MEASURING BELIEFS AND REWARDS: A NEUROECONOMIC APPROACH*

ANDREW CAPLIN
MARK DEAN
PAUL W. GLIMCHER
ROBB B. RUTLEDGE

The neurotransmitter dopamine is central to the emerging discipline of neuroeconomics; it is hypothesized to encode the difference between expected and realized rewards and thereby to mediate belief formation and choice. We develop the first formal tests of this theory of dopaminergic function, based on a recent axiomatization by Caplin and Dean (*Quarterly Journal of Economics*, 123 (2008), 663–702). These tests are satisfied by neural activity in the nucleus accumbens, an area rich in dopamine receptors. We find evidence for separate positive and negative reward prediction error signals, suggesting that behavioral asymmetries in responses to losses and gains may parallel asymmetries in nucleus accumbens activity.

I. INTRODUCTION

The neurotransmitter dopamine is central to the emerging discipline of neuroeconomics. Pioneering work by Wolfram Schultz, P. Read Montague, Peter Dayan, and their colleagues[1] suggests that dopamine not only participates in the encoding of information on crucial economic variables such as preferences and beliefs, but also plays a key role in choice and learning. The "dopaminergic reward prediction error" (DRPE) hypothesis states that instantaneous dopamine levels in the brain encode the difference between how rewarding an event is expected to be, and how rewarding it turns out to be. Largely based on this hypothesis, research informed by an understanding of the dopamine system has already had an impact on the social sciences.[2]

Reasons for economists to be interested in observing reward prediction errors are manifold.[3] Beliefs play a central role in theories of decision-making and learning, yet they are hard to

*We thank Larry Katz and four anonymous referees for their suggestions and open-mindedness. We thank Alberto Bisin, Peter Bossaerts, Mauricio Delgado, Laura deSouza, Eric DeWitt, Ernst Fehr, Souheil Inati, Joe Kable, Ifat Levy, Kenway Louie, P. Read Montague, Yael Niv, and Antonio Rangel for valuable guidance.

1. See Schultz, Apicella, and Ljungberg (1993), Mirenowicz and Schultz (1994), Montague, Dayan, and Sejnowski (1996), Schultz, Dayan, and Montague (1997), and Hollerman and Schultz (1998).

2. For example, Bernheim and Rangel (2004), McClure et al. (2004), and Bossaerts, Preuchoff, and Hsu (2008).

3. For more details, see Caplin and Dean (2008a, 2008b, 2008c).

The Quarterly Journal of Economics, August 2010

observe. Adding dopaminergic measurements to the arsenal of belief measurement techniques can bolster current methods based on structural models of the updating processes (e.g., Cheung and Friedman [1997]) or incentive-compatible elicitation methods (Nyarko and Schotter 2002). Similarly, theories of reference-dependent choice, such as loss aversion, give a central role to the decision-maker's reference point, yet little is known about how this is determined.[4] Dopamine provides a promising line of attack for those seeking to understand reference points and reward expectations, at least in tightly specified "neuroeconomic" experiments.

Despite its promise, Caplin and Dean (2008a) detail evidentiary gaps in the existing literature on the DRPE hypothesis. Critically, most tests of the DRPE hypothesis assume that expectations are derived through reinforcement learning and then test a highly parameterized version of the DRPE hypothesis.[5] Yet reinforcement learning often contradicts Bayesian learning, and behavioral experiments suggest that it serves as a good model of learning only in restrictive circumstances (Charness and Levin 2005). Moreover, the predictions of a given DRPE model based on reinforcement learning are often violated in the data.[6]

To address these issues, Caplin and Dean (2008c) (henceforth CDc) propose an axiomatically based testing protocol that disconnects the DRPE hypothesis entirely from learning behavior. CDc consider a simple data set in which dopamine activity is observed when prizes are obtained from different lotteries. In such an environment, the DRPE hypothesis can be characterized by three simple and easily testable axioms. These axioms target the DRPE hypothesis without any of the auxiliary hypotheses that have characterized previous tests.

The current paper thus contains the first tests of the axiomatic version of the DRPE hypothesis. We pick a simple setting in which the DRPE hypothesis can be completely characterized by the three intuitive axioms of CDc. In the experiment that we use to test these axioms, human subjects are endowed with lotteries from which a prize is drawn. We use functional magnetic resonance imaging (fMRI) to measure brain activity as the prize is revealed

4. Kahneman and Tversky (1979), Samuelson and Zeckhauser (1988), and Köszegi and Rabin (2006).

5. For example, O'Doherty et al. (2004) and Bayer and Glimcher (2005).

6. For example, Berridge and Robinson (1998), Zink et al. (2003), Delgado et al. (2005), Knutson and Peterson (2005), and Redgrave and Gurney (2006).

to the subject. By comparing fMRI measures of activity as different prizes are received from different lotteries, we test whether activity in a brain region known as the nucleus accumbens satisfies the axioms. This brain region is a principal anatomical target of the dopamine neurons hypothesized to encode the DRPE signal.

In broad terms, the results of our experimental tests support the basic DRPE model. Of the three axioms that we test, two are strongly supported, the third weakly. To a first approximation, measured activity in the nucleus accumbens does indeed satisfy the DRPE axioms. Our experiment also throws up one intriguing and unexpected finding. Our evidence suggests that overall dopaminergic activity may be an amalgamation of two different processes operating with different temporal dynamics: the signal recording "positive" prediction error acts at a shorter time lag, and with less intensity, than that recording negative prediction error. This suggests that further study of the dopamine system may be particularly valuable for those interested in understanding asymmetric responses to gains and losses of the form described in prospect theory (Kahneman and Tversky 1979).

II. Dopamine and the Reward Prediction Error Hypothesis

II.A. What Is Dopamine?

The brain is composed of tens of billions of neurons, tiny self-sustaining units about a thousandth of an inch in diameter. A connection between neurons across which communication can take place is called a synapse. Such connections allow (in general) one-way communication, with a presynaptic neuron communicating information to one, or possibly many, postsynaptic cells. A neurotransmitter is a chemical used in this process of communication. When a presynaptic neuron releases a neurotransmitter, it travels across the synaptic cleft, the physical gap across which the synaptic connection is made, and attaches itself to receptors in the postsynaptic cell. Thus, the state of the postsynaptic neuron comes to reflect the fact that the presynaptic neuron has released a neurotransmitter, a form of information transfer.

Dopamine is one such neurotransmitter, and the term *dopamine* (or *dopaminergic*) *neuron* refers to any neuron that uses dopamine as a neurotransmitter to communicate with its postsynaptic (downstream) partners. Although dopamine neurons exist in several different parts of the brain, this paper focuses on the midbrain dopamine neurons, a particular class of these

neurons located at the base of the brain. Interestingly, although the dendrites of these midbrain dopamine neurons (the structures through which these cells receive inputs from upstream neurons) are located in a relatively small region of the brain, the axons of these neurons distribute dopaminergic synapses throughout almost half of the human brain. This suggests that the information that they transmit might well be of importance to neurons in many different functional divisions of the nervous system.

II.B. The DRPE Hypothesis

It was observed early on that many addictive drugs mimic the effects of dopamine at the synapse, and that humans appear to place a high positive value (as measured by both self-report and choice) on processes that activate or mimic the activity of midbrain dopamine neurons (see Wise [2004] for a review). As a result of these early observations, midbrain dopamine neurons were presumed to carry some kind of hedonic pleasure signal.

This simple "dopamine-as-pleasure" theory was called into question by studies that showed that dopamine signals were stronger when the same reward was delivered unexpectedly than when it was expected. For example, Mirenowicz and Schultz (1994) measured the activity of dopaminergic neurons in a thirsty monkey as it learned to associate a tone with the receipt of fruit juice. Dopamine neurons were initially active in response to the juice but not the tone. However, after many repetitions (presumably once the monkey had learned that the tone predicted the arrival of juice), dopamine neurons responded to the tone rather than to the juice. Moreover, once learning had taken place, if the tone was played but the monkey did not receive the juice, then there was a pause or decrease in the background level of dopamine activity at the time that the juice was expected.

These findings led to the hypothesis that dopamine was encoding the difference between "experienced" and "predicted" reward, or a "reward prediction error" (Montague, Dayan, and Sejnowski 1996; Schultz, Dayan, and Montague 1997). In the above example, before learning had taken place, the receipt of the fruit juice was a positive surprise (in the sense of a positive utility shock) to the monkey, so dopamine responded in a positive way. However, after learning had taken place, although still rewarding, the fruit juice was no longer surprising, so dopamine did not respond to its arrival. However, the tone was now both

surprising and rewarding, as it was unexpected and predicted the imminent arrival of juice.

If correct, the DRPE hypothesis makes the observation of dopamine of great potential interest to economists. Not only does dopamine carry information on beliefs and rewards (or preferences), but subsequent studies have shown it to play an important role in choice and learning. We will return to this point in Section VI.

II.C. Testing the DRPE Hypothesis

The neuroscientific literature contains a number of tests of the DRPE hypothesis on both monkeys and humans (Schultz, Dayan, and Montague 1997; McClure, Berns, and Montague 2003; O'Doherty et al. 2003, 2004; Bayer and Glimcher 2005; Abler et al. 2006; Li et al. 2006; Pessiglione et al. 2006; Bayer, Lau, and Glimcher 2007; D'Ardenne et al. 2008). Although generally supportive, these tests have all failed to be taken as conclusive proof of the DRPE hypothesis. These tests typically operationalize the DRPE hypothesis by assuming fixed values for the "experienced reward" of different events and using a reinforcement learning model to construct a time path for "predicted reward." This allows the authors to construct a "reward prediction error" for a sequence of rewards and cues, which can then be compared to observed dopamine activity. Typically, these studies do find that dopamine activity, or neural activity in areas rich in dopamine receptors, is correlated with the reward prediction error signal. Although these restrictive tests have provided generally intriguing results, it is unsurprising to learn that they have not conclusively demonstrated that the DRPE theory is both necessary and sufficient to explain the role of dopamine in behavior.

Perhaps the main reason that alternative theories of dopamine remain plausible is that existing tests of the DRPE hypothesis have relied on auxiliary assumptions (arbitrary parameterizations lying outside the theory) and on very weak tests. It is easy to understand the attraction of such tests because they provide insight not only into the basic question of whether or not the DRPE hypothesis is correct, but also into the actual learning algorithm it may encode. Unfortunately, this makes it hard to separate out precisely how strong is the support for the broad hypothesis as opposed to the learning algorithm. In O'Doherty et al. (2003), for example, the authors use a model of reinforcement learning to fit neural responses. In support of the basic

DRPE hypothesis, the evidence did suggest that once a human had been repeatedly exposed to a tone that predicted a reward, dopamine neurons became active in response to the tone itself. Although it is clear that many parameterized versions of the DRPE hypothesis *do* make such a prediction, (i) many other theories of dopamine also make this prediction and (ii) many parameterizations of the DRPE theory do *not* make this prediction.

