular duration?

Let us return to the idea that the purpose of a saccade is to place a valuable stimulus on the fovea. It is better to acquire something valuable now than have to wait for it and get it later. That is, the value of most things declines as a function of time. For example, college students would rather

receive \$400 now than wait for five years to receive \$1000 [16]. This implies that for young people the value of \$1000 drops to less than \$400 in five years. For older people, this value drops more slowly, and for children, the value drops more quickly [17]. Psychologists have characterized this behavior via a hyperbolic reward discount function. If α represents the value of something at present, and β is the rate at which we discount this value in time, then the value at some time t in the future is:

$$V(t) = \frac{\alpha}{1 + \beta t} \quad (2)$$

For example, when the image of a face falls on the retina, perhaps we will assign a greater value α to it than to a non-face image. Furthermore, because the value of this image declines as a function of time at a rate specified by β , perhaps the duration of the saccade p (i.e., time to acquisition of the reward via placement on the fovea) carries a cost:

$$J_p = \alpha \left(1 - \frac{1}{1 + \beta p} \right) \quad (3)$$

Eq. (3) is a cost of time, implying that the longer it takes to get the target on the fovea, the larger the cost due to devaluation of the stimulus. Therefore, accuracy cost J_x declines with movement duration, encouraging one to be lazy and make slow movements. The cost of time J_p increases with movement duration, encouraging one to end the movement as soon as possible in order to minimize the devaluation of reward. The basic idea is that saccade duration and velocities, i.e., saccade kinematics, arise from a desire to be as lazy as possible while acquiring as much reward as possible.

Reward processing and the cost of time

Eq. (2) not only is a good fit to choices that people and other animals make regarding the temporal discounting of valuable commodities, it is also a good fit to discharge of dopamine cells in the brain of monkeys that have been trained to associate visual stimuli with delayed reward [18]. Dopaminergic cells tend to give a short burst of discharge in response to a stimulus that predicts future reward, and the magnitude of this burst declines hyperbolically as a function of the expected delay to reward. This hyperbolic temporal encoding of future reward is also present in the response of neurons in the lateral intraparietal area (LIP) [19], and neurons in the dorsolateral prefrontal cortex [20]. For example, in a task in which a stimulus promises future reward, LIP neurons with receptive fields that cover the stimulus discharge before a saccade to that stimulus, and the rate of firing is a hyperbolic function of the time to future reward [19]. One interpretation

is that the discharge in LIP before a saccade represents the overall value of the upcoming saccade. In a scenario in which a movement must be made to acquire a rewarding state, the encoding of the economic value of the stimulus is a hyperbolic function of time to reward.

In Parkinson's disease (PD), many of the dopaminergic cells die. If we hypothesize that this is reflected in a devaluation of the stimulus, i.e., a smaller than normal α , then the cost of time (Eq. 3) rises more slowly in PD, resulting in saccades that have longer than normal durations, and smaller than normal velocities, particularly for large amplitude saccades. This prediction is consistent with the available data [21-24].

If an abnormally low stimulus value can produce slow saccades, then an abnormally high value should produce fast saccades. In schizophrenia, saccade velocities are faster than in healthy controls [25]. Stone et al. [26] suggested that in the striatum of schizophrenic patients, there is greater than normal dopamine synthesis. Indeed, currently available antipsychotic medications have one common feature: they block dopamine D2 receptors. Remarkably, the reward discount function (Eq. 2) in schizophrenia has a higher slope than in controls [27;28], suggesting that the value of a given stimulus is discounted more steeply as a function of time than in controls. This implies that the cost of time rises faster than normal in schizophrenia, resulting in saccades that have shorter duration and greater peak speed.

As we age, peak saccade velocity declines. Saccades of children have higher velocities than young adults [29], which in turn have higher velocities than elderly [*30]. Remarkably, as we age the slope of the temporal discount function also declines [17]. In effect, aging decreases the rate of temporal discounting of reward. The availability of dopamine in the brain also declines with age. For example, rhesus monkeys exhibit a 50% decline in dopamine concentrations in the caudate and putamen from youth to old age [31], and squirrel monkeys exhibit a 20% decline [32]. All of this is consistent with a decrease in the slope of the temporal discount function, which in principle can account for the decreases in saccade velocities.

There are large differences across species in saccade velocities. Rhesus monkeys exhibit velocities that are about twice as fast as humans [33;34]. Although there are small differences in the eye plants of monkeys and humans, such differences cannot account for the remarkably faster saccades in monkeys [6]. However, rhesus monkeys exhibit a greater temporal discount rate: when making a choice between stimuli that promise water over a range of 1-20 seconds, thirsty

rhesus monkeys [18] exhibit discount rates that are many times that of thirsty humans [*35]. This translates into a cost of time that rises much faster than in humans, resulting in markedly faster saccades in monkeys as compared to humans.

