Neuron
Article

Neuronal Origins of Choice Variability in Economic Decisions

Camillo Padoa-Schioppa[1,2,3,*]
[1]Department of Anatomy and Neurobiology
[2]Department of Economics
[3]Department of Biomedical Engineering
Washington University in St. Louis, St. Louis, MO 63110, USA
*Correspondence: camillo@wustl.edu
http://dx.doi.org/10.1016/j.neuron.2013.09.013

SUMMARY

To investigate the mechanisms through which economic decisions are formed, I examined the activity of neurons in the orbitofrontal cortex while monkeys chose between different juice types. Different classes of cells encoded the value of individual offers (*offer value*), the value of the chosen option (*chosen value*), or the identity of the chosen juice (*chosen juice*). Choice variability was partly explained by the tendency to repeat choices (choice hysteresis). Surprisingly, near-indifference decisions did not reflect fluctuations in the activity of *offer value* cells. In contrast, near-indifference decisions correlated with fluctuations in the preoffer activity of *chosen juice* cells. After the offer, the activity of *chosen juice* cells reflected the decision difficulty but did not resemble a race-to-threshold. Finally, *chosen value* cells presented an "activity overshooting" closely related to the decision difficulty and possibly due to fluctuations in the relative value of the juices. This overshooting was independent of choice hysteresis.

INTRODUCTION

In recent years, significant progress has been made in understanding the neural underpinnings of economic choices. In particular, much work has focused on the computation and representation of subjective values. Lesion studies have shown that value-based decisions are selectively disrupted after lesions to the orbitofrontal cortex (OFC) and/or the amygdala, but effectively spared after lesions to other brain regions (Buckley et al., 2009; Camille et al., 2011; Gallagher et al., 1999; Rudebeck and Murray, 2011; West et al., 2011). Neurophysiology experiments have found that neurons in the primate OFC encode the subjective value of different goods during economic decisions and integrate multiple dimensions on which goods can vary (Kennerley et al., 2009; Padoa-Schioppa and Assad, 2006; Roesch and Olson, 2005). Functional imaging in humans has consistently confirmed and extended these results (Kable and Glimcher, 2007; Levy et al., 2010; Peters and Büchel, 2009; Plassmann et al., 2007). But in spite of these advances, fundamental questions remain open. Perhaps most pressingly, the precise mechanisms through which values are compared remain unclear. In this respect, OFC appears particularly noteworthy. In a computational sense, an economic decision is a process through which the values of different goods are compared and one good is eventually chosen. Studies in which monkeys chose between different juice types have shown that neurons in the OFC encode three variables: *offer value* (the value of individual goods, independent of the eventual choice), *chosen value* (the value of the chosen good, independent of its identity), and *chosen juice* (the identity of the chosen good, independent of its value) (Padoa-Schioppa and Assad, 2006, 2008). OFC thus appears to represent all the components of the decision process, suggesting that closer examination of activity in this area might shed light on key aspects of economic choice.

In the effort to unravel the neuronal mechanisms of economic decisions, it could be fruitful to establish an analogy between economic decisions and other behaviors frequently examined in neurophysiology (Sugrue et al., 2005). In particular, extensive research has focused on the decision process underlying the visual perception of motion (henceforth "perceptual decisions"). In a somewhat simplified account, two brain areas play a critical role. Neurons in the middle temporal (MT) area encode the direction of motion for the stimuli present in the visual scene at any given time. In contrast, neurons in the lateral intraparietal (LIP) area encode the binary result of the decision process. When stimuli are degraded such that the decision process stretches over longer periods of time, neurons in MT encode the instantaneous evidence from the visual stimuli, with no memory. In contrast, neurons in LIP encode the accumulated evidence in favor of one particular decision (Newsome, 1997; Shadlen et al., 1996). Tracing the analogy between economic and perceptual decisions, *offer value* cells in OFC may correspond to neurons in MT, whereas *chosen juice* may correspond to neurons in LIP. Indeed, the former seem to represent the main input to the decision process, whereas the latter seem to represent the binary outcome of the decision. In contrast, *chosen value* cells in OFC do not appear to have a clear counterpart in perceptual decisions.

The analogy with perceptual decisions highlights two fundamental and open issues in economic decision-making. First, extensive work on perceptual decisions has been devoted to understanding how fluctuations in the activity of different neuronal populations contribute to decisions near the indifference point (threshold). In particular, the observation that near-indifference decisions are mildly, but significantly, correlated with activity fluctuations in area MT (Britten et al., 1996; Cohen and Newsome, 2009) has provided a critical link between this area and the perception of motion. In contrast, the neuronal origins of variability in economic choices have not yet been examined, and we do not yet understand what drives decisions near the indifference point. Second, the time necessary to reach either a perceptual or an economic decision depends on the decision difficulty (Padoa-Schioppa et al., 2006; Roitman and Shadlen, 2002; Soltani et al., 2012). Building on this notion, much research has focused on neuronal activity reflecting the formation of a perceptual decision over time. In particular, the activity of neurons in LIP was found to increase gradually during perceptual decisions, suggesting that these cells encode the evolving decision state of the animal (Roitman and Shadlen, 2002; Shadlen and Newsome, 2001). By comparison, less is known about how economic decisions form over time, or about how economic decisions depend on the decision difficulty. In addition to these empirical questions, considerable work on perceptual decisions has been devoted to mathematical conceptualization. Specifically, activity profiles in area LIP have been described with a variety of models, including race-to-threshold processes and dynamical systems (Bogacz et al., 2006; Gold and Shadlen, 2007; Wang, 2002). In contrast, although several proposals were recently put forth (Hunt et al., 2012; Krajbich et al., 2010; Soltani et al., 2012; Solway and Botvinick, 2012), a comprehensive model for the neuronal mechanisms of economic decisions remains elusive (see Discussion). To address these issues and gather elements that would inform future models, I examined data recorded in the OFC of monkeys engaged in economic choices.

