



The temporal derivative of expected utility: A neural mechanism for dynamic decision-making

Xian Zhang ^{a,d,*}, Joy Hirsch ^{a,b,c,d}

^a Department of Psychology, Columbia University, USA

^b Department of Neuroscience, Columbia University Medical Center, USA

^c Department of Radiology, Columbia University Medical Center, USA

^d fMRI Research Center, Columbia University Medical Center, USA

ARTICLE INFO

Article history:

Accepted 21 August 2012

Available online 30 August 2012

Keywords:

Expected utility

Temporal derivative of expected utility

Dynamic game

fMRI and decision-making

ABSTRACT

Real world tasks involving moving targets, such as driving a vehicle, are performed based on continuous decisions thought to depend upon the temporal derivative of the expected utility ($\partial V/\partial t$), where the expected utility (V) is the effective value of a future reward. However, the neural mechanisms that underlie dynamic decision-making are not well understood. This study investigates human neural correlates of both V and $\partial V/\partial t$ using fMRI and a novel experimental paradigm based on a pursuit–evasion game optimized to isolate components of dynamic decision processes. Our behavioral data show that players of the pursuit–evasion game adopt an exponential discounting function, supporting the expected utility theory. The continuous functions of V and $\partial V/\partial t$ were derived from the behavioral data and applied as regressors in fMRI analysis, enabling temporal resolution that exceeded the sampling rate of image acquisition, hyper-temporal resolution, by taking advantage of numerous trials that provide rich and independent manipulation of those variables. V and $\partial V/\partial t$ were each associated with distinct neural activity. Specifically, $\partial V/\partial t$ was associated with anterior and posterior cingulate cortices, superior parietal lobule, and ventral pallidum, whereas V was primarily associated with supplementary motor, pre and post central gyri, cerebellum, and thalamus. The association between the $\partial V/\partial t$ and brain regions previously related to decision-making is consistent with the primary role of the temporal derivative of expected utility in dynamic decision-making.

© 2012 Elsevier Inc. All rights reserved.

Introduction

Expected utility (V) is the product of the probability and subjective utility of a goal. Initiated by Bernoulli (1738) and formalized by Morgenstern and Von Neumann (1944), expected utility has been a central concept in decision science. This paper investigates the role of temporal change in V , or its temporal derivative ($\partial V/\partial t$) during decision-making. Many real-world tasks can be modeled as a dynamic game, such as a pursuit–evasion game, where decisions and actions are made continuously, not only responding to, but also altering, the game state. According to dynamic decision theory (Isaacs, 1965), one of the most important decision principles in a dynamic game is maximization of $\partial V/\partial t$ (Appendix 1). The importance of temporal change in V has also been implicated in prominent theories such as the temporal difference

algorithm (TD) (Sutton and Barto, 1987), and the prediction error theory (Schultz et al., 1997). Further, single unit data have shown that the activation pattern of some dopamine neurons are similar to $\partial V/\partial t$ (Figure 1 of Schultz et al., 1997). It has been hypothesized (Figure 4 of Schultz et al., 1997), but not demonstrated, that the temporal derivative of utility is coded in the nervous system as a mechanism to enable immediate and reflexive responses. This paper presents a paradigm that allows both V and $\partial V/\partial t$ to be quantified and their neural correlates observed.

Currently, decision-making is most commonly studied with static paradigms, such as a conventional event-related design, where each trial consists of discrete events such as a stimulus, a response and a reward event. Those paradigms are “static” because they consist of discrete decisions where the continuous $\partial V/\partial t$ is not well defined (Basar and Olsder, 1999). Due to the limited temporal resolution of functional magnetic imaging (fMRI) and the absence of independent manipulations of V and $\partial V/\partial t$, it is difficult to separate $\partial V/\partial t$ from V for human subject studies using conventional paradigms. On the other hand, single unit recordings not only describe a variety of temporal profiles of neuronal action potentials but also allow a measure of the neural activity related to the expected utility, prior to the receiving of the reward, and the response to the actual reward separately. However, single unit

Abbreviations: V , expected utility; $\partial V/\partial t$, the temporal derivative of the expected utility; SMA, supplementary motor area; MFC, medial frontal cortex; ACC, anterior cingulate; PCC posterior, cingulate; VP, ventral pallidum; CPU, caudate-putamen; NAC, nucleus accumbens; CN, caudate nucleus; SPL, superior parietal lobule.

* Corresponding author at: fMRI Research center of Columbia University Medical Center, 710 W. 168th Street, New York, NY 10032, USA.

E-mail addresses: xz63@columbia.edu (X. Zhang), Joyhirsch@yahoo.com (J. Hirsch).