Movement vigor and dopamine

Movement speeds are not only affected by the value of the goal stimulus, but also by the subject's global motivational state. For example, Niv et al. [36] noted that "hungrier rats are more jumpy, performing all actions at a faster pace." They focused on the tonic discharge of dopamine neurons and suggested that this baseline level of discharge encodes the long-term average reward available per unit of time. When this rate is high, the opportunity cost associated with taking time to move is increased and faster movements are preferred. An animal with higher baseline levels of dopamine would therefore be more active and perform all actions more vigorously. In comparison, phasic discharge encodes the temporally discounted value of the current stimulus [18]. Regardless of the general motivational state, a stimulus that is valued more might produce a faster movement than one that is valued less. Indeed, a smaller phasic discharge of dopamine neurons precedes a slow reaching movement toward a food reward, whereas a larger discharge precedes a fast reaching movement toward the same reward (tables 1 and 2 of Ljungberg et al. [37]). It is possible that tonic dopamine sets a baseline for reward per unit of time as applied for all actions, while phasic dopamine sets the reward per unit of movement time for the specific stimulus that affords the upcoming movement.

State dependent value of a stimulus

Animals do not assign a value to a stimulus based on its inherent properties, but based on their own state when the stimulus was encountered. For example, birds that are initially trained to obtain equal rewards after either large or small effort, and are then offered a choice between the two rewards without the effort, generally choose the reward previously associated with the greater effort [38]. This paradoxical result can be understood in terms of a greater utility (i.e., relative usefulness, rather than absolute value) for the reward that was attained following a more effortful action. This phenomenon is called state-dependent valuation learning, and is present in a wide variety of species from mammals to invertebrates [39].

The state-dependent valuation of stimuli allows us to consider a curious fact: kinematics of saccades to target of a reaching movement are affected by the load on the arm. For example, the peak speed of a saccade is higher when there is a load that resists the reach, and lower when the

load assists the reach [4]. Why should varying the effort required to perform a reach to a target affect saccade velocities to that target? A reaching movement that is resisted by a load arrives at the target after a larger effort than one that is assisted. The more effortful state in which the reward is encountered favors assignment of a greater utility for that stimulus. This greater utility may contribute to a faster saccade.

Why hyperbolic discounting of reward?

The hyperbolic form of the reward discount function is favored by psychologists, whereas the exponential form is favored by economists and other theorist. Here I chose the hyperbolic form because empirically, it is a better fit to choices that animals make [16]: the exponential discounting tends to produce a poor fit to data as time to reward increases. However, in simulating saccades, the timescales are too short to allow us to dissociate between hyperbolic and exponential temporal discount functions.

An important question is why should the brain temporally discount reward hyperbolically [40]. One possibility is that our cost function (which added the cost of time to cost of effort and accuracy) is really a consequence of a more fundamental normative law: actions are performed in such a way as to maximize reward per unit of time, while minimizing the effort expended per unit of time. That is, what matters is the rate of reward, which is the difference between the reward that we hope to attain and the effort we expect to expend, per unit of time. It remains to be seen whether maximizing the rate of reward is sufficient to explain movement patterns in biology, as well as the economic decision making processes that are reflected in temporal discounting of reward.

Conclusions

Suppose that the objective of any voluntary movement is to place the body at a more valuable state. Further suppose that the value associated with this state is not static, but is discounted in time: we would rather receive the reward now than later. The value that one assigns a stimulus, and the rate at which this value declines, forms a reward temporal discount function. The temporal discounting of reward forms an implicit cost of time, i.e., a penalty for the duration of the movement. This penalty is a hyperbolic function. If one assumes that motor commands are programmed to minimize effort while maximizing reward, and if one further assumes that reward loses value hyperbolically as a function of movement duration, then one can mathematically reproduce the relationship between movement duration, amplitude, and velocity in saccades.

Saccade kinematics can change due to development, due to disease, and due to evolution. The reward temporal discount function is also affected by disease, is affected by development and aging, and is different across species. In principle, a change in the temporal discount function will produce a change in movements. Indeed, there is a correlation between changes in the reward temporal discount function and changes in duration and velocities of saccades as assayed in various diseases, in various age groups, and in various species. This correlation raises the possibility that the motor commands that move our body during goal-directed behavior reflect a specific cost of time; one in which passage of time discounts reward.

