DOPAMINE, REWARD PREDICTION ERROR, AND ECONOMICS*

Andrew Caplin and Mark Dean

The neurotransmitter dopamine has been found to play a crucial role in choice, learning, and belief formation. The best-developed current theory of dopaminergic function is the "reward prediction error" hypothesis—that dopamine encodes the difference between the experienced and predicted "reward" of an event. We provide axiomatic foundations for this hypothesis to help bridge the current conceptual gap between neuroscience and economics. Continued research in this area of overlap between social and natural science promises to overhaul our understanding of how beliefs and preferences are formed, how they evolve, and how they play out in the act of choice.

I. Introduction

Dopamine is a neurotransmitter—a substance that transmits information from one nerve cell of the brain to another. A long line of neurobiological studies have shown that dopamine plays a crucial role in many aspects of behavior that are important to economic decision-making (see Section II).[1] The goal of this paper is to characterize the "dopaminergic reward prediction error" (DRPE) hypothesis, which has become the standard model within neuroscience. This model asserts that neurons that contain this neurotransmitter release dopamine in proportion to the difference between the "predicted reward" and the "experienced reward" of a particular event. The model was developed following experimental work on monkeys by Schultz, Apicella, and Ljungberg (1993) and Mirenowicz and Schultz (1994). They established that the extent to which dopamine is released in response to a juice "reward" depends critically on whether or not receipt of the reward has already been signaled. In the presence of a clear prior cue, the release of dopamine shifts forward in time to coincide with the cue rather than with receipt of the reward.

Recent theoretical and experimental work on dopamine release has focused on the role this neurotransmitter plays in

---

* We thank Ed Glaeser, Paul Glimcher, and Antonio Rangel for their profound help at different stages in this research project. We also thank Roland Benabou, Alberto Bisin, Eric Dewitt, Kfir Eliaz, Guillaume Frechette, Greg Kaplan, John Leahy, Laurence Maloney, Efe Ok, Ariel Rubinstein, Robb Rutledge, and Andy Schotter for insightful comments.

[1] There are many excellent review articles, including Berridge and Robinson (1998), Schultz (2002), and Montague, Hyman, and Cohen (2004). Bernheim and Rangel (2004) discuss related material from the economic point of view.

The Quarterly Journal of Economics, May 2008

learning and the resulting choice behavior. Schultz, Dayan, and Montague (1997) noted that the shift forward in time of dopamine release in response to information about a future reward suggests that it plays a role in reinforcement learning (Bush and Mosteller 1951; Rescorla and Wagner 1972). A quantitative signal analogous to the level of dopamine release turns out to be needed in reinforcement learning algorithms to drive convergence toward a standard dynamic programming value function (Barto and Sutton 1982). Combining these insights, neuroscientists have hypothesized that the role of dopamine is to update the "value" that humans and animals attach to different actions and stimuli, which in turn affects the probability that such actions will be chosen. If true, this theory suggests that deeper understanding of dopamine will expand economists' understanding of how beliefs and preferences are formed, how they evolve, and how they play out in the act of choice. Even if the theory turns out to be in need of refinement, the findings to date are striking, given their interaction with the core concepts of economics. Deeper qualitative and quantitative understanding of dopamine release represents one of the most promising areas of neuroeconomics, a field that is itself one of the most vital areas of convergence between social and natural science (Glimcher 2003; Camerer, Loewenstein, and Prelec 2005).

In the current paper we begin the process of integrating the study of dopamine with economic understanding. One factor preventing such convergence is that experimental findings are consistent with alternative theories of dopaminergic function, such as the "incentive salience" theory of Berridge and Robinson (1998) and the "attention switching" theory of Redgrave and Gurney (2006). A second and related factor is that the current DRPE hypothesis is analogous to early economic theory, with an unobservable "reward" posited to mediate the relationship between dopamine, choice, and external stimuli. Samuelson (1938) pioneered the use of axiomatic techniques with his insistence that possible existence of a utility function be inferred only on the basis of observable choice data. Following his example, we formulate the DRPE hypothesis in axiomatic terms. Our treatment has precisely the "revealed preference" characteristic of identifying any possible reward function directly from the observables. This axiomatic approach solves the language barrier between economics and neuroscience by being explicit in its definition of the term "reward." It also guides the design of definitive neuroeconomic testing of the underlying model, which is currently going on in conjunction with the Center for Neural Science at NYU (Caplin et al. 2008).