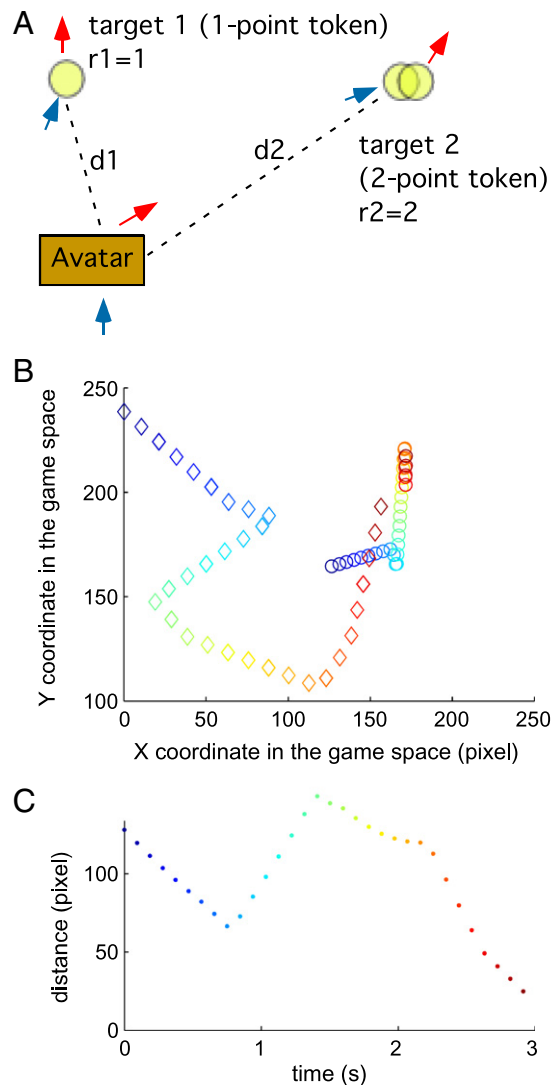


Fig. 1. A. An example of the target-choice task. Based on the values, $r1$ and $r2$, and the distances, $d1$ and $d2$, the player makes a choice, which, in this case, is to pursue target 2. The blue arrows indicate the directions of movement prior to the event. The red arrows indicate the new directions of the movement as the result of the actions of the characters. B. A pursuit process. The diamond and the circles represent the location of the avatar and the target, respectively. The time is coded in color, where the blue and red colors indicate the starting and the ending time points, respectively. Each dot represents a 100-ms period. The distance between the avatar and the target is defined with dots of identical color, i.e. at the same time. Note that, due to the random perturbation of the game program, not all actions appear optimal. C. The game state is represented as the distance between the avatar and the target. Each dot represents the distance between the diamond and the circle shown in Fig. 1B with identical color, or at the same time.

recordings are invasive and only sample a few neurons and thus do not provide the description of the global neural networks related to the V , $\partial V/\partial t$ and reward.

To meet those challenges, we adopted a pursuit–evasion game, the classic game that was used for developing the dynamic game theory (Isaacs, 1965), to determine both V and $\partial V/\partial t$ as continuous functions of the game states, as well as the capture event as an impulse function. The underlying neural activities associated with those dynamic variables were isolated using fMRI. Our hypothesis is that brain activity related to dynamic decision-making would be correlated with $\partial V/\partial t$ and distinguishable from neural activity associated with V .

Materials and methods

Task

The task was a continuous pursuit–evasion video game played during fMRI scans. The game was modeled after the familiar Pacman game and

subjects aimed to collect 1-point and 2-point rewards and to avoid 2-point losses, which were treated as gaining negative 2 points. All the characters moved in continuous game space. The character corresponding to the “pellet” as in the Pacman game could also move away from a predator. This computer game adopted a first person’s viewpoint. In other words, the avatar of the player was always shown at the center of the monitor. See Appendix 2.1–2.4 for the rules and detailed description of the game. The avatar was controlled by the player with actions such as turning up, down, left and right using an MRI compatible track ball, while other characters were controlled by the game program. The goal of the player was to accumulate a maximum number of points, the unit of utility, by capturing targets. Both the controls and the movements of all the game characters, including the avatar, were modeled as a vehicle with constant speed and minimum turning radius. The speed of a pursuer was 15% greater than that of an evader. Randomly, the game program generated perturbation to the decisions of game characters, modeled as the effect of random wind gusts as on a sail-boat (Appendix 2.5). Such perturbation is important. First, for the player, it simulates the unpredictable nature of the environment. Second, it simulates the

random strategy of a computer character. Finally and most importantly, without the perturbation, every pursuit process would be identical, trivially easy and thus failing to provide a rich and independent manipulation of V and $\partial V/\partial t$. In fact, without perturbation, V and $\partial V/\partial t$ would be highly correlated because V is an exponential function.

The continuous game can be divided into two basic phases. First, the avatar faces two targets with comparable expected utility and the player performs a target-choice task (Fig. 1A). Second, when there is no comparable alternative target, the player pursues the chosen target. Fig. 1B represents a dynamic pursuit process in the 2-dimensional game space, where the time is coded in color. V , for the first approximation, depends on the distance between the avatar and the target (Fig. 1C), as well as the value of the target. In summary, this game represents a departure from a conventional trial structure where each event is isolated. Rather, the continuous acquisition of V enables a disassociation among V , $\partial V/\partial t$ and the capture event.

Subjects

Subjects were recruited from the Columbia University Medical Center community and provided written informed consent according to the guidelines of the local institutional review board. Twenty-one subjects (7 females, age = 26 ± 8 years) volunteered for the study and participated without compensation. They were given an instruction (Appendix 2.6) and practiced in the scanner prior to image acquisition.

fMRI scan parameters and analysis

To reduce the susceptibility-induced artifacts, we employed a spiral-in/out sequence for the fMRI data acquisition (TE = 36 ms, TR = 2000 ms, FA = 84, FOV = 21 cm, voxel size = $3 \times 3 \times 4$ mm) (Glover and Thomason, 2004). Functional magnetic resonance imaging (fMRI) images were analyzed using the general linear model for individual data

and the mixed effect analysis for the group data using the FSL software package (Smith et al., 2004).

fMRI regressors

V of the entire game state as defined by Eq. (1) was applied to the target that the player is currently pursuing. Potential targets other than the immediate one were ignored because the game was designed so that over 97% of the time, the expected utility values of other potential targets were less than half of the value of the current target. $\partial V/\partial t$ was the temporal derivative of V . In addition, following the standard event related paradigm, a capture event was defined as an impulse function at the time of the capture with the magnitude as the value of the target (Fig. 2). V , $\partial V/\partial t$ and capture event were employed as major regressors for a general linear model analysis.

Note that, the captured event in the game differs from the reward event in previous studies, where once chosen, the target will be obtained with certain probability and without further effort. The merit of the conventional approach is that the utility for the reward event is clearly defined, as the product of value and probability of the target. In our game, there is a substantial chance that the target may escape the pursuit in spite of the effort. In other words, the brain activity shown associated with the capture event could also include the activity that previously associated with the prediction error, because a capture event provides information that was uncertain before.

