NeuroImage 65 (2013) 223–230

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

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.

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).

1053-8119/$ – see front matter
http://dx.doi.org/10.1016/j.neuroimage.2012.08.063

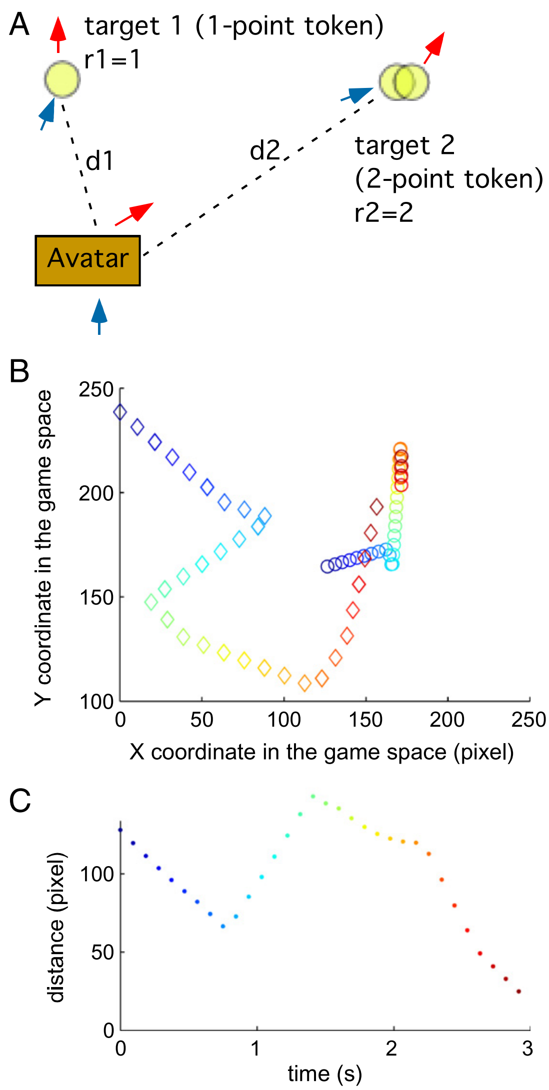

Fig. 1. A. An example of the target-choice task. Based on the values, r1 and r2, and the distances, d1 and d2, the player makes a choice, which, in this case, is to pursue target 2. The blue arrows indicate the directions of movement prior to the event. The red arrows indicate the new directions of the 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