Acknowledgements

This work was supported by grants from the NIH (NS057814 and EY019581). I am grateful to Adrian Haith, who came up with the idea of maximizing reward minus effort per unit of time as a normative way of formulating hyperbolic temporal discounting.

Reference List

1. Xu-Wilson M, Zee DS, Shadmehr R: The intrinsic value of visual information affects saccade velocities. *Exp.Brain Res.* 2009, 196:475-481.
2. Epelboim J, Steinman RM, Kowler E, Pizlo Z, Erkelens CJ, Collewijn H: Gaze-shift dynamics in two kinds of sequential looking tasks. *Vision Res.* 1997, 37:2597-2607.
3. Snyder LH, Calton JL, Dickinson AR, Lawrence BM: Eye-hand coordination: saccades are faster when accompanied by a coordinated arm movement. *J.Neurophysiol.* 2002, 87:2279-2286.
4. van Donkelaar P, Siu KC, Walterschied J: Saccadic output is influenced by limb kinetics during eye-hand coordination. *J.Mot.Behav.* 2004, 36:245-252.
5. Takikawa Y, Kawagoe R, Itoh H, Nakahara H, Hikosaka O: Modulation of saccadic eye movements by predicted reward outcome. *Exp.Brain Res.* 2002, 142:284-291.
6. Shadmehr R, Orban de Xivry JJ, Xu-Wilson M, Shih TY: Temporal discounting of reward and the cost of time in motor control. *J.Neurosci.* 2010, 30:10507-10516.
* This paper hypothesizes that the value that the brain assigns to the stimulus, and the rate at which it discounts this value in time, form a cost that influences the motor commands that move our body.
7. Hirsch J, Curcio CA: The spatial resolution capacity of human foveal retina. *Vision Res.* 1989, 29:1095-1101.
8. Green DG: Regional variations in the visual acuity for interference fringes on the retina. *J.Physiol* 1970, 207:351-356.
9. Jacobs RJ: Visual resolution and contour interaction in the fovea and periphery. *Vision Res.* 1979, 19:1187-1195.
10. Walls GL: The evolutionary history of eye movements. *Vision Res.* 1962, 2:69-80.
11. YARBUS AL: Eye movements and vision. New York: Plenum Press; 1967.
12. Cerf M, Harel J, Einhasuer W, Koch C: Predicting human gaze using low-level saliency combined with face detection. In *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press; 2008.
13. Hayden BY, Parikh PC, Deaner RO, Platt ML: Economic principles motivating social attention in humans. *Proc.Biol Sci.* 2007, 274:1751-1756.
14. Gottlieb JP, Kusunoki M, Goldberg ME: The representation of visual salience in monkey parietal cortex. *Nature* 1998, 391:481-484.
15. Harris CM, Wolpert DM: Signal-dependent noise determines motor planning. *Nature* 1998, 394:780-784.
16. Myerson J, Green L: Discounting of delayed rewards: Models of individual choice. *J.Exp.Anal.Behav.* 1995, 64:263-276.
17. Green L, Myerson J, Ostaszewski P: Discounting of delayed rewards across the life span: age differences in individual discounting functions. *Behavioural Processes* 1999, 46:89-96.

18. Kobayashi S, Schultz W: Influence of reward delays on responses of dopamine neurons. *J.Neurosci.* 2008, 28:7837-7846.

** The authors measured the discharge of dopamine cells in the brain of monkeys that had been trained to associate visual stimuli with delayed reward. The magnitude of the discharge in response to the stimulus was a hyperbolic function of the time delay to future reward. The behavioral reward temporal discount function, estimated from the choices that the animals made, was also hyperbolic in time. Among the two animals, the one that was behaviorally more impulsive also had dopamine discharge that declined more steeply.

19. Louie K, Glimcher PW: Separating value from choice: delay discounting activity in the lateral intraparietal area. *J.Neurosci.* 2010, 30:5498-5507.

* Some monkeys are more impulsive than others. In this paper, the authors measured discharge of cells in the lateral intraparietal area (LIP) and found that during a period before saccade onset, cells discharged according to the temporally discounted value of the saccade stimulus. This discounting was hyperbolic, and the discharge was steeper in the impulsive animal. Therefore, the subjective value of the stimulus is reflected in the LIP discharge before the saccade.

20. Kim S, Hwang J, Lee D: Prefrontal coding of temporally discounted values during intertemporal choice. *Neuron* 2008, 59:161-172.

21. Shibasaki H, Tsuji S, Kuroiwa Y: Oculomotor abnormalities in Parkinson's disease. *Arch.Neurol.* 1979, 36:360-364.

22. White OB, Saint-Cyr JA, Tomlinson RD, Sharpe JA: Ocular motor deficits in Parkinson's disease. II. Control of the saccadic and smooth pursuit systems. *Brain* 1983, 106 (Pt 3):571-587.