Motion correction and global mean were included as confounding regressors along with the major regressors, including V , $\partial V/\partial t$ and the capturing event described in following sections. In addition, 30% of the capturing events were associated with unexpected outcomes. For example, capturing a 1-point token could result in either a gain of 2 points or a gain of 0 points (see Appendix 3 and Table A3 for a detailed description and Fig. A2 for the fMRI result). Both the capture event and the unexpected payoff regressors included all types of targets, combining both

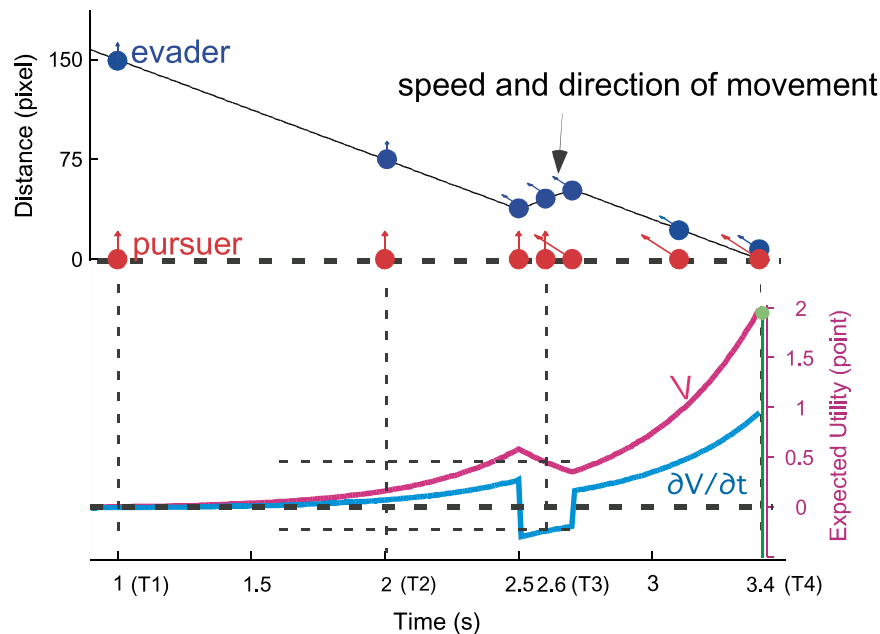


Fig. 2. Schematic of a pursuit epoch. The solid black line in the upper panel indicates the game state, similar to Fig. 1C plotted against time. Both the relative location and the direction of movement of the avatar (red) and the target, which is a 2-point token, (blue) are shown at certain time points. The speed and the direction of movement at a time point are represented as the length and the angle, respectively, of the arrow attached to the dot. Between 1 and 2.5 s, the distance shortens because the avatar is faster. At 2.5 s, the target changed its direction but the avatar did not. As a consequence, the separation distance became greater. The avatar corrected the direction at 2.7 s. Finally, the avatar captured the target at 3.4 s and received the reward of 2 units of utility (points). In the lower panel, V (magenta) and $\partial V/\partial t$ (cyan) represent the expected utility and its temporal derivative calculated according to the game state presented in the upper panel. The function of V is determined as an exponential discounting function, based on the behavioral target-choice data. The capture event is represented as an impulse function (green bar) at T4 with the magnitude of the value of the target.

positive (capturing a target) and negative (being captured by the opponent) events.

All regressors were convoluted with the canonical hemodynamic response function, except for the motion correction parameters and the global mean. The orthogonality among all the regressors was confirmed using a correlation measure (Hare et al., 2008). The largest correlation coefficient among all the regressors in this task was 0.03, indicating the independence of the variables for the fMRI analysis.

Task description and theoretical basis for determining the function of expected utility

The role of $\partial V/\partial t$ is illustrated with a typical epoch in a pursuit–evasion game in Fig. 2. During the pursuit process, a sudden change in the direction of movement of the target requires the pursuer, the avatar, to make a decision. Note that, at time $T3$, the direction of movement was non-optimal because the avatar failed to respond to the change. In contrast, at time $T2$, the avatar's direction of movement was optimal. However, the expected utility at $T3$, $V(T3)$, is positive and even greater than $V(T2)$. In contrast, $\partial V(T2)/\partial t$ is positive and $\partial V(T3)/\partial t$ is negative, corresponding to whether the direction of the movement is optimal. Such a situation illustrates the dissociation between V and $\partial V/\partial t$. In addition, Fig. 2 also shows that the function of V , $\partial V/\partial t$ and capturing are distinct to each other, confirming that a pursuit–evasion game may enable us to separate the brain activity specific to those components of decision-making. In summary, a neuron that codes V cannot make the decision that, at $T3$, it needs to change the direction of movement. Nor can it tell the difference when two targets with the same value and same distance are moving at different speeds. In contrast, coding $\partial V/\partial t$ provides a natural and effective way to utilize game state information.

To understand the function of V for such a complex game, methods from dynamic game theory and microeconomics were combined. First, the game scene was segmented into epochs that ended with a capture. Second, the game state, i.e. the distances between the avatar and the targets, and the action, i.e. the direction control, were converted into a continuous function of V . Since the distance is by and large linearly related to the temporal delay for the reward, we assume that the function V of distance is similar to a temporal discounting function.

Estimating the form and the parameters of the discounting function

Since the game is engineered in the way that random events occurred independently of each other, the rational function of V is expected to be an exponential discounting function (Samuelson, 1937), where the time variable is replaced by distance:

$$V(d, r) = r \times e^{-d/d_0} \quad (1)$$

where d is the distance measured at a time, r is the value of the target and d_0 , the discounting constant, is a parameter to be determined. This exponential function (Eq. (1)) will be qualitatively compared with a general hyperbolic function (Eq. (2)) commonly used in previous decision science research (Loewenstein and Prelec, 1992; Richards et al., 1997; Scholten and Read, 2010):

$$V(d, r) = r \times (1 + \alpha \times d)^{-\beta/\alpha} \quad (2)$$

where α and β are positive parameters to be determined.