Tests such as this one are therefore joint tests of the underlying DRPE model, the reinforcement learning model of belief formation, and a set of arbitrary parameterizations. For this reason, a set of alternative theories of dopamine function (and hence alternative theories of the biological basis of belief formation) persist. The qualitative fMRI studies of Zink et al. (2003), Delgado et al. (2005), and Knutson and Peterson (2005) have, for example, suggested that dopamine responses may be modulated by "salience," or how surprising an event is. Redgrave and Gurney (2006) suggest that dopamine plays a role in switching attention between different activities. The incentive salience hypothesis of Berridge and Robinson (1998) holds that dopamine influences the subject's assessment of a reward's salience, but in a way that is not causally related to belief formation. By stripping away the need for these additional assumptions, and by anchoring experimental data to conditions of necessity and sufficiency, the axiomatic approach provides for tests of the underlying DRPE hypothesis without relying on a particular model of belief formation or on arbitrary parameterizations.

III. The Axiomatic Model

In this paper, we use an axiomatic representation based on the work of CDc to design and implement a test of the DRPE hypothesis. The axiomatic representation provides a set of necessary and sufficient conditions for the entire class of DRPE models. Moreover, these tests do not require *ad hoc* auxiliary assumptions on the nature of belief formation or "reward." Thus the axioms provide a simple and parsimonious way of testing the concepts that lie at the heart of the DRPE hypothesis.

III.A. Definitions

The environment in which we formalize and test the DRPE hypothesis is one in which an experimental subject is endowed

with a lottery (or probability distribution over prizes) from which a specific prize is then realized.[7] The key observable is the firing rate of dopamine neurons, $\delta(z, p)$, when the prize z is obtained from the lottery p. The characterizations in CDc are based on an idealized data set in which the dopaminergic firing rate is observed for any such conceivable combination of prizes and lotteries. For experimental purposes, it is important to deal with cases in which we observe δ only on some finite subset A of all possible lottery–prize pairs, as these are the data that will be generated by any real-world experiment. We therefore define a finite version of the data set described in CDc.[8]

DEFINITION 1. Let Z be a set of prizes with generic element $z \in Z$. The set of all simple lotteries over Z is denoted Λ, with generic element $p \in \Lambda$. We define the set $\Lambda(z)$ as all lotteries with z in their support, and denote as z the degenerate lottery that assigns probability 1 to prize $z \in Z$,

$$z \in \Lambda(z) \equiv \{p \in \Lambda | p_z > 0\}.$$

A dopaminergic data set comprises a finite set A consisting of pairs (z_n, p_n), with $z_n \in Z$ and $p_n \in \Lambda(z_n)$ for all $1 \leq n \leq N$, and with $\{z, z\} \in A$, $\forall\ z \in Z$, together with a dopaminergic firing rate $\delta : A \to \mathbb{R}$ for each observation $(z_n, p_n) \in A$.

In the two-prize case, a dopaminergic data set can be represented easily in graphical form, as demonstrated in Figure I. The space of lotteries, Λ, can be represented by a single number: the probability of winning prize 1. This forms the x-axis of these figures. We represent the function δ using two lines—the solid line indicates the dopamine firing rate after prize 1 is obtained from each of these lotteries (i.e., $\delta(z_1, p)$), whereas the dashed line represents the dopamine firing rate when prize 2 is obtained from each lottery (i.e., $\delta(z_2, p)$).

The definition of a DRPE representation is as in CDc. Effectively, we say that dopamine has a DRPE representation if we can find an expected reward function for lotteries and

7. We do not allow observation of dopaminergic activity from a prize that is impossible according to the given lottery (i.e., a prize from outside the support of a particular lottery).

8. Given that the DRPE hypothesis has quite specific information on what happens when there is no surprise, we will also insist that all no-surprise outcomes of the form (z, z) are in the domain of observation, although this has no technical impact on the availability of a DRPE representation.

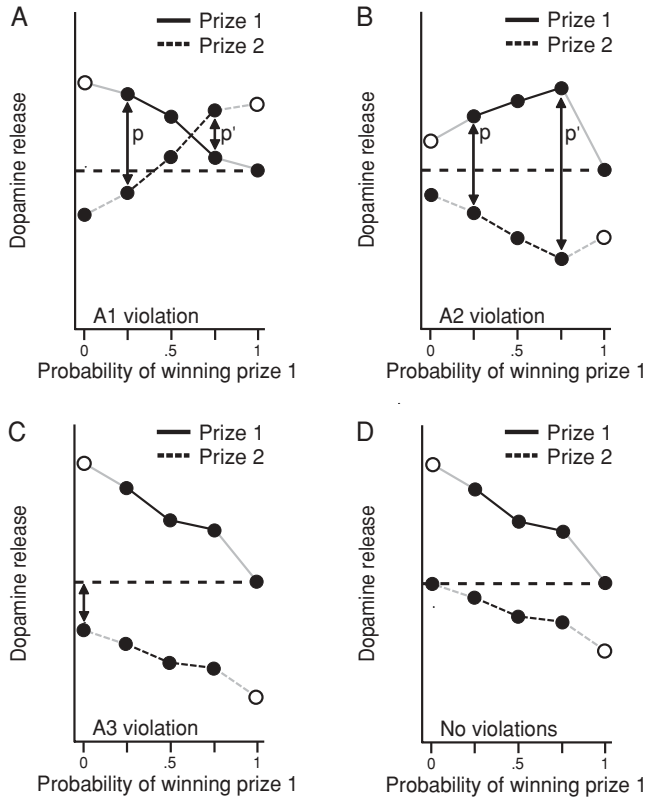

FIGURE I

Graphical Representation of Violations of the Axioms for the Two-Prize Case

Solid points represent example experimental measurements. Open points represent unobservable outcomes. (A) A violation of Axiom 1: Coherent Prize Dominance. When received from lottery p, prize 1 leads to higher dopamine release than does prize 2, indicating that prize 1 has higher experienced reward. This order is reversed when the prizes are realized from lottery p', suggesting that prize 2 has higher experienced reward. Thus a DRPE representation is impossible. (B) A violation of Axiom 2: Coherent Lottery Dominance. More dopamine is released when prize 1 is obtained from lottery p' than from lottery p, suggesting that p has a higher predicted reward than p'. The reverse is true for prize 2, making a DRPE representation impossible. (C) A violation of Axiom 3: No Surprise Equivalence. The dopamine released when prize 1 is obtained from its degenerate lottery is higher than when prize 2 is obtained from its degenerate lottery. (D) No axioms are violated in this graph.

an experienced reward function for prizes such that dopamine activity is decreasing in the former and increasing in the latter. Furthermore, all situations in which experienced reward is equal to actual reward, and thus there is no "reward prediction error,"

must be treated equivalently by dopamine. These properties capture the notion that dopamine encodes the difference between experienced and predicted rewards.

DEFINITION 2. The finite dopaminergic data set (A, δ) admits a dopaminergic reward prediction error (DRPE) representation (r, E) if there exist functions $r : \Lambda \to \mathbb{R}$ and $E : r(Z) \times r(\Lambda) \to \mathbb{R}$ such that $\delta(z, p) = E(r(z), r(p))$, with $E(.,.)$ strictly increasing in its first and strictly decreasing in its second argument, and such that $E(r(z), r(z)) = E(r(z'), r(z'))$ for all $z, z' \in Z$.[9]

III.B. The Three Axioms

CDc introduce three necessary conditions for the existence of a DRPE representation. In the case in which there are three or more prizes, these conditions are necessary but not sufficient for a DRPE representation. Yet in the two-prize case one can prove directly that such equivalence does indeed hold.

The first axiom, "Coherent Prize Dominance," demands that dopamine "rank" prizes consistently, regardless of what lottery these prizes were obtained from. If winning prize 1 produces more dopaminergic activity than winning prize 2 from the same lottery, it must be the case that prize 1 has a higher experienced reward. Thus, it must be the case that, from *any lottery,* there is more dopamine released when prize 1 is obtained than when prize 2 is obtained.

AXIOM 1 (A1: Coherent Prize Dominance). Given $(z, p), (z', p), (z, p'), (z', p') \in A$,

$$\delta(z, p) > \delta(z', p) \Rightarrow \delta(z, p') > \delta(z', p').$$

Figure IA shows a violation of this axiom, which in this graphical space is equivalent to the requirement that the lines $\delta(z_1, p)$ and $\delta(z_2, p)$ cannot cross.

The second axiom, "Coherent Lottery Dominance," demands that the ordering of lotteries by dopamine firing rate be independent of the obtained prize. If a higher dopamine firing rate is observed when prize 1 is obtained from lottery p' than from p, this indicates that p' has a lower *predicted reward* than p. Thus

9. Note that we additionally assume that people make perfect predictions in the case of degenerate lotteries: the predicted reward of the lottery p that gives prize z for sure is equal to the experienced reward of prize z. Thus the experienced reward function can be derived directly from the predicted reward function.

it must also be true that we observe a higher dopamine firing rate when prize 2 is obtained from p' than when it is obtained from p.

AXIOM 2 (A2: Coherent Lottery Dominance). Given $(z, p), (z', p), (z, p'), (z', p') \in A$,

$$\delta(z, p) > \delta(z, p') \Rightarrow \delta(z', p) > \delta(z', p').$$

Graphically, coherent lottery ordering is equivalent to the requirement that the lines $\delta(z_1, p)$ and $\delta(z_2, p)$ be co-monotonic—that they have the same direction of slope between any two points. Figure IB shows a case that contradicts this—higher dopamine activity is observed when prize 1 is obtained from lottery p' than when it is obtained from lottery p, yet the exact opposite is true for prize 2.

"No Surprise Equivalence" deals directly with situations in which a particular prize is expected with certainty. These are situations that dopamine must treat equivalently, regardless of the prize, as there is no reward prediction error.

AXIOM 3 (A3: No Surprise Equivalence). Given $z, z' \in Z$,

$$\delta(z', z') = \delta(z, z).$$

Figure IC shows a violation of this axiom, in which more dopamine is released when prize 1 is obtained from its degenerate lottery than when prize 2 is obtained from its degenerate lottery. No Surprise Equivalence demands that the points $\delta(z_1, 1)$ and $\delta(z_2, 0)$ take the same value.

In the case of two prizes, A1–A3 are necessary and sufficient conditions for dopamine activity to be described by the DRPE model.

THEOREM 1. With two pure prizes, a finite dopaminergic data set admits a DRPE if and only if it satisfies A1–A3.[10]

Thus, in the two-prize case, if A1–A3 hold, we will be able to extract consistent orderings over lotteries and prizes, which we can label "dopaminergic" predicted and experienced reward, respectively. Figure ID illustrates such a case. How these orderings might relate to more traditional notions of reward and prediction is a matter we discuss in the conclusion.

10. A proof of this theorem is available from the authors on request.

III.C. Other Models of Dopamine Function

Although we have not explicitly axiomatized other models of dopamine activity, it is clear how some of the alternative hypotheses, if true, would lead to violations of the three axioms described above. Here we focus on how the "hedonia" and "salience" hypotheses would lead to violations of the representation.