23. Nakamura T, Kanayama R, Sano R, Ohki M, Kimura Y, Aoyagi M, Koike Y: Quantitative analysis of ocular movements in Parkinson's disease. *Acta Otolaryngol.Suppl* 1991, 481:559-562.

24. Blekher T, Siemers E, Abel LA, Yee RD: Eye movements in Parkinson's disease: before and after pallidotomy. *Invest Ophthalmol.Vis.Sci.* 2000, 41:2177-2183.

25. Mahlberg R, Steinacher B, Mackert A, Flechtner KM: Basic parameters of saccadic eye movements--differences between unmedicated schizophrenia and affective disorder patients. *Eur.Arch.Psychiatry Clin.Neurosci.* 2001, 251:205-210.

26. Stone JM, Morrison PD, Pilowsky LS: Glutamate and dopamine dysregulation in schizophrenia--a synthesis and selective review. *J.Psychopharmacol.* 2007, 21:440-452.

27. Heerey EA, Robinson BM, McMahon RP, Gold JM: Delay discounting in schizophrenia. *Cogn Neuropsychiatry* 2007, 12:213-221.

28. Kloppel S, Draganski B, Golding CV, Chu C, Nagy Z, Cook PA, Hicks SL, Kennard C, Alexander DC, Parker GJ, Tabrizi SJ, Frackowiak RS: White matter connections reflect changes in voluntary-guided saccades in pre-symptomatic Huntington's disease. *Brain* 2008, 131:196-204.

29. Munoz DP, Armstrong IT, Hampton KA, Moore KD: Altered control of visual fixation and saccadic eye movements in attention-deficit hyperactivity disorder. *J.Neurophysiol.* 2003, 90:503-514.

30. Irving EL, Steinbach MJ, Lillakas L, Babu RJ, Hutchings N: Horizontal saccade dynamics across the human life span. *Invest Ophthalmol.Vis.Sci.* 2006, 47:2478-2484.

* Saccade kinematics change with development and aging. Very young children make fairly slow saccades, but peak speeds increase as children reach the teen years. Teenagers make saccades that have the highest velocities of any age. However, within fifteen additional years the peak velocities drop significantly. This decline continues until peak velocities reach their lowest values in the 8th and 9th decade of life.

31. Collier TJ, Lipton J, Daley BF, Palfi S, Chu Y, Sortwell C, Bakay RA, Sladek JR, Jr., Kordower JH: Aging-related changes in the nigrostriatal dopamine system and the response to MPTP in nonhuman primates: diminished compensatory mechanisms as a prelude to parkinsonism. *Neurobiol.Dis.* 2007, 26:56-65.

32. McCormack AL, Di Monte DA, Delfani K, Irwin I, DeLanney LE, Langston WJ, Janson AM: Aging of the nigrostriatal system in the squirrel monkey. *J.Comp Neurol.* 2004, 471:387-395.

33. Straube A, Fuchs AF, Usher S, Robinson FR: Characteristics of saccadic gain adaptation in rhesus macaques. *J Neurophysiol.* 1997, 77:874-895.

34. Chen-Harris H, Joiner WM, Ethier V, Zee DS, Shadmehr R: Adaptive control of saccades via internal feedback. *J.Neurosci.* 2008, 28:2804-2813.

35. Jimura K, Myerson J, Hilgard J, Braver TS, Green L: Are people really more patient than other animals? Evidence from human discounting of real liquid rewards. *Psychon.Bull.Rev.* 2009, 16:1071-1075.

* Comparison of temporal discounting of reward is difficult between humans and other species because humans are typically offered money, whereas other animals are offered food or water. In this paper, the authors considered temporal discounting in humans for water reward, allowing more direct comparison with other animals.

36. Niv Y, Daw ND, Joel D, Dayan P: Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 2007, 191:507-520.

37. Ljungberg T, Apicella P, Schultz W: Responses of monkey dopamine neurons during learning of behavioral reactions. *J.Neurophysiol.* 1992, 67:145-163.

38. Clement TS, Feltus JR, Kaiser DH, Zentall TR: Work ethic in pigeons: reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin Review* 2000, 7:100-106.

39. Pompilio L, Kacelnik A: Context-dependent utility overrides absolute memory as a determinant of choice. *Proc.Natl.Acad.Sci.U.S.A* 2010, 107:508-512.

40. Kacelnik A: Normative and descriptive models of decision making: time discounting and risk sensitivity. In *Characterizing human psychological adaptations*. Chichester: Wiley; 1997:51-70.