The decision function

Although dynamic game theory has established the importance of the temporal derivative of the game state in decision-making, the real time decision function for human subjects in a pursuit–evasion game, to our knowledge, has not been described. Here the real time decision function, which consists of both perception and decision components, is defined as

the mathematical operation that transforms input information into action, in real time.

The game data include the x and y coordinates of all the characters (see Fig. 1B for an example). Both the input information and the avatar's action can be extracted from those data. In this game all characters moved at a constant speed. The only action was the change in the direction of the movement. Therefore, we focus on the direction aspect of the information. Other information, such as the distance, although related to the expected utility, is irrelevant for a specific action. For example, regardless of the distance, if the target is at one o'clock, the player moves the avatar to that direction. The direction of movement (*Direction*) of the avatar is obtained as:

$$Direction_t = \text{cart2pol}(X_{t+1} - X_t, Y_{t+1} - Y_t)$$

where cart2pol is the mathematical function transforming a Cartesian coordinate into an angle, t indicates a point in time, $t+1$ and t are 10 ms apart, and X, Y are the coordinates of the avatar in the 2-dimensional game world. The action of the player is measured as the change of *Direction*:

$$Action_t = Direction_{t+1} - Direction_t$$

Note that, $Action_t$ is the result of a decision responding to information received at some time prior. The direction of location between the avatar and the target (*Angle*) is defined as:

$$Angle_t = \text{Cart2pol}(x_t - X_t, y_t - Y_t)$$

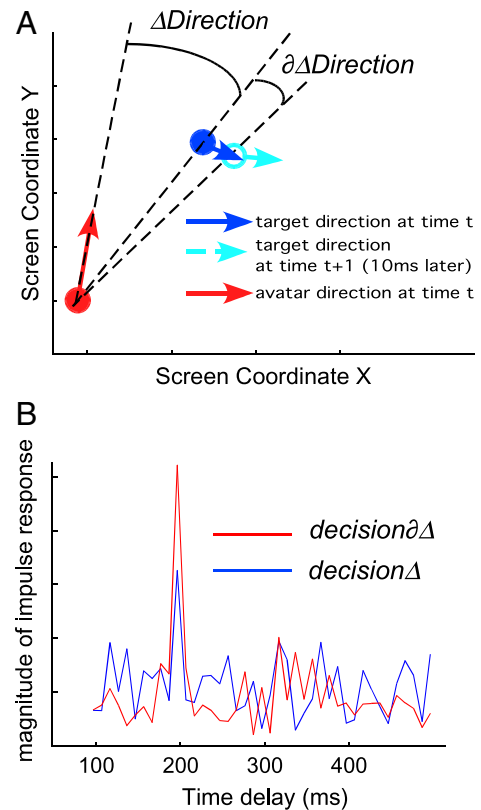


Fig. 3. A. The positions (circles) and the directions (arrows) of the avatar (red) and the target (blue) at two time points 10 ms apart. (The direction and position changes are exaggerated for illustration purposes.) The angle arcs are graphic presentations of $\Delta Direction$ and $\partial \Delta Direction$, as labeled. B. The decision functions $decision\partial\Delta$ (red) and $decision\Delta$ (blue) during the last 1.5 s of the pursuit process. The peak of the decision function at 200 ms indicates that, given an input time series, the action is similar to the input but delayed for 200 ms. The greater magnitude of the decision function related to $\partial \Delta Direction$ indicates the significance of the temporal derivative.

where x and y are the coordinates of the target. The difference between *Direction* and *Angle*, referred to as $\Delta\text{Direction}$, can be defined as:

$$\Delta\text{Direction}_t = \text{Angle}_t - \text{Direction}_t \text{ (Fig. 3A)}.$$

For example, if the avatar is moving toward three o'clock, while the target is at one o'clock, $\Delta\text{Direction}_t$ is two o'clock or 60° . According to dynamic game theory, the temporal derivative of the game state plays an important role in decision-making. If $\Delta\text{Direction}$ is an aspect