The hedonia hypothesis states that, rather than encoding a reward prediction error, dopamine encodes simply the reward value of events. In other words, there is some reward function r that attaches reward values to different events, and dopamine activity is an increasing function of this reward value. Although a system that encodes hedonia might satisfy A1 and A2, it would violate A3: No Surprise Equivalence. Unless every object in the observation set has the same reward value, different prizes would lead to different dopaminergic responses, even when received from degenerate lotteries. Thus A3 provides a test between the hedonia and DRPE hypotheses.

The salience hypothesis states that dopamine responds to the salience, or surprise associated with a particular event. Although the concept of salience is often not well defined, it does seem that for any sensible definition, a system that encoded salience would violate both A1: Coherent Prize Dominance and A2: Coherent Lottery Dominance. To see this, consider a case with two prizes, x and y, and two lotteries. The first, p, gives prize x with 99% probability and prize y with 1% probability, whereas the second, q, gives prize x with 1% and y with 99%. In this case, the salient event is getting prize y from lottery p or getting prize x from lottery q, as these are the "surprising" events. Thus, a salience encoder would imply that $\delta(y, p) > \delta(x, p)$ but $\delta(x, q) > \delta(y, q)$, violating A1. Similarly, $\delta(y, p) > \delta(y, q)$ but $\delta(x, q) > \delta(x, p)$, violating A2. Thus, A1 and A2 provide a test between salience and the DRPE hypothesis.

IV. The Experiment

We describe now the methodology by which we test the axioms described above, and thus the DRPE hypothesis. In an ideal world, we would make real-time observations directly from dopamine neurons as agents chose among, and received prizes from, various lotteries. Unfortunately, such measurements, although feasible in animals (see, for example, Mirenowicz and Schultz [1994], Phillips et al. [2003], and Bayer and Glimcher [2005]), are infeasible in

humans due to the invasiveness of the procedure. Instead, we measure dopamine activity indirectly using fMRI. This technique, described in more detail below, relies on a difference in the magnetic susceptibility of oxygenated and deoxygenated blood to measure a blood oxygen level–dependent (BOLD) signal, which is in turn related to brain activity. By focusing on an area of the basal ganglia called the nucleus accumbens, which is known to receive substantial inputs from the midbrain dopamine neurons, one can obtain an estimate of dopamine-related activity in real time.[11] Unfortunately, the data produced by this technique are noisy, so we use repeated observations (both within and across subjects) to construct estimates of δ. The assumptions we make in doing so are discussed below.

IV.A. Experimental Design

The experimental paradigm we use is designed to endow subjects with lotteries so that we can observe brain activity when they are informed of what prize they have won from each lottery. On each trial, subjects choose between two lotteries, represented by pie charts, and experience the outcome of their chosen lottery. A fixation cross signals the beginning of a trial. After 12.5 seconds, two lotteries appear on either side of the fixation cross. After 5 seconds, the fixation cross is extinguished and the subject has 1.25 seconds to press a button to indicate which of the lotteries he or she wishes to play. The chosen lottery moves to the center of the display and after a delay period of 7.5 seconds, the outcome of the lottery is determined (by a random number generator) and revealed to the subject for 3.75 seconds. The prize which the subject receives is indicated by a change in the color of that prize's segment of the pie chart.[12] If the subject fails to press a button during the response window, he or she receives the worst prize available from any lottery in the experiment, a loss of \$10. Figure II shows the timeline of a typical trial.

As we describe below, brain activity is measured at the point at which the prize that the subject has won is revealed from the

11. It should be noted that this technique measures overall activity in this brain area, to which dopaminergic action potentials are a major, although not unique, contributor. This imposes on our measurement a limitation shared by all fMRI-based studies of dopaminergic activity. If anything, however, this limitation should bias our empirical results *against* observing the axiomatic behavior we seek.

12. All the colors used in the experiment are approximately isoluminent, reducing brain activity that comes about due solely to visual stimulation induced by the changing display.

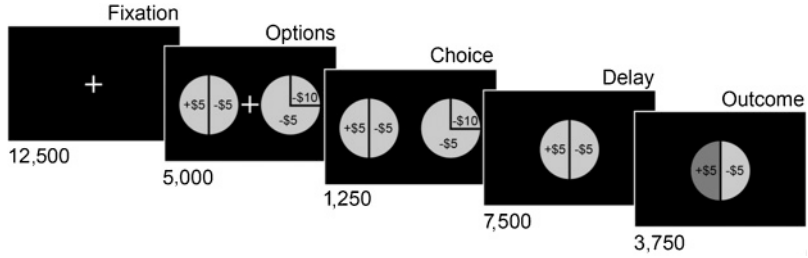

FIGURE II
Experimental Design

Following a fixation period, subjects were presented with two lotteries. When the fixation cross was extinguished, subjects had 1,250 milliseconds to indicate their choices by button presses. Following a delay period, the outcome was revealed by a change in the color of the prize received. Durations of each period in the 30-second trial are given in milliseconds. In this example, the subject chose the lottery on the left and won $5.

lottery he or she owns. It should be noted that, at this stage, this is the only uncertainty that is resolving itself for the subject. Subjects do not learn any more about future lotteries that they will receive, or choices that they will be offered. Thus, we interpret measured neural activity at this stage as the response to the receipt of a particular prize from a particular lottery.

Each subject takes part in two separate scanning sessions consisting of multiple blocks of 16 trials each. Before each session, subjects are given instructions and complete one or more unpaid practice blocks of trials outside the scanner. At the start of each session, subjects are endowed with $100, given to them in cash, with money added to or subtracted from this amount on the basis of the outcome of each trial. How much they have won or lost is reported at the end of each block. The final amount awarded to a subject for a session is the $100 endowment, plus the cumulative outcome (positive or negative) of all lotteries, plus a $35 show-up fee. A full set of instructions is included in Appendix I.

It is worth commenting on some features of the experimental design. First, although we ask subjects to choose between lotteries, we do not make use of the choice data in this paper. The reason we ask for choices is to keep the subject alert and engaged in the experiment. An experimental session lasts for about two hours, and if the subjects are not asked to perform any task during this time they could lose concentration and, in some cases, fall asleep inside the scanner. Second, each trial includes several relatively long delays. The reason for this is that the BOLD signal measured by the scanner is the convolution of the neural activity we wish to

measure with a twenty-second-long "hemodynamic response function," which approximately takes the form of a gamma function. Thus, by spacing out events within a trial, differentiation between activity associated with different events becomes more accurate. Third, we make the somewhat unusual choice to reward subjects based on the outcome of *every* trial, rather than on the basis of some randomly selected subset of trials. The reason for this is also to keep subjects engaged in the experiment. Finally, as subjects can win or lose money on each trial, there is a chance that the subjects will lose all of their \$100 in the course of a scanning session. Although we designed the experiment to minimize the risk of this happening, it is possible. In such an occurrence, the experiment is stopped as soon as the subject's account reaches zero, and the scan terminated by this event is excluded from all further analysis.

Our choice of the lotteries to present to subjects was governed by the need for repeated observations of lottery–prize pairs. As fMRI data have a low signal-to-noise ratio, we need to observe a subject receiving a particular prize from a particular lottery several times to accurately estimate the underlying neural activity. Thus, the set of lottery–prize pairs from which we make observations over a two-hour experiment is relatively small. We restrict ourselves to two prizes (+\$5, −\$5) and five lotteries (probabilities of winning \$5 of 0, .25, .5, .75, and 1), giving eight possible lottery–prize pairs.

In each trial, the subject is offered a choice between one lottery from the above *observation set* and a second lottery from a larger *decoy set*, which included lotteries that have \$0 and −\$10 in their support. To ensure that the lottery from the observation set is chosen in most trials, the decoy lottery has an expected value of between \$1.25 and \$5 less than the observation lottery. In each sixteen-trial scan, assuming the observation lottery is always chosen, the subject receives the degenerate lotteries (those that have a 100% chance of winning a particular prize) twice each and the other lotteries four times each. The ordering of lottery presentation is randomized in each scan.

IV.B. Measuring δ

This experiment provides us with repeated occurrences of a subject receiving a particular prize from a particular lottery. There are four steps to using the experiment to construct measures of δ, and so test our axioms:

1. Use fMRI to obtain data on BOLD activity for all locations within the subject's brain.
2. Define an anatomically restricted region of interest (ROI) within the brain (a subarea very densely populated with dopaminergic synapses), the activity in which we will use as a proxy for dopaminergic activity.
3. Construct a time series of activity in the ROI, and use this time series to construct estimates of δ.
4. Use these estimates of δ to test our axioms.

The following sections describe each of these steps in detail.

(1) From Functional Magnetic Resonance Imaging to Dopamine.[13] The signal measured by an MRI scanner is now very well understood and the mapping of that signal to neural activation is heavily constrained. The scanner works by placing a subject in a strong and highly structured magnetic field and then subjecting him or her to brief radiofrequency pulses of energy. As different chemical substances respond to these pulses as a function of the local magnetic field, this allows the scanner to reveal the chemical structure of tissue at any location inside the brain with tremendous precision.

Relating information about the local chemical structure of the brain to neural activity, however, is significantly more complicated. The local shifts in electrical equilibrium produced by brain activity lie well beyond the resolution of these devices. Instead, the scanners measure brain activity indirectly by observing a small change in the local chemical environment induced by neural activity. When a brain cell becomes active, it consumes energy. This demand for energy leads to an increase in blood flow. The response of the blood flow system to increased demand is now well characterized and approximates a linear process. The vascular system responds to an impulse in demand with a delayed and graded increase in blood flow, with an onset delayed by about two seconds and a peak at a delay of about six seconds, a process known as the *hemodynamic response*. Fortunately for neurobiologists, the molecule hemoglobin, which carries oxygen to the cells, and the density of which is controlled by the hemodynamic response, has a magnetic signature that can be measured by the brain scanner.

13. For technical details of the imaging protocol and initial data analysis, see Appendix II. For more details on magnetic resonance imaging, the reader is referred to Huettel, Song, and McCarthy (2004).

The brain scanner thus allows us to measure the hemodynamic response as a time series at almost any location in the brain. Signal-to-noise considerations, however, limit the precision of this measurement. In practice the scanner yields, with each measurement, the local oxygenation of the blood in little cubes of brain tissue, typically 3 mm on a side, known as *voxels*. The BOLD signal in each voxel is therefore an estimate of the average metabolic demand by all of the neurons within that voxel—on the order of 10,000,000 neurons. By repeating this measurement at intervals of 1–2 seconds, intervals knows as repetition times (TRs), one can construct a time series that reports average metabolic activity in each 3-mm voxel in a human brain. A brain scan typically consists of approximately 150,000 voxels, so this yields approximately 150,000 different time series for each brain scanned.

How can BOLD activity be related to the activity of dopamine neurons? Recall that the MRI scanner averages the activity of the roughly 10,000,000 neurons within each voxel. Unfortunately, the average human brain contains only about 100,000 dopamine neurons, which are distributed spatially over dozens of voxels. The result is that direct measurement of the hemodynamic response induced by the dopamine neurons is at present difficult. However, each dopamine neuron connects to on the order of 10,000 other cells, the locations of which are well known. This means that the activity of on the order of one billion neurons are influenced by dopamine activity, and we know the locations of these neurons. The strategy for measuring dopamine activity in a living human is thus to identify, *ex ante*, the locations in the brain containing high densities of dopaminergic synapses and then to measure the metabolic activity in these regions as a function of behavioral manipulations hypothesized to influence dopaminergic activity.