We outline three economic applications of our model. First, we discuss the potential for measured dopamine release to provide insight into belief formation in repeated games, a topic of great interest in experimental economics (Nyarko and Schotter 2002). The second application relates directly to learning theory, in which Erev and Roth (1998) and Camerer and Ho (1999) pioneered experimental application of the reinforcement model of animal learning. Finally, we outline an application to addiction, strongly related to the work of Bernheim and Rangel (2004).[2]

II. Dopamine and Economic Decision Making

II.A. Neuroscientific Foundations

A connection between nerve cells (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. When a presynaptic neuron releases a neurotransmitter, that chemical 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 postsynaptic neuron "knows" that the presynaptic neuron has released a neurostransmitter, a form of information transfer. This may, in turn, lead to further communication with neurons to which the "newly informed" neuron is connected. Dopamine is one such neurotransmitter and the term dopamine (or dopaminergic) neuron refers to any neuron that uses dopamine as a neurotransmitter.

Although dopamine neurons exist in several different parts of the brain, the ones that we are interested in form two *dopaminergic pathways*, long thin bundles of cells that connect areas in the midbrain to areas in the forebrain. These pathways are evolutionarily very old and exist in all mammals, reptiles, and birds, as well as humans. The *mesostriatal* pathway links the

---

[2] The current research forms part of a broader agenda designed to establish rigorous methods for introducing nonstandard data into economic analysis. Caplin and Dean (2008) provide an entirely separate example focused around the question of how long a subject takes to make a decision and the light this sheds on search and the decision-making process. Caplin (2008) provides a general statement concerning the methodology being employed in these various studies.

substantia nigra in the midbrain to the *striatum* (*caudate nucleus* and *putamen*) in the forebrain. The *mesolimbic* pathway connects the *ventral tegmental area* in the midbrain to the to the *amygdala*, the *nucleus accumbens*, and the *medial prefrontal cortex*, again located in the forebrain. All the studies we discuss below measure dopamine activity within one of these brain structures. Although these pathways consist of many individual dopaminergic cells, there is evidence that activity within these groups of cells tends to be highly correlated (see Schultz [2002] for evidence that dopamine neurons all tend to behave in the same way, and Grace and Bunney [1983] for evidence on the electrical connection between dopaminergic neurons). This has led neuroscientists to hypothesize that dopamine neurons have a single, coherent, and important message to send to several different areas of the brain.

There are two broadly popular techniques for measuring activity in dopaminergic neurons. Single-cell "recordings" in animals measure nerve cell activity directly, using electrodes inserted into the animals' brains. This technique gives excellent spatial and temporal resolution, but is not feasible in humans for obvious reasons. Instead, human studies use functional magnetic resonance imaging (fMRI), which measures the amount of oxygenated blood present in a particular brain area at a given time, providing an indirect measure of brain activity in real time. This technique produces noisy data and has poor temporal resolution, but it is an improvement on previous techniques. All the observational studies described below use either single-unit recording or fMRI technology.

In addition to gathering information by observing how dopamine neurons respond to different stimuli, researchers can also use biological or pharmacological interventions to alter dopamine release. These include direct stimulation of dopamine neurons, lesions (destruction) of dopaminergic areas (in animals), genetic alteration to reduce or enhance dopamine production (again in animals), and drug-induced reduction or enhancement of dopamine activity (in both animals and humans). Many of the studies described below make use of one of these techniques.

II.B. From Reward to Information

A sequence of early experiments led neuroscientists to the conclusion that dopamine played a crucial role in behavior by mediating "reward." Essentially the idea was that dopamine converted experiences into a common scale of "reward" and that animals (and by extension, people) made choices in order to maximize this reward.

The seminal study in this literature was that of Olds and Milner (1954), who showed that rats would repeatedly return to locations in which they were given direct electrical stimulation of brain regions that included many dopamine neurons. Later studies showed that, when given the opportunity to self-administer such stimulation, rats would do so rather than feed, drink, or mate with attractive female rats (Gallistel, Shizgal, and Yeomans 1981). Moreover, they would choose to do so even if this stimulation was coupled with painful electrical shocks (see Gardner and David [1999] for a review). Further evidence relating dopamine to choice comes from Berridge (1999), who showed that when they are deprived of dopamine, rats stop responding to "rewarding" stimuli in their environment, such as food.

A second set of findings that seemingly support the hedonia theory relates activity levels of dopamine neurons to the receipt of "rewards." These rewards include food (Kiyatkin and Gratton 1994), liquid (Apicella et al. 1991), and access to sexual partners (Fiorino, Coury, and Phillips 1997). Studies using fMRI in humans have shown that dopamine areas also respond positively to the receipt of more abstract rewards, such as money (Montague and Berns 2002; Knutson and Peterson 2005). More recent work has shown that dopamine response correlates with stated subjective preferences over types of fruit juice (O'Doherty et al. 2006).