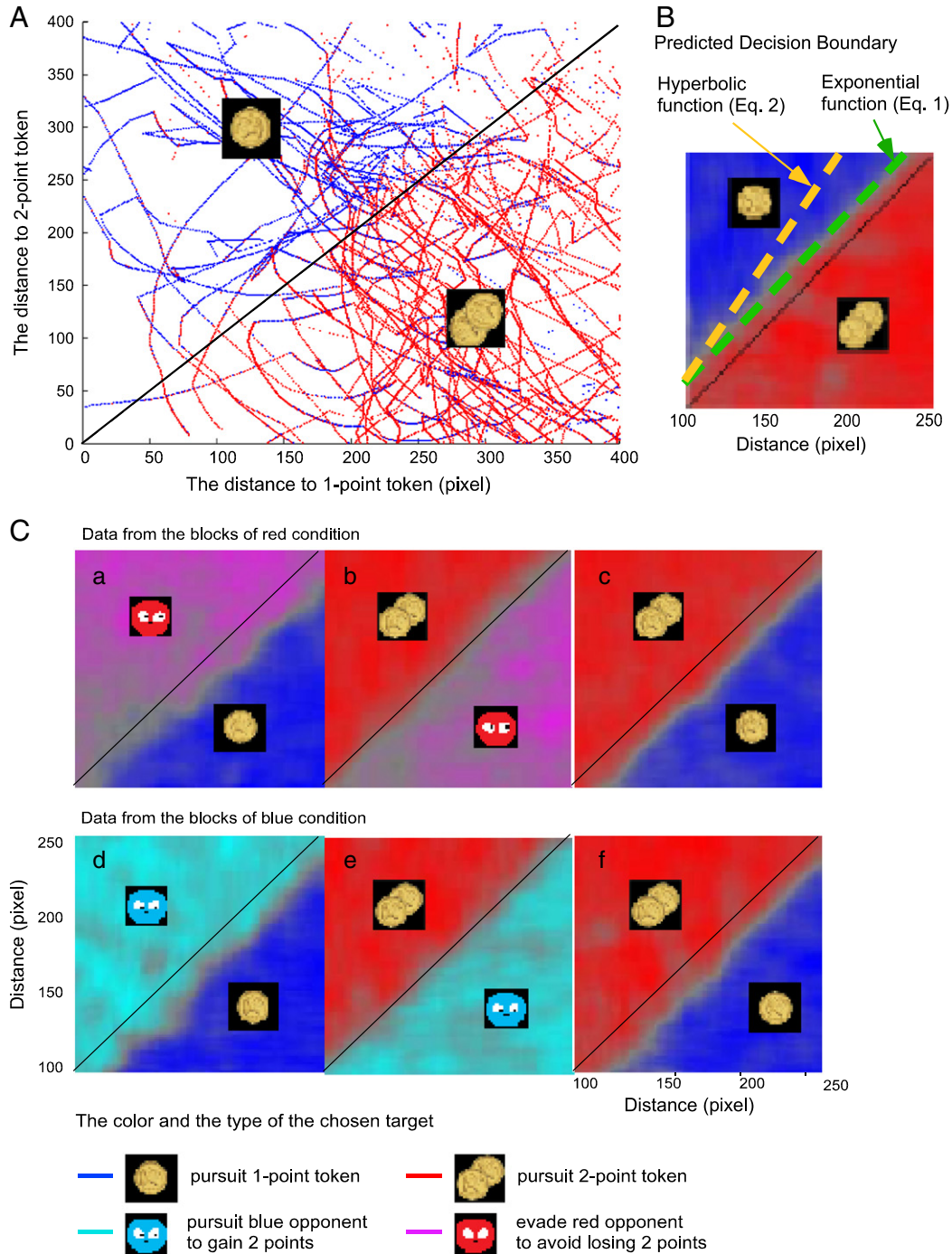


Fig. 4. A. Distance vs. Distance scatter-plot showing an example of a subject's pursuit choices as a function of distance and the value of the targets during an 8-min run. Each dot represents a 10 ms time point and the distance in pixels to the 1-point token (X axis) and the distance to the 2-point token (Y axis). Red and blue indicate that the player is pursuing the 2-point token and the 1-point token, respectively. A line represents a series of consecutive measures. If a red line reaches the X-axis, the 2-point token is captured and if a blue line reaches the Y-axis, the 1-point token is captured, i.e. the distance becomes zero. The two tokens shown inside either the upper or the lower triangle indicate the target of choice for most of the time. B. The averaged and smoothed pursuit choices for 21 subjects and 4 runs in the distance range of between 100 and 250 pixels. Saturated colors indicate that the icon shown in the region was chosen most of the time, while gray indicates the decision boundary, where the probability for choosing either target is close to 0.5. The theoretical predictions by both the exponential discounting function (green dashed line) and the hyperbolic function (orange dashed line) are also presented. C. The average pursuit-choices across 21 subjects and 4 runs. See Table A1 in the Appendix for the payoff tables related to these decisions. In short, panels e, d, c and f involve a choice between two points and one point, while panels b and e involve a choice between two options of two points. The distance ranges from 100 to 250 pixels, where the data quality is high.

of the game state, its temporal derivative ($\partial\Delta\text{Direction}$) can be defined as:

$$\partial\Delta\text{Direction}_t = \Delta\text{Direction}_{t+1} - \Delta\text{Direction}_t \text{ (Fig. 3A).}$$

Based on conventional static decision theory, the optimal action should maximize the expected utility (V) by minimizing the distance. The most intuitive strategy is to make $\Delta\text{Direction}$ zero, or $\text{Action} = \Delta\text{Direction}$, so that the avatar will be moving toward the target. This strategy is optimal when the target is stationary. However, in a dynamic pursuit process, where the target is evading, moving toward the current target position is often not effective. Instead, one should perceive the direction change of the target ($\partial\Delta\text{Direction}$) and make a move toward the anticipated position. Therefore, dynamic game theory emphasizes the importance of the temporal derivative of the game state. We hypothesize that the decision function should consist of two distinct components, one for $\Delta\text{Direction}$ and the other for $\partial\Delta\text{Direction}$, referred to as $\text{decision}\Delta$ and $\text{decision}\partial\Delta$. Since there is a time delay between the input information and the action due to reaction time, we use the convolution method to study those decision mechanisms:

$$\text{Action} = \text{decision}\Delta^* \Delta\text{Direction} + \text{decision}\partial\Delta^* \partial\Delta\text{Direction}$$

where $*$ is the symbol of convolution. For one trial, Action , $\Delta\text{Direction}$ and $\partial\Delta\text{Direction}$ are time series of 150 data points obtained at the final 1.5 s of the pursuit process, when both the player and the target are highly engaged in the process and other characters in the game are irrelevant. A simple way to obtain both decision components, $\text{decision}\Delta$ and $\text{decision}\partial\Delta$, is to assume that $\Delta\text{Direction}$ and $\partial\Delta\text{Direction}$ are approximately independent of each other:

$$\text{decision}\Delta = \text{Action}(*-1)\Delta\text{Direction}$$

$$\text{decision}\partial\Delta = \text{Action}(*-1)\partial\Delta\text{Direction},$$

where $(*-1)$ is the symbol for de-convolution.