Studies in animals, where it is feasible to measure both the BOLD signal or dopamine chemically and the activity of nerve cells directly, fortunately provide further constraints on the relationship between dopamine activity and the BOLD signal. A number of studies have now indicated that, at a biological level of analysis, activity in the dopamine neurons and the BOLD signal in our regions of interest are co-monotonic. (For a review of this issue see Knutson and Gibbs [2007].)

(2) Defining Regions of Interest. Scanning subjects using fMRI provides us with an enormous amount of information about BOLD activity; for each of the 150,000 voxels in a scan of a

typical subject's brain it provides a time series of data points for the entire scanning period. The next stage of our analysis is to identify the areas of the brain that we will use to test our theory. As discussed above, several experiments have shown patterns of BOLD activity in the nucleus accumbens and ventral putamen that are strikingly similar to patterns of dopamine activity measured in animals using more direct techniques. Because the nucleus accumbens receives particularly dense projections from a large number of dopamine neurons and can be accurately defined anatomically using data obtained from a brain scanner, we focus on activity in this area as a proxy for dopamine activity. There are two standard ways of identifying regions of interest within fMRI data:

1. Anatomical ROI: Identified as a particular brain structure using an understanding of the physical geography of the brain.
2. Functional ROI: Defined by the way activity in that area is related to a particular stimulus.

In this paper, we focus mainly on an anatomical definition of the nucleus accumbens. For individual subjects, we defined the nucleus accumbens according to the algorithm described in Neto et al. (2008).[14] Figure III shows the ROIs for three of our subjects.

As a robustness check for our results, we also employed a functionally defined ROI, using the assumption that dopaminergic neurons should, as a first approximation, respond positively at the time of prize receipt to the difference between the value of the prize and the expected value of the lottery from which it came. We therefore regress brain activity in each voxel on this difference (as well as other variables described in Appendix II). We used a random-effects group-level analysis[15] to identify activity positively correlated with this "expected reward prediction error" regressor. Figure IV shows the significant areas at a threshold of $p < .0005$ (uncorrected), areas that overlap considerably with

14. The dorsal limit of the nucleus accumbens is the horizontal plane passing under the caudate nucleus head from the inferior border of the lateral ventricle to the edge of the internal capsule. The lateral limit is the internal capsule. The medial limit is the diagonal band of Broca. The ventral limit is the anterior hypothalamic nucleus and the external capsule laterally. The posterior limit is the anterior border of the anterior commissure. The anterior limit begins where the anterior caudate head and putamen are clearly divided by the internal capsule. The nucleus accumbens was defined bilaterally in this manner on the individual high-resolution anatomical images in Talairach space (Talairach and Tournoux 1988).

15. See Appendix II for details.

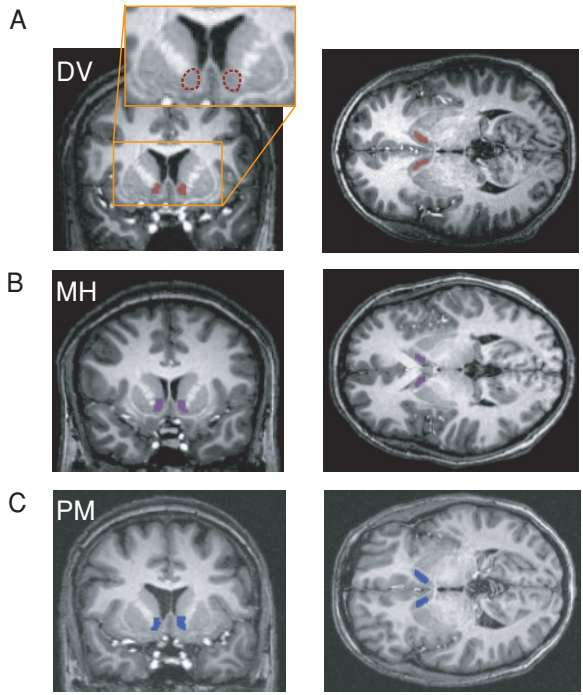

FIGURE III
The Nucleus Accumbens Defined Anatomically in Three Subjects
(A–C) Regions defined in three subjects (DV in A, MH in B, and PM in C). Coronal sections (left, $y = +7$) and horizontal sections (right, $z = +0$) are shown for each subject. The inset in (A) shows the outlined nucleus accumbens for subject DV. The nucleus accumbens was defined by anatomical landmarks using the algorithm described in Neto et al. (2008). Data are shown in radiological convention with the right hemisphere on the left in the coronal sections and on the bottom in the horizontal sections.

the typical anatomically defined nucleus accumbens. Unlike our anatomical ROIs, which were defined in individual subjects, functional ROIs were defined at the group level. In order to make the definition of the ROI statistically independent of later tests of the axioms, we split the data set into two halves, data sets a and b, with set a containing odd-numbered scanning runs for the first session and even-numbered runs for the second session, and set b containing all other runs. We then collect data from set b using the ROI defined using data from set a, and vice versa.

The next task is to combine BOLD data from the voxels identified in an ROI into a single time series. We do this by averaging across all voxels in an ROI and then converting the average signal

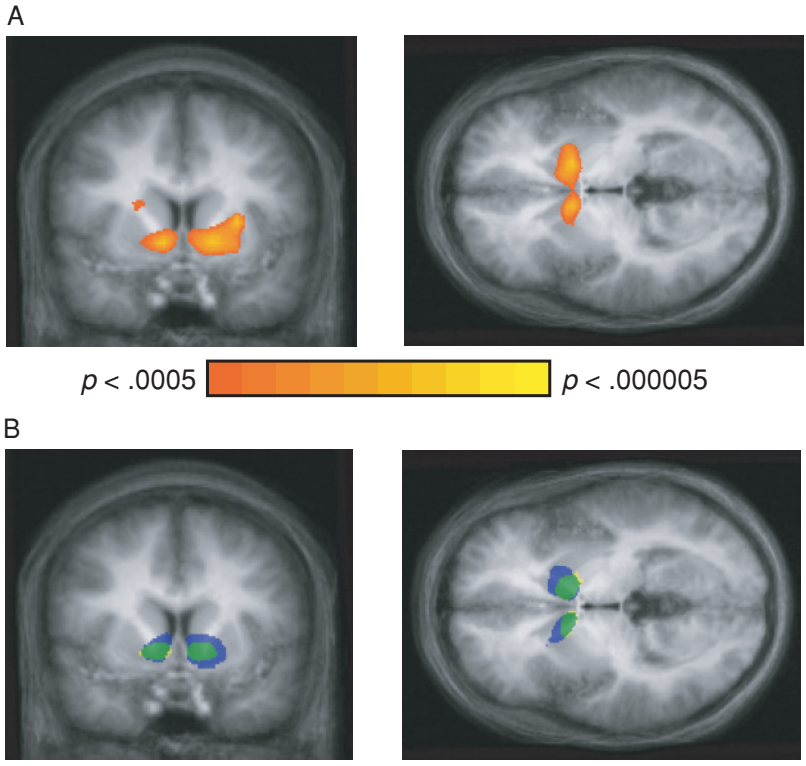

FIGURE IV

Group Analysis Showing the Brain Areas in Which Activity Is Correlated with "Expected Reward Prediction Error"

(A) A region of correlation ($p < .0005$, uncorrected), which overlaps considerably with the anatomically defined nucleus accumbens, can be seen in a coronal (left, $y = +7$) and a horizontal section (right, $z = +0$), overlaid on a mean normalized anatomical image. (B) When the data set is split in half, independent regions of correlation ($p < .005$, uncorrected) are defined for data set *a* (blue), odd-numbered runs in the first session and even-numbered runs in the second, and data set *b* (yellow), the rest of the runs. The region of overlap between the two regions is indicated as (green). The random-effects analyses include regressors for the options onset, button press, outcome onset, and a parametric variable at the time of the outcome onset. This variable is computed as the difference between the outcome and the expected value of the lottery in dollars. All regressors are one time point convolved with the canonical double-gamma hemodynamic response function. Data are shown in radiological convention with the right hemisphere on the left in the coronal sections and on the bottom in the horizontal sections.

in each trial to percent signal change according to standard fMRI protocol; by using the last two time points of the fixation period as a baseline and dividing the signal in a trial by the average signal in those two time points.

(3) Constructing δ. In an ideal world, we would use a within-subjects design to test the axioms on an individual-by-individual basis. However, fMRI data are still too noisy for such a test. We therefore combine data across subjects, effectively treating our data as all coming from a single person. In general, finding the axioms satisfied at such a group level is neither necessary nor sufficient to say that they are satisfied at the individual level. Effectively, we rely on an assumption of homogeneity—that subjects order prizes and lotteries the same way. In this case, this only requires that all subjects find winning \$5 more rewarding than losing \$5, and that all subjects expect a greater reward from lotteries with higher objective probability of winning the better prize. (Although we acknowledge the limitation of this approach, we also note that this assumption of homogeneity has now been used in literally tens of thousands of papers.)

We now use our time series data to provide estimates of δ. We do this by regressing the time series of dopamine activity on a sequence of dummy variables for each of the eight lottery–prize pairs in the experiment, and using the estimated coefficients as an estimate of activity caused by each pair. Specifically, we use a separate dummy to represent the event of getting each given prize from each given lottery (eight dummies). There is therefore one dummy variable that takes the value 1 when the \$5 prize is revealed from the lottery that had a 50% chance of +\$5 and 50% chance of −\$5, another that takes the value 1 when the −\$5 is revealed from the same lottery, and so on. Dummies take the value 1 for a time window starting 4 TRs (5 seconds) and finishing 10 TRs (12.5 seconds) after a prize has been revealed. This time window is chosen to take into account the hemodynamic response, the lag between brain activity and the change in blood chemistry that can be detected by fMRI. The coefficients on these dummies we use as our estimates δ. Notationally, we will use $\hat{\delta}(x, p)$ to indicate the estimated parameter on the dummy that is set to 1 when prize x is received from the lottery that gives the prize \$5 with probability p. In addition, we include scan-level dummies to capture scan-specific effects (i.e., a separate dummy for each scan run—remembering that each subject takes part in multiple scans). The regression is performed using ordinary least squares, with Huber/White/sandwich robust standard errors (Huber 1967; White 1980).