RESULTS

Neuronal activity in OFC was recorded in two experiments during which monkeys chose between different juice types (see Experimental Procedures). In experiment 1, animals chose between two juices labeled A and B, with A preferred (Padoa-Schioppa and Assad, 2006). Offers were represented by sets of colored squares on a computer monitor and the animals indicated their choices with an eye movement. Juice quantities varied from trial to trial and behavioral choice patterns typically presented a quality-quantity trade-off (Figures 1A and 1B). In experiment 2, the procedures were very similar except that three juices were used in each session (Padoa-Schioppa and Assad, 2008). Two of the three juices were offered in each trial, with the three juice pairs randomly interleaved.

From Neuronal Responses to Cell Classes

Previous analyses were based on neuronal responses, defined as the activity of one cell in one time window (see Experimental Procedures). It was shown that the vast majority of neuronal responses encoded one of three variables: *offer value* (Figure 1C), *chosen value* (Figure 1D), and *chosen juice* (Figure 1E). Notably, the firing rate could increase or decrease as a function of the encoded variable (positive or negative encoding). However, two important questions were not previously addressed. First, because variables *offer value*, *chosen value* and *chosen juice* were intrinsically correlated, individual responses were often explained by more than one variable. For example, one response could have a nonzero slope when regressed onto either *offer value* or *chosen value*. In such case, the response was assigned to the variable with the highest R^2. However, this criterion did not assess whether *offer value* and *chosen value* were distinct classes of responses or, alternatively, whether the two variables represented “poles” of a continuum. Second, previous studies did not test whether *offer value*, *chosen value*, and *chosen juice* corresponded to separate groups of cells. In principle, any given neuron could encode different variables at different times. Alternatively, each cell could consistently encode a single variable. I addressed these issues as follows.

To assess whether *offer value* and *chosen value* are distinct classes of responses, I computed for each response the linear regression onto variables *offer value* and *chosen value*, from which I obtained the two R^2. I then defined $\Delta R^2 = R^2_{offer\ value} - R^2_{chosen\ value}$, which ranged from -1 to $+1$. For a response perfectly explained by *offer value* (*chosen value*) and poorly explained by *chosen value* (*offer value*), ΔR^2 is close to $+1$ (-1). If *offer value* and *chosen value* are two poles of a continuum, the distribution of ΔR^2 should be unimodal with a peak close to zero. Conversely, if *offer value* and *chosen value* are distinct classes of responses, the distribution of ΔR^2 should be bimodal with a dip close to zero. As illustrated in Figures 1F and 1G, the distribution obtained for ΔR^2 was indeed bimodal ($p < 0.02$, Hartigan’s dip test). Thus, *offer value* and *chosen value* appeared to be distinct classes of responses. I repeated this analysis for the two other pairs of variables (*offer value* versus *chosen juice* and *chosen value* versus *chosen juice*). In both cases, the distribution for ΔR^2 was clearly bimodal (both $p < 10^{-10}$, Hartigan’s dip test; Figures 1H–1K). In conclusion, *offer value*, *chosen value*, and *chosen juice* responses are best thought of as different classes of responses, not as poles of a continuum.

To assess whether different neurons encoded different variables, I first examined data from experiment 1. Neuronal responses were classified as encoding one of four variables: *offer value A*, *offer value B*, *chosen value*, or *chosen juice*. Responses that were not task-related or that were not explained by any variable were unclassified. Given a neuron and two time windows, I defined a “classification conflict” if the neuron was classified in both time windows but it encoded different variables. A conflict was also detected if a cell encoded the same variable but with different sign. I thus sought to establish whether the incidence of classification conflicts in the population was greater, comparable, or lower than expected by chance. Chance level was estimated with a bootstrap technique (see Supplemental Experimental Procedures available online). This analysis showed that the number of classification conflicts present in the data was significantly lower than expected by chance ($p < 10^{-10}$, t test; Figure S1A). Conversely, for each pair of time windows, cells with consistent classification were significantly more frequent than expected by chance (Figure S1B; all $p < 10^{-10}$, t test). Data from experiment 2 provided very similar results both for the analysis of classification conflicts ($p < 10^{-10}$, t test) and for that of classification consistency (all $p < 10^{-10}$, t test). In other words, OFC neurons typically encoded the same variable across time windows.