Results

Expected utility function determined by behavioral data

In this game, the player selected a target and controlled the direction of movement for pursuit and capture the target, referred to as a token. *Pursuit-choice* denotes the object toward which the avatar is moving. Examples of *pursuit-choices* between a 1-point (the unit of utility in the game) token and a 2-point token for a typical subject are shown in Fig. 4A, and the group averaged data are shown in Fig. 4B. Note that the decision boundary (the gray area in Fig. 4B) aligns with the prediction of the exponential discounting function (green dashed line), for being parallel to the equal distance (black 45°) line. In other words, when the 1-point token and the 2-point token were chosen with equal probability, the difference between the distances (Δ) is constant regardless of the average distances for both. Data show that Δ is approximately constant regardless of the average distances for both targets for all types of decisions in the game (Fig. 4C, see Appendix 2.4 for the details for those choices). Translating distance into temporal delay, this result can be characterized as dynamic consistency, the signature of an exponential discounting function (Schweighofer et al., 2006; Strotz, 1956) (Eq. (1)), in agreement with previous findings (see Schweighofer et al., 2006) or Appendix 4 for a detailed inference). We combined data from all types of decisions in the game to compute the parameter d_0 individually for each subject. Due to the narrow distribution of d_0 , the relationship between the performance and the discounting constant is by and large independent for our data (Appendix 5).

The decision function and the temporal derivative of the game state

Fig. 3B shows both $\text{decision}\partial\Delta$ (red) and $\text{decision}\Delta$ (blue) based on combined data from all subjects and runs. Both $\text{decision}\Delta$ and $\text{decision}\partial\Delta$ contain a significant signal, indicated by the relative magnitude of the peak at 200 ms vs. the base line noise. In other words, players apparently encode both $\Delta\text{Direction}$ and $\partial\Delta\text{Direction}$ and, after a 200 ms delay, probably due to the reaction time, react following those inputs. In particular, our analysis shows that $\text{decision}\partial\Delta$ is more prominent than $\text{decision}\Delta$ during the final 1.5 s of pursuit processes, consistent with dynamic game theory. Note that the decision functions are narrow peaks, while the reaction time is expected to have a broad distribution. This is the result of the continuous and smooth nature of our input information and limitation of the de-convolution method. The narrow peak indicates the median reaction time.

$\Delta\text{Direction}$ and $\partial\Delta\text{Direction}$ represent a game state and the temporal derivative, and are assumed to be related to the V in a complex way. In subsequent analysis, we study the neural correlates for V and $\partial V/\partial t$ calculated directly with distances.

Neural correlates for V and $\partial V/\partial t$

The regressor for V co-varied with activity in the cerebellum, brainstem, thalamus, supplementary motor area (SMA), pre and postcentral gyri, regions often associated with motor functions (Fig. 5A, Table 1A). The regressor for $\partial V/\partial t$ co-varied with activity in the anterior and posterior cingulate cortices (ACC and PCC) and ventral pallidum (VP) (Fig. 5B and Table 1B), regions that have previously been related to decision-making and executive functions (McClure et al., 2004; Sarinopoulos et al., 2010). The neural activity correlated with the capture event was found in caudate-putamen and subcallosal cortex, which includes nucleus accumbens (NAC) and the caudate nucleus (CN) (Fig. 5C, Table 1C), regions that are part of the ventral striatum and previously related to reward and learning (McNamara et al., 2002). For example, it has been reported that CN showed deactivation in response to failure in a shooting game (Mathiak et al., 2011).

Discussion

The relationships between the choice and the distance of the targets during the games were best described by an exponential function used to model the relationship between the expected utility and the temporal delay and thus, are consistent with the classic expected utility theory and previous observations. This finding does not necessarily contradict more frequently reported results supporting a hyperbolic discounting function. In our game, as well as in real world, random events, such as the emergence of either an unexpected better choice or an unexpected danger; can render a chosen target irrelevant. Therefore, the effective value of a target decreases exponentially with time. In contrast, in an inter-temporal choice paradigm, even the most delayed target has a greater than zero chance to be obtained. As a consequence, along the time dimension, the expected utility asymptotes to a fixed value and thus assumes a hyperbolic function. Therefore, it is possible that the difference in the form of the discounting function may reflect similar decision principles applied to different situations.

The dynamic game paradigm and the behavioral data that are used for generating regressors allow dissociation of important components of decision-making such as V , $\partial V/\partial t$ and capturing events. The major finding in this study is that a neural system consisting of ACC, PCC, SPL, and VP is related with $\partial V/\partial t$, the basis for dynamic decision-making. In other words, those regions are correlated with the temporal derivative rather than the value of expected utility. Such a distinction cannot be made with a conventional experimental paradigm because, for a trial lasting a few seconds, the neural activity related with both V and $\partial V/\partial t$ occurs at essentially the same time for fMRI. On the other hand, single