(4) Testing the Axioms. We now face the challenge of using our estimates, $\hat{\delta}$, to test our axioms. If these observations were

deterministic then the test would be easy—by Theorem 1, all we would have to do would be to take the numbers $\hat{\delta}(x, p)$ and check whether Coherent Prize Dominance, Coherent Lottery Dominance, and No Surprise Equivalence hold. Unfortunately, $\hat{\delta}(x, p)$ are noisy estimates of underlying brain activity $\delta(x, p)$. Ideally we would like to take the route of standard statistical hypothesis testing, by stating the null hypothesis that the underlying parameters $\delta(x, p)$ violate our axioms. We would then wish to calculate the probability of observing $\hat{\delta}(x, p)$ given this null hypothesis. Such tests rely on our ability to use the null hypothesis to generate a suitable test statistic. In the case of simple linear restrictions, this presents no difficulty. However, in this case, it is extremely difficult to do. We therefore take an alternative approach, consisting of pairwise Wald tests. In particular, for each $\{x, p\}, \{y, q\} \in A$, we perform a test of the restriction that $\delta(x, p) = \delta(y, q)$. If we cannot reject this hypothesis, we treat the two values as equal. If we can, then we treat them as unequal in the same direction as the relation of $\hat{\delta}(x, p)$ and $\hat{\delta}(y, q)$.

We are now in a position to test our axioms. Let the function $\text{sign}(x)$ equal $+$ if x is positive, $-$ if x is negative, and $=$ otherwise. The test of our axioms can therefore be written as follows:

- Axiom 1. Coherent Prize Dominance:

$$\begin{aligned}&\text{sign}(\delta(5, 0.25) - \delta(-5, 0.25))\\ &= \text{sign}(\delta(5, 0.5) - \delta(-5, 0.5))\\ &= \text{sign}(\delta(5, 0.75) - \delta(-5, 0.75)).\end{aligned}$$

- Axiom 2. Coherent Lottery Dominance:

$$\begin{aligned}&\text{sign}(\delta(5, 0.25) - \delta(5, 0.5))\\ &= \text{sign}(\delta(-5, 0.25) - \delta(-5, 0.5))\end{aligned}$$

and

$$\begin{aligned}&\text{sign}(\delta(5, 0.25) - \delta(5, 0.75))\\ &= \text{sign}(\delta(-5, 0.25) - \delta(-5, 0.75))\end{aligned}$$

and

$$\begin{aligned}&\text{sign}(\delta(5, 0.5) - \delta(5, 0.75))\\ &= \text{sign}(\delta(-5, 0.5) - \delta(-5, 0.75)).\end{aligned}$$

- Axiom 3. No Surprise Equivalence:

$$\delta(5, 1) = \delta(-5, 0).$$

One thing to note is that these criteria would be met by any δ function that ordered prizes and lotteries consistently—for example, one that ranked losing \$5 above winning \$5, or that was everywhere constant. We therefore also provide a more restrictive test based on the idea that reward should be increasing in monetary value, and that predicted reward should be increasing in lottery expected value, which we refer to as *Strong Coherent Prize Dominance* and *Strong Coherent Lottery Dominance*.

V. Experimental Results

V.A. Subjects

Fourteen paid volunteers participated in the experiment (nine women, five men, all right-handed, mean age = 26.0 years (S.D. 8.1 years)). All participants gave informed consent in accordance with the procedures of the University Committee on Activities Involving Human Subjects of New York University. All subjects completed at least 13 scans (of approximately 8 minutes each) over two sessions. Excessive motion during the experiment rendered the fMRI data for two subjects unusable.[16] Of the remaining twelve subjects, all completed 14–16 scans, with most subjects ($n = 9$) completing 8 scans in each session.[17]

Subjects earned an average of \$125 (S.D. \$39) per session including the endowment and show-up fee. One subject lost the entirety of the endowment during the second scanning session, and the final scan of that session is excluded from analysis. That subject was also the only subject who failed to respond within the required time window on more than 2 trials, missing 6 trials in total. The average reaction time for successful responses was 382 ms (S.D. 103 ms). In total, 17 trials were missed of a possible 3,024. Due to a programming error, a further 4 trials erroneously resulted in missed trials, despite the response being within the specified time window. These 4 trials were excluded from further

16. Both subjects had nine scans with at least 0.1 mm per TR or 0.1° per TR average motion in any direction; no other subject had more than three scans with as much motion. These subjects were excluded from all further analysis, as is common practice in fMRI studies.

17. To an experimental economist, the small number of experimental subjects in this and other neuroscientific studies may be disturbing. This is particularly so given that there are significant individual differences in neurological structure. Unfortunately, it is a necessary feature of current experiments, given technological constraints. Because dopaminergic responses are of interest to many research groups, robustness of results is uncovered through independent replication.

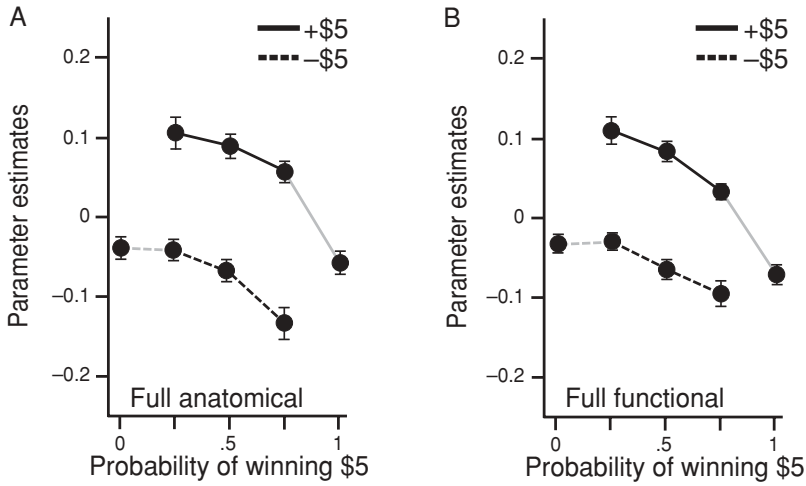

FIGURE V

Parameter Estimates Using the Full Time Window (TR 4–10)

Parameter estimates are shown for regions of interest in the nucleus accumbens defined both (A) anatomically and (B) functionally. Error bars show ±1 robust standard errors. Regressions run on 74,088 data points (TRs) from 2,975 trials, 189 scans, and 12 subjects.

analysis. Subjects usually chose the lottery with the higher expected value, with 6 subjects making such a choice on every trial. In total, 28 choices of lotteries in the decoy set were made. Thus out of a possible 3,024 trials in 189 completed scans, 2,975 trials are included in further analysis.

V.B. Results

Figure VA shows the parameter estimates $\hat{\delta}$ for the anatomically defined ROI. These estimates are shown in the graphical format introduced in Section III.A. For each prize, we plot a line showing the parameter estimates when that prize is received from each observed lottery. Recall from Section III.B that our three axioms are equivalent to three properties of these graphs: that the lines do not cross, that they are co-monotonic, and that $\hat{\delta}(-5, 0)$ is equal to $\hat{\delta}(5, 1)$.

An examination of Figure VA suggests that activity in the anatomically defined nucleus accumbens is consistent with Strong Coherent Prize Dominance, Strong Coherent Lottery Dominance, and No Surprise Equivalence: the line for the +$5 prize lies everywhere above that for the −$5 prize, and both lines are

TABLE I
STATISTICAL TESTS ON THE DIFFERENCE BETWEEN PARAMETER ESTIMATES

	Anatomical ROI		Functional ROI	
	Sign	Prob.	Sign	Prob.
A1: Coherent prize dominance				
{5, 0.25}–{−5, 0.25}	+	.0	+	.0
{5, 0.50}–{−5, 0.50}	+	.0	+	.0
{5, 0.75}–{−5, 0.75}	+	.0	+	.0
A2: Coherent lottery dominance				
{−5, 0.50}–{−5, 0.25}	=	11.1	−*	5.3
{5, 0.50}–{5, 0.25}	=	73.9	−*	9.7
{−5, 0.75}–{−5, 0.50}	−	.9	−*	9.3
{5, 0.75}–{5, 0.50}	−	4.4	−	.2
{−5, 0.75}–{−5, 0.25}	−	.0	−	.0
{5, 0.75}–{5, 0.25}	−	4.8	−	.1
A3: No surprise equivalence				
{−5, 0}–{5, 1}	=	34.0	+	.7

Notes. The Prob. column reports the probability that each hypothesis holds according to a Wald test of linear restriction using robust standard errors. The Sign column shows a + or − if the test is significant in that direction at the 5% level, with a * appended if significant at the 10% level. Regressions run on 74,088 data points (TRs) from 2,975 trials, 189 scans, and 12 subjects.

downward-sloping. Furthermore, $\hat{\delta}(-5, 0)$ looks very similar to $\hat{\delta}(5, 1)$, suggesting that No Surprise Equivalence might also hold.

Table I performs the statistical tests discussed in Section IV.B(4) above. These largely confirm that the data satisfy the three axioms. The evidence for Strong Coherent Prize Dominance is overwhelming: the hypothesis that $\hat{\delta}(-5, p) = \hat{\delta}(5, p)$ is rejected at below the 0.1% level for each $p \in \{0.25, 0.5, 0.75\}$ (with $\hat{\delta}(-5, p) < \hat{\delta}(5, p)$). $\hat{\delta}(-5, 0)$ is not significantly different from $\hat{\delta}(5, 1)$, so No Surprise Equivalence also holds. Coherent Lottery Dominance also holds, but only in the weak sense: for neither prize is $\hat{\delta}(x, 0.25)$ statistically different from $\hat{\delta}(x, 0.5)$, but, for both prizes $\hat{\delta}(x, 0.5)$ is significantly higher that $\hat{\delta}(x, 0.75)$ and $\hat{\delta}(x, 0.25)$ is significantly higher that $\hat{\delta}(x, 0.75)$. Thus, our key result is that the BOLD signal recorded from the anatomically defined nucleus accumbens region meets the necessary and sufficient criteria required of a reward prediction error encoder. Moreover, the ordering of prizes and lotteries is as one would expect—more money is rated as "more rewarding" than less money, and lotteries with a higher probability of winning $5 have a higher predicted reward.

V.C. Robustness Tests

Functionally Defined ROI. Figure VB shows the parameter estimates for the functionally defined ROIs (the statistical tests are also reported in Table I). In most major respects, the results are the same: the line for the +$5 prize lies everywhere above that for the −$5 prize, and both lines are downward-sloping. In fact, for the functionally defined ROI, Axiom 2 holds in the strong as well as the weak sense, as both lines are significantly downward-sloping between all points. However, for this ROI, No Surprise Equivalence does not hold: the amount of activity observed when $5 is lost for sure is significantly higher than for when $5 is won for sure.

Temporal Window. As a second check of the robustness of our results, we examine the temporal window, or the time within each trial during which $\hat{\delta}$ was estimated. To do this we construct a plot of the average BOLD activity as a function of time for trials of each lottery–prize pair. This is shown in Figure VI for both anatomically and functionally defined ROIs. The temporal window used in the preceding analysis of $\hat{\delta}$ is shown in gray. For our results to be robust to different time windows, we would require that the ordering of these lines not change through the course of the trial. Figure VI suggests that this is in fact *not* the case: Early time periods (immediately after the lottery outcome is revealed) seem to show clear differentiation between lotteries when the *positive* prize is received, whereas the later time periods show differentiation between lotteries when the *negative* prize is received. Moreover, activity for the degenerate lotteries seems to follow a rather different pattern from that seen for nondegenerate lotteries. For all nondegenerate lotteries, BOLD activity peaks soon after the prize has been received, then falls. For the degenerate lotteries, activity shows no spike in response to the revelation of the prize.

In order to further examine this apparent temporal variation in $\hat{\delta}$, we reestimate our eight parameters on two different temporal windows: an early window (consisting of TR 4–6, where TR 0 is the time at which outcome is displayed) and a late window (TR 7–10) for both the anatomically and functionally defined ROIs. These estimates are shown in Figures VII and VIII. Although still satisfying Coherent Prize Dominance, the early window graph (Figure VII) suggests that Coherent Lottery Dominance does not hold in this period—the positive prize line remains downward-sloping, whereas the negative prize line is largely flat. In contrast,

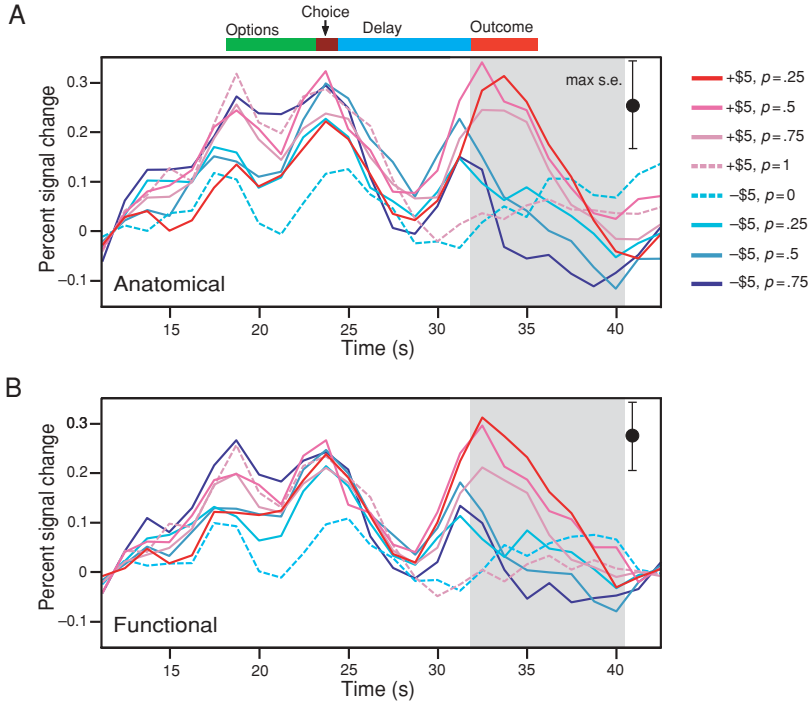

FIGURE VI

Group-Level Time Courses Are Shown Averaged over All Voxels in a Region of Interest for Twelve Subjects and Then Replotted as Trial Averages

Trial averages are shown for regions of interest in the nucleus accumbens defined both (A) anatomically and (B) functionally. Trial averages are color-coded by lottery–prize pair with the probability of winning $5 indicated for each. The largest standard error for any timepoint for any lottery–prize pair is shown at right. The timeline above the plot shows the expected time of responses to each period using a 5-second (4 TRs) lag to account for the delay in the hemodynamic response function. Peak responses typically coincided with the options onset, button press, and outcome onset (hereafter referred to as TR 0). The time window (TR 4–10) used for the analysis in Section V.B is shown in gray.

although Coherent Lottery Dominance does seem to hold approximately in the late window (Figure VIII), it seems that the responsiveness of activity to changes in lottery is much stronger for the negative prize than the positive prize. This pattern is borne out by Figure IX, which shows how the difference between $\hat{\delta}(x, 0.25)$ and $\hat{\delta}(x, 0.75)$ changes with the estimation period for each prize for the anatomically defined ROI. The figure plots these differences for estimates made on different 2-TR windows, starting at the TR indicated on the x-axis. Thus the graph provides an indication of how the slope of the $\hat{\delta}(5, x)$ and $\hat{\delta}(-5, x)$ lines varies with

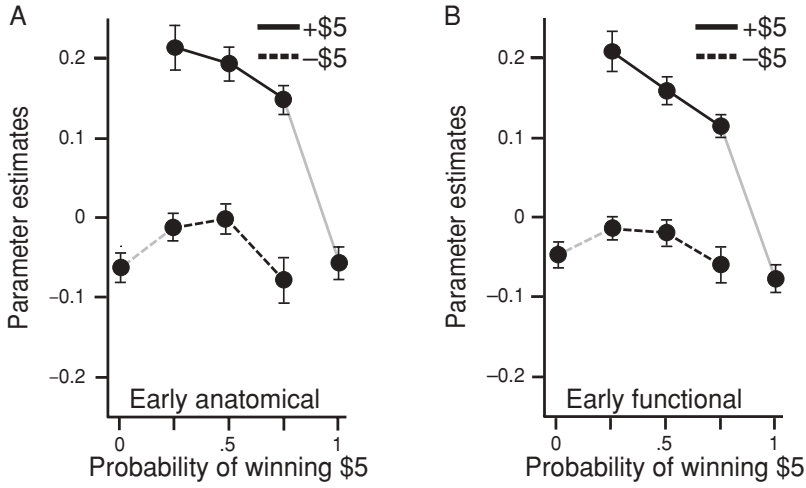

FIGURE VII
Parameter Estimates Using the Early Time Window (TR 4–6)

Parameter estimates are shown for regions of interest in the nucleus accumbens defined both (A) anatomically and (B) functionally. Error bars show ±1 robust standard errors. Regressions run on 74,088 data points (TRs) from 2,975 trials, 189 scans, and 12 subjects.

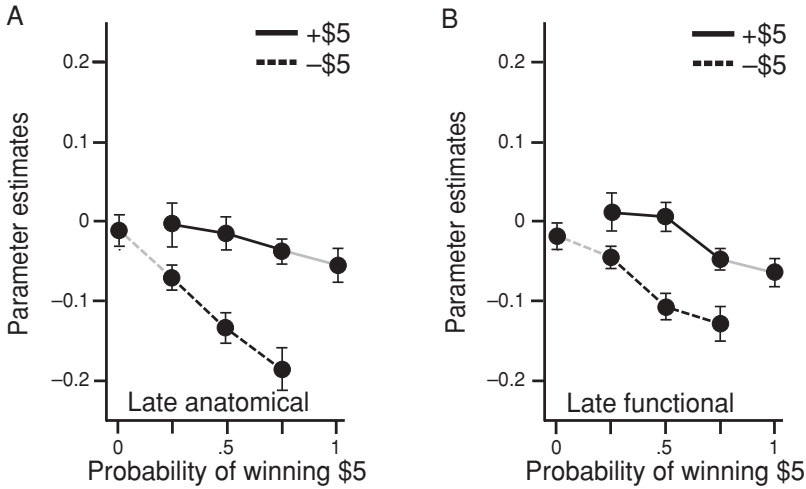

FIGURE VIII
Parameter Estimates Using the Late Time Window (TR 7–10)

Parameter estimates are shown for regions of interest in the nucleus accumbens defined both (A) anatomically and (B) functionally. Error bars show ±1 robust standard errors. Regressions run on 74,088 data points (TRs) from 2,975 trials, 189 scans, and 12 subjects.

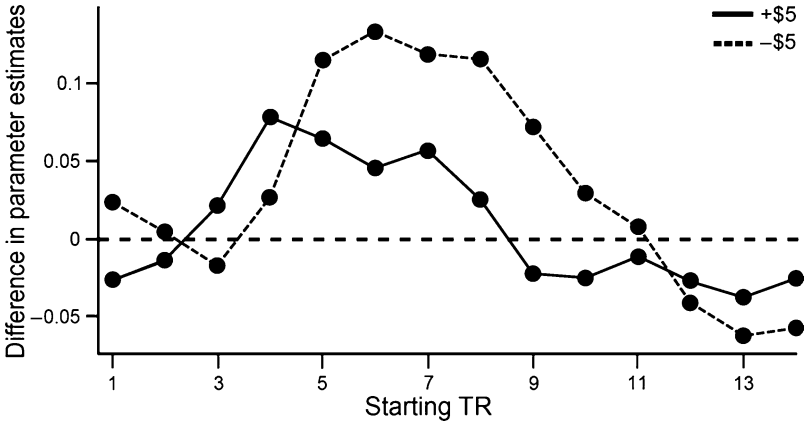

FIGURE IX

Difference in Parameter Estimates of Activity in the Anatomically Defined Nucleus Accumbens between Cases Where Each Prize Is Received from the 25% Lottery and the 75% Lottery

Each point represents this difference for a sliding 2-TR window starting at the TR indicated on the x-axis, where TR 0 is the time of outcome onset and TR 4–10 is the time window used for prior analyses.

the time window considered. This graph indicates that the peak differentiation between lotteries occurs around TR 4 for the positive prize and around TR 6 for the negative prize. Perhaps even more surprisingly, the size of the differentiation for the negative prize is also roughly twice as large as that for the positive prize.[18] The economic and neurobiological implications of this result are discussed below.

It should be noted that the original time window we selected is not an *ad hoc* "knife edge" case for which the axioms hold. First of all, the time window was selected in order to match what is known about standard hemodynamic response functions. Second, our original results are robust to *balanced* changes in the time window—that is, changes in the time window that change the start and end point of the window while keeping the central point the same.[19]

18. It should be noted that interpreting differences in magnitudes in BOLD signals is a complicated matter, particularly when increases and decreases in the signal are compared. Thus the difference in magnitude should be interpreted with caution.

19. In such windows Axioms 1 and 3 hold, whereas support for Axiom 2 is mixed.

V.D. Discussion

Our results can be summarized as follows:

1. Strong and robust support for Axiom A1: Coherent Prize Dominance.
2. Support for Axiom A2: Coherent Lottery Dominance for the average signal across the full time window of the hemodynamic response.
3. Weak support for Axiom A3: No Surprise Equivalence in the anatomical but not the functional ROI.

The results of this study are broadly a success for proponents of the DRPE hypothesis. The average BOLD signal measured by fMRI from the anatomically defined nucleus accumbens satisfies the three necessary and sufficient conditions for a reward prediction–error encoder, although support is weak for the third axiom. Certainly, this renders false previous claims that nucleus accumbens activity (as measured by BOLD) cannot encode a reward prediction error. In light of the axioms being satisfied, there is a strong theoretical basis for using dopaminergic measurements to define consistent measurements of "reward" and "belief" based on neurobiological measurements of activity in this area. In our experiment, these measurements satisfy basic rationality conditions: more money is more rewarding than less money, and lotteries have a higher predicted reward if they have a higher probability of winning the higher prize. Thus, our work rigorously tests and confirms the conclusions of previous authors who have found evidence indicative of the DRPE hypothesis in fMRI data (McClure, Berns, and Montague 2003; O'Doherty et al. 2003, 2004; Abler et al. 2006; Li et al. 2006; Pessiglione et al. 2006; D'Ardenne et al. 2008).

Note that, although we do not axiomatically charactize salience or hedonia models of dopamine activity, our results do not look promising for these other potential explanations for the information encoded in nucleus accumbens activity. Recall that, from Section III.C, the key axiom that appears to be inconsistent with the "hedonia" hypothesis was No Surprise Equivalence: hedonia would imply that better prizes would lead to higher responses even from degenerate lotteries. In our data, either No Surprise Equivalence holds (in the anatomically defined ROI) or we find that the *worse* prize gives rise to higher dopamine activity (functional ROI). Neither of these cases appears consistent with the hedonia hypothesis.