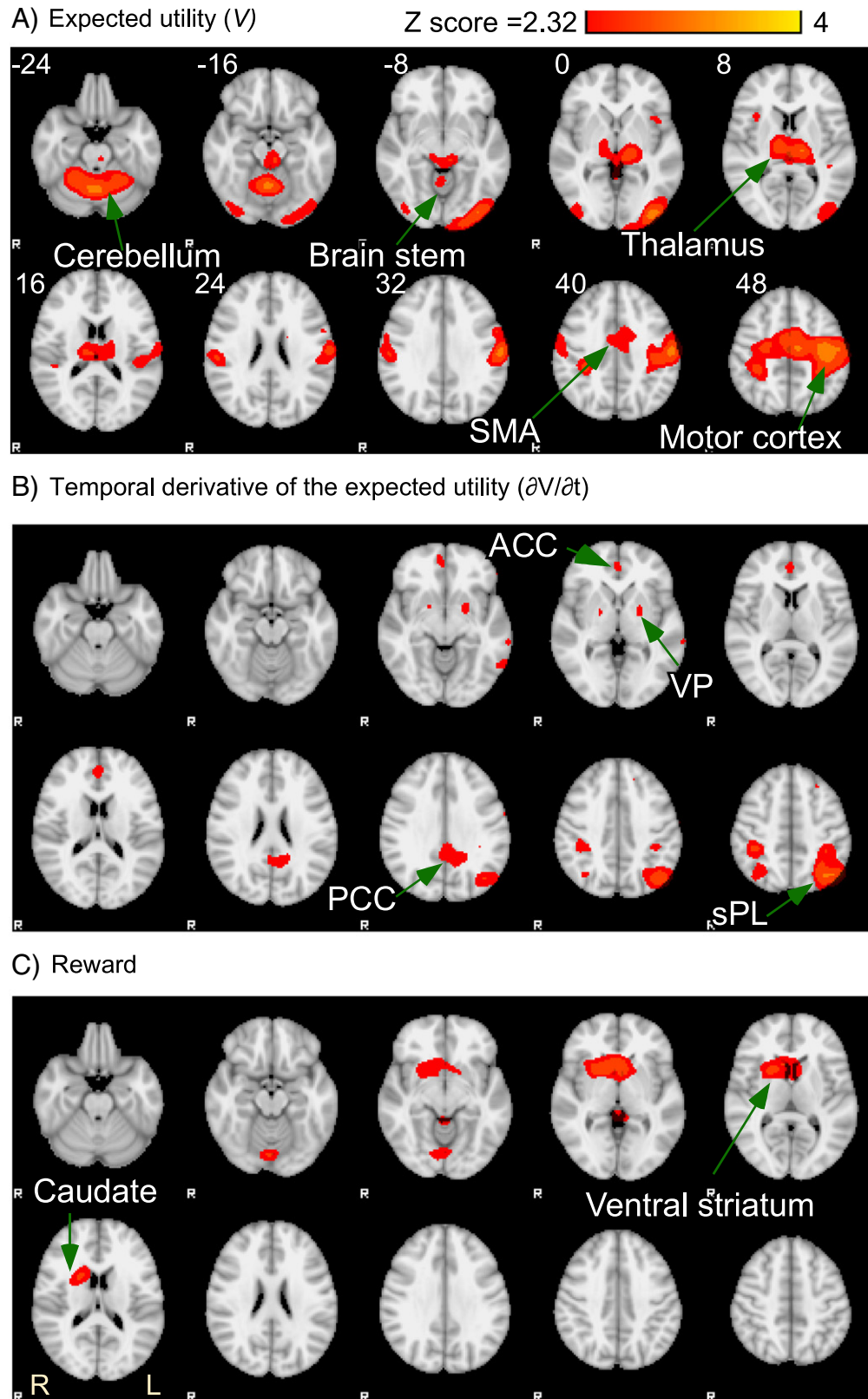


Fig. 5. Areas for which brain activity is correlated with regressors V (A), $\partial V/\partial t$ (B) and Capture event (C). Red/yellow indicates that the BOLD signal is positively correlated with the regressors and the Z scores range from 2.33 to 4. The images are presented following radiological convention, where the right brain is presented as image left. The abbreviations in the figure include anterior cingulate gyrus (ACC), posterior cingulate gyrus (PCC), ventral pallidum (VP), supplementary motor (SMA) and superior parietal lobule (SPL).

Table 1

Z-scores, MNI coordinates and brain areas of the local maxima. Brain areas are identified with the Harvard–Oxford cortical structure atlas. The significance of the activity clusters were corrected for multiple-comparisons at p value < 0.05.

Brain area for the local maxima	Hemi	X	Y	Z	Z score
A. V					
Precentral gyrus (PRECG)	L	−36	−14	54	4.39
Precentral gyrus (PRECG)	R	24	−12	60	4.3
Thalamus (THAL)	L	−14	−20	6	3.9
Thalamus (THAL)	R	12	−16	10	3.45
Inferior lateral occipital cortex (ILOC)	L	−38	−88	0	4.09
Inferior lateral occipital cortex (ILOC)	R	42	−86	0	2.89
Postcentral gyrus (POSTCG)	L	−40	−24	60	4.32
Supplementary motor area (SMA)	L	−4	−2	56	3.98
Brain-stem	LR	−6	−30	−14	3.18
B. $\partial V/\partial t$					
Superior lateral occipital cortex (SLOC)	L	−42	−68	46	3.84
Superior lateral occipital cortex (SLOC)	R	30	−56	56	3.25
Posterior cingulate gyrus (PCC)	L	−14	−48	28	2.79
Posterior cingulate gyrus (PCC)	R	0	−44	30	2.74
Ventral pallidum (VP)	L	−22	2	−6	2.69
Posterior middle temporal gyrus (PMTG)	R	38	−38	48	3.3
Superior parietal lobule (sPL)	L	−36	−44	52	3.21
Anterior cingulate gyrus (ACC)	R	2	44	14	2.84
Medial frontal cortex (MFC)	R	6	52	−10	2.62
Frontal inferior orbital gyrus (IFG)	L	−54	38	−4	2.46
Precentral gyrus (PRECG)	L	−62	−2	34	2.4
Postcentral gyrus (POSTCG)	L	−64	−10	40	2.34
Superior frontal gyrus (SFG)	L	−26	28	46	2.38
Frontal superior medial gyrus	R	8	70	12	2.37
C. Capture event					
Subcallosal cortex (SCC)	R	2	20	0	3.56
Putamen (CPU)	R	22	12	−2	3.46
Caudate nucleus (CN)	R	20	12	12	3.39
Lingual gyrus (LING)	R	4	−78	−14	3.23
Brain-stem	LR	2	−40	−4	2.89

cell recordings, with high temporal resolution, have shown that decision-making related neurons have a phasic response pattern (Moncayo et al., 2000; Opris et al., 2009; Schultz et al., 1997), consistent with a derivative operation. In summary, our results provide additional evidence in support of a neural mechanism related to the derivative of expected utility using fMRI data and human subjects.

Although it is frequently reported that ventral striatum, ACC, PCC and other frontal areas are correlated with V in studies using static decision-making tasks (Kable and Glimcher, 2007; McClure et al., 2004; O'Doherty et al., 2007; Pine et al., 2009), our observation, based on the dissociation of V and $\partial V/\partial t$, adds granularity to those findings. The pattern of the activity related to the regressor V was mainly associated with motor related regions. A similar pattern has also been reported with monkey single unit recordings, where the activity of SMA and motor cortex was enhanced when the primate either anticipated a short delay or experienced a long elapsed delay (Roesch and Olson, 2005; Watanabe, 1986) and similar findings were also reported in a human fMRI study (Wunderlich et al., 2009).

The pursuit–evasion video game employed in this study is intended to elicit complex, real-world dynamic decision-making. This departure from the conventional approach to study decision-making as a single input–output event enables the separation of components of a decision-making process. Behavioral and fMRI results suggest that dynamic decision-making is based both upon the exponential discounted utility (V) and its temporal derivative ($\partial V/\partial t$), consistent with dynamic game theory as described in mathematics. The separation between factors of decision-making advances our understanding of dynamic decision-making and understanding the neural activity during complex and dynamic real world situations.

Acknowledgments

This study is partially supported by:
NIAAA-09-07 (NIH) HHSN275200900019C Mechanisms of Behavior Change (subaward JH, PI Jon Morgenstern) and NIH R01 HD051912-01A2 Mechanisms of recovery following severe brain injury (subaward JH, PI Nicholas Schiff).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2012.08.063>.

References

- Basar, T., Olsder, G., 1999. Dynamic Noncooperative Game Theory, 2nd ed. Society for Industrial and Applied Math.
- Bernoulli, D., 1738. Exposition of a new theory on the measurement of risk. *Econometrica* 22, 22–36.
- Glover, G.H., Thomason, M.E., 2004. Improved combination of spiral-in/out images for BOLD fMRI. *Magn. Reson. Med.* 51, 863–868.
- Hare, T.A., O'Doherty, J., Camerer, C.F., Schultz, W., Rangel, A., 2008. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.* 28, 5623–5630.
- Isaacs, R., 1965. The Theory of Differential Games. Wiley, New York, NY.
- Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.
- Loewenstein, G., Prelec, D., 1992. Anomalies in intertemporal choice — evidence and an interpretation. *Q. J. Econ.* 107, 573–597.
- Mathiak, K.A., Klaser, M., Weber, R., Ackermann, H., Shergill, S.S., Mathiak, K., 2011. Reward system and temporal pole contributions to affective evaluation during a first person shooter video game. *BMC Neurosci.* 12, 66.
- McClure, S.M., Laibson, D.I., Loewenstein, G., Cohen, J.D., 2004. Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507.
- McNamara, F.N., Clifford, J.J., Tighe, O., Kinsella, A., Drago, J., Fuchs, S., Croke, D.T., Waddington, J.L., 2002. Phenotypic, ethologically based resolution of spontaneous and D(2)-like vs D(1)-like agonist-induced behavioural topography in mice with congenic D(3) dopamine receptor 'knockout'. *Synapse* 46, 19–31.
- Moncayo, J., de Freitas, G.R., Bogousslavsky, J., Altieri, M., van Melle, G., 2000. Do transient ischemic attacks have a neuroprotective effect? *Neurology* 54, 2089–2094.
- Morgenstern, O., Von Neumann, J., 1944. Theory of Games and Economic Behavior. Princeton University Press.
- O'Doherty, J.P., Hampton, A., Kim, H., 2007. Model-based fMRI and its application to reward learning and decision making. *Ann. N. Y. Acad. Sci.* 1104, 35–53.
- Opris, I., Hampson, R.E., Deadwyler, S.A., 2009. The encoding of cocaine vs. natural rewards in the striatum of nonhuman primates: categories with different activations. *Neuroscience* 163, 40–54.
- Pine, A., Seymour, B., Roiser, J.P., Bossaerts, P., Friston, K.J., Curran, H.V., Dolan, R.J., 2009. Encoding of marginal utility across time in the human brain. *J. Neurosci.* 29, 9575–9581.
- Richards, J.B., Mitchell, S.H., DeWit, H., Seiden, L.S., 1997. Determination of discount functions in rats with an adjusting-amount procedure. *J. Exp. Anal. Behav.* 67, 353–366.
- Roesch, M.R., Olson, C.R., 2005. Neuronal activity dependent on anticipated and elapsed delay in macaque prefrontal cortex, frontal and supplementary eye fields, and premotor cortex. *J. Neurophysiol.* 94, 1469–1497.
- Samuelson, P., 1937. A note on measurement of utility. *Rev. Econ. Stud.* 4, 155–161.
- Sarinopoulos, I., Grupe, D.W., Mackiewicz, K.L., Herrington, J.D., Lor, M., Steege, E.E., Nitschke, J.B., 2010. Uncertainty during anticipation modulates neural responses to aversion in human insula and amygdala. *Cereb. Cortex* 20, 929–940.
- Scholten, M., Read, D., 2010. The psychology of intertemporal tradeoffs. *Psychol. Rev.* 117, 925–944.
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science* 275, 1593–1599.
- Schweighofer, N., Shishida, K., Han, C.E., Okamoto, Y., Tanaka, S.C., Yamawaki, S., Doya, K., 2006. Humans can adopt optimal discounting strategy under real-time constraints. *PLoS Comput. Biol.* 2, 1349–1356.
- Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niaz, R.K., Saunders, J., Vickers, J., Zhang, Y.Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 23, S208–S219.
- Strotz, R.H., 1956. Myopia and inconsistency in dynamic utility maximization. *Rev. Econ. Stud.* 23, 165–180.
- Sutton, R.S., Barto, A.G., 1987. A temporal-difference model of classical conditioning. Proceedings of the Ninth Annual Conference of the Cognitive Science Society, pp. 355–378.
- Watanabe, M., 1986. Prefrontal unit-activity during delayed conditional go/no-go discrimination in the monkey. 2. Relation to go and no-go responses. *Brain Res.* 382, 15–27.
- Wunderlich, K., Rangel, A., O'Doherty, J.P., 2009. Neural computations underlying action-based decision making in the human brain. *Proc. Natl. Acad. Sci. U. S. A.* 106, 17199–17204.