Our data also seem inconsistent with the possibility that the nucleus accumbens encodes salience. Again, recall from Section III.C that a standard reading of the salience hypothesis would imply that dopamine activity should lead to a failure of Coherent Prize Dominance. From the lottery $p = .25$, winning \$5 is more surprising, and so arguably more salient than losing \$5, so winning \$5 should lead to a higher dopamine response. From the lottery $p = .75$, losing \$5 is more salient than winning \$5, so losing \$5 should lead to the higher dopamine response. We find no evidence for such an effect.

The success of the DRPE hypothesis is largely robust to the choice of functional or anatomical ROI. In both cases Coherent Prize Dominance and Coherent Lottery Dominance hold. The only difference between the two results is that No Surprise Equivalence holds in the anatomical ROI and not in the functional ROI. An examination of Figure VI suggests that this result may be part of a richer story involving the degenerate lottery, which has not yet received attention in either neurobiological or economic circles. Clearly, the time course of activity following the revelation of prizes is very different for the degenerate lotteries than for all nondegenerate lotteries. Although revelation from the nondegenerate lotteries leads to a sharp increase in BOLD activity, followed by a gradual decline in all cases, revelation for the degenerate lotteries leads to a much slower, gentler increase in activity for both the +\$5 and −\$5 prizes. For the anatomical ROI, the path is the same for both prizes, whereas for the functional ROI, the response for the −\$5 line is somewhat higher than that for the +\$5. This result suggests that the degenerate lotteries are treated differently at a neurological level than are nondegenerate lotteries.

Perhaps the most novel feature of the data is that, although average activation for the entire time window satisfies the DRPE hypothesis, this seems to be due to the amalgamation of two different processes, each with different temporal dynamics. This result supports earlier controversial theoretical proposals (Daw, Kakade, and Dayan 2002; Bayer and Glimcher 2005; Bayer, Lau, and Glimcher 2007), which hypothesized that dopamine responses may be asymmetric—recording positive but not negative reward prediction error. Our findings raise the possibility that the nucleus accumbens is indeed receiving, and possibly amalgamating, signals from two different processes that, between them, provide an encoding of an RPE signal. A high priority in future research is to

understand the robustness and significance of the distinct pattern of dopaminergic responses to losses and gains that we identify.

As we note above, the observations that we make have to do with activity in the nucleus accumbens, and not dopaminergic activity *per se*. Thus, we cannot conclude from these findings that *dopamine* is an RPE encoder. In fact, the evidence we find for two different systems points to the possibility that dopamine may only be encoding part of the RPE signal we observe here, as suggested in Daw, Kakade, and Dayan (2002) and Bayer and Glimcher (2005) and a recent detailed proposal by Dayan and Huys (2009). If this is the case, then the signal we observe could reflect activity induced in part by dopamine and in part by some other source that may serve as the negative RPE encoder. To say more about the role of dopamine and RPE, one would have to perform more direct measurements of dopamine, such as single-unit recording from dopamine neurons in monkeys. We see such a project as important future research.

VI. Conclusions

This paper presents the first use of an axiomatic representation theorem to test a neurobiological hypothesis using neurobiological data. We show that BOLD activity measured by fMRI in the dopamine-rich nucleus accumbens can be modeled as encoding a reward prediction error—the difference between the experienced and predicted rewards of an event. In doing so, we believe that this paper makes three contributions. First, it provides a concrete answer to the question of whether activity in the nucleus accumbens can encode a reward prediction error. Second, it increases the tools that economists have for studying economic behavior. Third, it introduces the tools of axiomatic modeling into the study of neuroscience.

Promising as our results are, they do not immediately advance our understanding of choice, the acid test of neuroeconomic progress proposed by Bernheim (2009). Yet they point the way to just such advances, in particular through the potential of dopaminergic measurements to provide fresh insights into the evolution of beliefs and of expectation-based reference points. Given that the DRPE hypothesis holds, we can now define both dopaminergic reference points (the expected reward of an event) and beliefs (the probabilities attached to states of the world that would generate such an expectation). The next stage of

our research agenda is to link these concepts to the equivalent constructs in standard decision theory—via experiments that relate dopamine activation to choice. If such a link exists, then dopamine can provide a new tool for understanding how beliefs and reference points evolve, rather than having to infer this from choice data alone.[20]

Given their importance to play, understanding of beliefs is particularly important in repeated games (Stahl and Wilson 1995; Cheung and Friedman 1997; Fudenberg and Levine 1998). In this arena, dopaminergic evidence will strengthen the arsenal of belief elicitation techniques. Nyarko and Schotter (2002) were able to explain play in various games far better using beliefs estimated from an incentive-compatible mechanism than using model-based estimates. Rutström and Wilcox (2006) provide an opposing example in which model-estimated beliefs are superior. In contrast to incentive-compatible mechanisms, dopaminergic techniques offer a potential window into beliefs that does not interrupt the course of play.

With respect to methodology, it is our belief that the axiomatic approach has a significant role to play in the field of behavioral neuroscience for the reasons discussed in more detail in Caplin and Dean (2008a). This paper provides a proof of method, by using this approach to provide clear answers to a previously open question within neuroscience—whether or not activity in the nucleus accumbens encodes a reward prediction error signal.

Until now, model testing, comparison, and improvement in neuroscience has taken place largely through a regression-based approach, in which highly parameterized models of reward, belief, and learning have been correlated with brain activity. In essence, this approach constitutes a form of gradient-descent through modeling space toward what is hoped to be a globally best model. We believe that the axiomatic approach, which has characterized so

20. One open question is the extent to which our results will generalize beyond the simple experimental environment tested here. For example, do we know that dopamine will respond the same way when there is a longer gap between the predictive stimulus and reward, or if probabilities are subjective rather than objective? To some extent these are open questions, though previous studies give some guide. Gallistel and Gibbon (2000) show that dopamine does still seem to encode a DRPE if signals and reward are temporally separated, as long as there is not too much variation in the length of the intervals. Moreover, many previous studies have attempted to test the DRPE hypothesis in environments in which subjects have to learn probabilities from past rewards drawn from an unknown distribution—which is much closer to the idea of subjective probabilities than it is to objective probabilities. Of course, none of these studies test our axioms directly.

much of economic modeling during this same period, can provide a powerful alternative to this nonstructural tradition, which at present dominates neurobiological research. By clearly encapsulating conditions of necessity and sufficiency for describing a class of models, the axiomatic approach allows us to ask not whether a particular model fits well but rather whether an entire class of models can be falsified. What makes the axiomatic approach uniquely powerful is that it presents a model in the clearest and most easily falsifiable form possible. This represents a fundamental contribution that the economic approach can make to neuroscience and one that we believe can have broad impact in that discipline. Economic tools can help shape future neurobiological discourse.

In summary, the present results indicate that brain activity in the nucleus accumbens, as measured by fMRI, meets the criteria of necessity and sufficiency for carrying a reward prediction error signal. This fundamentally strengthens the conclusion that reward prediction error-based learning of value occurs in the human brain. Axiomatic modeling, an approach that offers many advantages over traditional neurobiological modeling, which is often necessarily *ad hoc* in nature, can be used to provide novel insights into brain function. In the converse direction, our broad confirmation of the DRPE hypothesis suggests concrete ways in which neurobiology will be able to return the compliment by providing new insights into economic behavior.

Appendix I: Instructions

We are interested in understanding how people choose and value uncertain financial options, like lotteries. You will be asked to make a series of choices between lotteries. For example, one lottery might be the one pictured at right in Figure A.1. When you play this lottery, you have a 50% probability of gaining $5 (of real money) and a 50% probability of losing $5. Before you start the game, we will give you $100 in real money. Put it in your pocket. You will play the game with this money. If you win more money over the course of the game, we will give you those winnings when you finish. If you lose money during the game, you will return it to the experimenter and you can keep the rest of the $100. If at any point in the game, you lose all of your $100, the game ends and you must return the money. You will play 8 rounds of 16 trials each. At the start of each trial, a white cross appears at the center

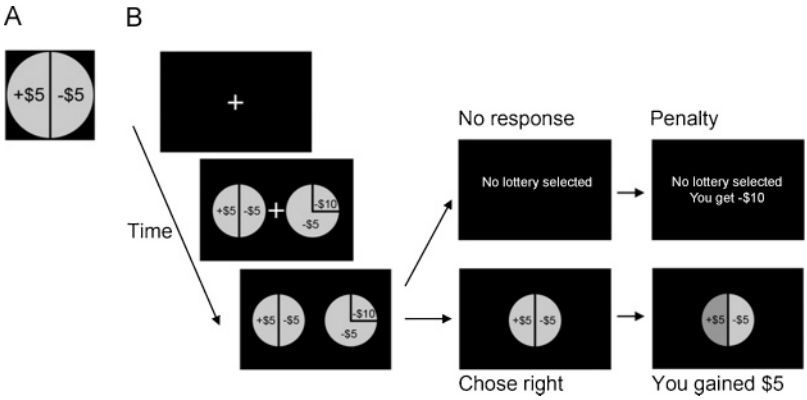

FIGURE A.1
Figures Included in the Instructions Given to Subjects before the Experiment (Appendix I)
(A) Example lottery. (B) Task diagram.

of the screen (shown in Figure A.1). Then two lotteries will be presented on the screen. Your task is to decide which of the two lotteries you would prefer to play with the $100 in your pocket. The amounts on the screen are in real money, which you can win and lose on every trial. Press the left button for the lottery on the left, the right for the lottery on the right. The lottery you chose will then be shown in the center of the screen. There is no right answer. We just want to know what lottery you would prefer to play. The computer then rolls the dice and tells you which prize you received. In the example below, you would have won $5 of real money. After each block of trials, the computer tells you how much you won or lost for that block and what your total earnings are up to that point in the game. If you do not make a choice within the 1.25-second time limit, the trial will end and the screen will display "No Lottery Selected" and you will receive a penalty of −$10 (the worst prize; shown in Figure A.1). Regardless of your performance in the game, you will be paid a show-up fee of $35. If you decide to quit playing the game before its conclusion, you will be paid the show-up fee but you must return the $100. Good luck!

APPENDIX II: DETAILS OF IMAGING PROTOCOL AND DATA PROCESSING

A. Imaging

We used a Siemens Allegra 3-Tesla head-only scanner equipped with a head coil from Nova Medical to collect the blood

oxygen level–dependent (BOLD) signal. We collected 23 axial slices of T2*-weighted functional images with an echo planar imaging (EPI) pulse sequence. Our slices were oriented parallel to the anterior–posterior commissure (AC–PC) plane. Sequence parameters were as follows: 23 axial slices, repetition time (TR) = 1.25 s, echo time (TE) = 30 ms, flip angle = 73°, 64 × 64 acquisition matrix, in-plane resolution = 3 × 3 mm, field of view (FOV) = 192 mm, slice thickness = 3 mm). Each scan consisted of sixteen 30-second trials with an additional fixation period of 15 seconds at the end of each scan, for a duration of 8 minutes and 15 seconds per scan. Thus each scan consisted of 396 images. We also collected high-resolution T1-weighted anatomical images using a magnetization-prepared rapid-acquisition gradient echo (MP-RAGE) pulse sequence (144 sagittal slices, TR = 2.5 s, TE = 3.93 ms, inversion time (TI) = 900 ms, flip angle = 8°, 1 × 1 × 1 mm, 256 × 256 matrix in a 256-mm FOV). The display was projected onto a screen at the back of the scanner and subjects viewed the display through a mirror attached to the head coil. To minimize head movements, subjects' heads were stabilized with foam padding.

B. Data Analysis

Data were analyzed with the BrainVoyager QX software package (Brain Innovation) with additional analyses performed in MATLAB (MathWorks) and Stata (StataCorp). Preprocessing of functional images included discarding the first four images to avoid T1 saturation effects, sinc-interpolation for slice scan time correction, intersession and intrasession 3D motion correction using six-parameter rigid body transformations, and linear trend removal and high-pass filtering (cutoff of 3 cycles per scan) to remove low-frequency drift in the signal. Images were coregistered with each subject's anatomical scan, rotated to the AC–PC plane, and transformed into Talairach space (Talairach and Tournoux 1988) using trilinear interpolation. For group-level random-effects analyses only, data were also spatially smoothed with a Gaussian kernel of 8 mm (full width at half maximum). We used the summary statistics approach to test when the mean effect at each voxel was significantly different from zero across subjects. We modeled the time course of activity as transient responses at the following times convolved with the canonical double-gamma hemodynamic impulse response function (peak = 6 s, undershoot peak = 15 s, peak–undershoot ratio = 6): lotteries onset, button press, and

outcome onset. We also included a parametric regressor at outcome onset equal in magnitude to the difference between the outcome and the expected value of the lottery in dollars. This regressor allowed us to perform a traditional regression analysis on our data.

NEW YORK UNIVERSITY
BROWN UNIVERSITY
NEW YORK UNIVERSITY AND CENTER FOR NEUROECONOMICS, NEW YORK UNIVERSITY
CENTER FOR NEURAL SCIENCE, NEW YORK UNIVERSITY

REFERENCES

Abler, Birgit, Henrik Walter, Susanne Erk, Hannes Kammerer, and Manfred Spitzer, "Prediction Error as a Linear Function of Reward Probability Is Coded in Human Nucleus Accumbens," *NeuroImage*, 31 (2006), 790–795.
Bayer, Hannah M., and Paul W. Glimcher, "Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal," *Neuron*, 47 (2005), 129–141.
Bayer, Hannah M., Brian Lau, and Paul W. Glimcher, "Statistics of Midbrain Dopamine Neuron Spike Trains in the Awake Primate," *Journal of Neurophysiology*, 98 (2007), 1428–1439.
Bernheim, B. Douglas, "On the Potential of Neuroeconomics: A Critical (but Hopeful) Appraisal," *American Economic Journal: Microeconomics,* 1 (2009), 1–41.
Bernheim, B. Douglas, and Antonio Rangel, "Addiction and Cue-Triggered Decision Processes," *American Economic Review*, 94 (2004), 1558–1590.
Berridge, Kent C., and Terry E. Robinson, "What Is the Role of Dopamine in Reward: Hedonic Impact, Reward Learning, or Incentive Salience?" *Brain Research Reviews,* 28 (1998), 309–369.
Bossaerts, Peter, Kerstin Preuschoff, and Ming Hsu, "The Neurobiological Foundations of Valuation in Human Decision-Making under Uncertainty," in *Neuroeconomics: Decision Making and the Brain*, Paul W. Glimcher, Colin F. Camerer, Ernst Fehr, and Russell A. Poldrack, eds. (New York: Academic Press, 2008).
Caplin, Andrew, and Mark Dean, "Axiomatic Methods, Dopamine and Reward Prediction Error," *Current Opinion in Neurobiology*, 18 (2008a), 197–202.
——, "Axiomatic Neuroeconomics," in *Neuroeconomics: Decision Making and the Brain*, Paul W. Glimcher, Colin F. Camerer, Ernst Fehr, and Russell A. Poldrack, eds. (New York: Academic Press, 2008b).
——, "Dopamine, Reward Prediction Error, and Economics," *Quarterly Journal of Economics*, 123 (2008c), 663–702.
Charness, Gary, and Dan Levin, "When Optimal Choices Feel Wrong: A Laboratory Study of Bayesian Updating, Complexity, and Affect," *American Economic Review*, 95 (2005), 1300–1309.
Cheung, Yin-Wong, and Daniel Friedman, "Individual Learning in Normal Form Games: Some Laboratory Results," *Games and Economic Behavior*, 19 (1997), 46–76.
D'Ardenne, Kimberlee, Samuel M. McClure, Leigh E. Nystrom, and Jonathan D. Cohen, "BOLD Responses Reflecting Dopaminergic Signals in the Human Ventral Tegmental Area," *Science,* 319 (2008), 1264–1267.
Daw, Nathaniel D., Sham Kakade, and Peter Dayan, "Opponent Interactions between Serotonin and Dopamine," *Neural Networks,* 15 (2002), 603–616.
Dayan, Peter, and Quentin J. M. Huys, "Serotonin in Affective Control," *Annual Review of Neuroscience,* 32 (2009), 95–126.
Delgado, Mauricio R., Melinda M. Miller, Souheil Inati, and Elizabeth A. Phelps, "An fMRI Study of Reward-Related Probability Learning," *NeuroImage*, 24 (2005), 862–873.
Fudenberg, Drew, and David Levine, "Learning in Games: Where Do We Stand?" *European Economic Review*, 42 (1998), 631–639.

Gallistel, C. Randy, and John Gibbon, "Time, Rate and Conditioning," *Psychological Review*, 107 (2000), 289–344.
Hollerman, Jeffrey R., and Wolfram Schultz, "Dopamine Neurons Report an Error in the Temporal Prediction of Reward during Learning," *Nature Neuroscience,* 1 (1998), 304–309.
Huber, Peter J., "The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1 (Berkeley: University of California Press, 1967).
Huettel, Scott A., Allen W. Song, and Gregory McCarthy, *Functional Magnetic Resonance Imaging* (Sunderland, MA: Sinauer Associates, 2004).
Kahneman, Daniel, and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47 (1979), 263–292.
Knutson, Brian, and Sasha E. B. Gibbs, "Linking Nucleus Accumbens Dopamine and Blood Oxygenation," *Psychopharmacology*, 191 (2007), 813–822.
Knutson, Brian, and Richard Peterson, "Neurally Reconstructing Expected Utility," *Games and Economic Behavior*, 52 (2005), 305–315.
Köszegi, Botond, and Matthew Rabin, "'A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 121 (2006), 1133–1165.
Li, Jian, Samuel M. McClure, Brooks King-Casas, and P. Read Montague, "Policy Adjustment in a Dynamic Economic Game," *PlosONE*, 20 (2006), 103–113.
McClure, Samuel M., Gregory S. Berns, and P. Read Montague, "Temporal Prediction Errors in a Passive Learning Task Activate Human Striatum," *Neuron*, 38 (2003), 339–346.
McClure, Samuel M., David I. Laibson, George Loewenstein, and Jonathan D. Cohen, "Separate Neural Systems Value Immediate and Delayed Monetary Rewards," *Science,* 306 (2004), 503–507.
Mirenowicz, Jacques, and Wolfram Schultz, "Importance of Unpredictability for Reward Responses in Primate Dopamine Neurons," *Journal of Neurophysiology*, 72 (1994), 1024–1027.
Montague, P. Read, Peter Dayan, and Terrence J. Sejnowski, "A Framework for Mesencephalic Dopamine Systems Based on Predictive Hebbian Learning," *Journal of Neuroscience,* 16 (1996), 1936–1947.
Neto, Lia L., Edson Oliveira, Francisco Correia, and António G. Ferreira, "The Human Nucleus Accumbens: Where Is It? A Stereotactic, Anatomical and Magnetic Resonance Imaging Study," *Neuromodulation*, 11 (2008), 13–22.
Nyarko, Yaw, and Andrew Schotter, "An Experimental Study of Belief Learning Using Real Beliefs," *Econometrica*, 70 (2002), 971–1005.
O'Doherty, John P., Peter Dayan, Karl J. Friston, Hugo D. Critchley, and Raymond J. Dolan, "Temporal Difference Models and Reward-Related Learning in the Human Brain," *Neuron*, 38 (2003), 329–337.
O'Doherty, John P., Peter Dayan, Johannes Schultz, Ralf Deichmann, Karl Friston, and Raymond J. Dolan, "Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning," *Science*, 304 (2004), 452–454.
Pessiglione, Mathias, Ben Seymour, Guillaume Flandin, Raymond J. Dolan, and Chris D. Frith, "Dopamine-Dependent Prediction Errors Underpin Reward-Seeking Behavior in Humans," *Nature*, 442 (2006), 1042–1045.
Phillips, Paul E. M., Garret D. Stuber, Michael L. A. V. Heien, R. Mark Wightman, and Regina M. Carelli, "Subsecond Dopamine Release Promotes Cocaine Seeking," *Nature,* 422 (2003), 614–618.
Redgrave, Peter, and Kevin Gurney, "The Short-Latency Dopamine Signal: A Role in Discovering Novel Actions?" *Nature Reviews Neuroscience*, 7 (2006), 967–975.
Rutström, E. Elisabet, and Nathaniel T. Wilcox, "Stated Beliefs versus Empirical Beliefs: A Methodological Inquiry and Experimental Test" (http://www.uh.edu/econpapers/RePEc/hou/wpaper/2006-02.pdf, 2006).
Samuelson, William, and Richard Zeckhauser, "Status Quo Bias in Decision Making," *Journal of Risk and Uncertainty*, 1 (1988), 7–59.
Schultz, Wolfram, Paul Apicella, and Tomas Ljungberg, "Responses of Monkey Dopamine Neurons to Reward and Conditioned Stimuli during Successive

Steps of Learning a Delayed Response Task," *Journal of Neuroscience*, 13 (1993), 900–913.
Schultz, Wolfram, Peter Dayan, and P. Read Montague, "A Neural Substrate of Prediction and Reward," *Science*, 275 (1997), 1593–1599.
Stahl, Dale, and Paul Wilson, "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10 (1995), 218–254.
Talairach, Jean, and Pierre Tournoux, *Co-planar Stereotaxic Atlas of the Human Brian* (New York: Thieme Medical Publishers, 1988).
White, Halbert, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica,* 48 (1980), 817–830.
Wise, Roy A., "Dopamine, Learning, and Motivation," *Nature Reviews Neuroscience*, 5 (2004), 1–12.
Zink, Caroline F., Giuseppe Pagnoni, Megan E. Martin, Mukeshwar Dhamala, and Gregory S. Berns, "Human Striatal Response to Salient Nonrewarding Stimuli," *Journal of Neuroscience*, 23 (2003), 8092–8097.