The simple hypothesis of "dopamine as reward" was spectacularly disproved by a sequence of experiments highlighting the role of *beliefs* in modulating dopamine activity: whether or not dopamine responds to a particular reward depends on whether or not this reward was *expected.* This result was first shown by Schultz, Apicella, and Ljungberg (1993) and Mirenowicz and Schultz (1994). 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 a small amount of time later. Initially (i.e., before the animal had learned to associate the tone with the juice), dopamine neurons fired in response to the *juice* but not the *tone.* However, once the monkey had learned that the tone predicted the arrival of juice, dopamine responded to the tone, but now did *not* respond 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 drop in the background level of dopamine activity when the juice was expected.

These findings made it clear that expectations play a key role in the response of dopaminergic neurons to stimuli. This finding has been supported by fMRI studies in humans (e.g., Montague and Berns [2002]).

The dramatic findings concerning the apparent role of information about rewards in mediating the release of dopamine led many neuroscientists to abandon the hedonic theory of dopamine in favor of the DRPE hypothesis: that dopamine responds to the difference between how "rewarding" an event is and how rewarding it was expected to be.[3] One reason that this theory has generated so much interest is that a reward prediction error of this type is a key algorithmic component of reinforcement models of learning: such a signal is used to update the value attached to different actions. This has led to the further hypothesis that dopamine forms part of a reinforcement learning system that drives behavior. Several studies have successfully correlated dopaminergic activity with the error signal from calibrated reinforcement learning models (e.g., Montague and Berns [2002]; Bayer and Glimcher [2005]; Bayer, Lau, and Glimcher [2007]; O'Doherty et al. [2003, 2004]; Daw et al. [2006]; and Li et al. [2007]).

That there is a link between dopamine and learning has been confirmed in humans in a recent study by Pessiglione et al. (2006), which builds on earlier work in animals by, for example, Wise (2004). Pessiglione et al. examined learning behavior as people were asked to choose between an option that stochastically provided a reward and one that never provided a reward. Participants in the study were divided into three groups, one of which was given a dopamine agonist (a drug that enhances dopamine levels), one a dopamine antagonist (a drug that diminishes dopamine levels), and one a placebo. The rate of learning (measured by the proportion of "correct" choices of the stochastically rewarding option) was highest for the group with enhanced dopamine, and lowest for the group with reduced dopamine.

In addition to the interaction with learning, dopamine has long been associated with addiction and addictive behavior. Evidence for this comes principally from the observation that many

---

[3] The above discussion makes it clear that reward is used in a somewhat unusual way. In fact, what dopamine is hypothesized to respond to is effectively unexpected changes in lifetime "reward": dopamine responds to the bell not because the bell itself is rewarding, but because it indicates increased probability of future reward. We will return to this issue in Section IV.

drugs of addiction, such as cocaine and amphetamines, act directly on the dopamine system. Even those drugs of addiction that do not affect the dopamine system directly seem to do so indirectly.

II.C. Dopamine and Economics

Validity of the DRPE hypothesis and the related hypothesis concerning the role of dopamine in learning would open the door to fundamental economic insights:

1. *Dopamine and the construction of utility*: According to the DRPE hypothesis, dopamine contains information on "reward," which in turn acts as an input into choice. As such, we can see dopamine as a key input into the construction of "utility," or a building block in determining the choices people make.
2. *Dopamine as a carrier of information on expectations*: The DRPE hypothesis states that dopamine responds to the difference between experienced and anticipated rewards. As such, it should be possible to use observation of dopamine to back out a measure of what was expected in different circumstances, allowing a new window into decision-making under uncertainty (see Section IV).
3. *Dopamine as a building block for learning*: The importance of learning theory in economics is apparent from the effort that has been dedicated to it both within macroeconomic theory (e.g., Marcet and Sargent [1989] and Evans and Honkapohia [2001]) and within microeconomic theory (e.g., Fudenberg and Levine [1998]). The DPRE hypothesis suggests that dopamine forms one of the building blocks of learning within the brain. Understanding the DRPE can therefore be thought of as the first step in developing a "neuroeconomic theory of learning." Again, we return to this theme in Section IV.

There are many intriguing economic questions that may be better understood as we drive forward with neuroeconomic research related to dopamine. If dopamine participates in the "construction of utility," we can begin to explore the mechanisms by which external stimuli affect choice. If dopamine can be linked to beliefs, then researchers have an additional window through which to explore the belief formation process. The relationship between dopamine and learning raises a particularly rich set of questions: When is the dopaminergically intermediated reinforcement learning system active? Is reinforcement a process that can take place at different speeds depending on internal and external cues? To what extent do the speed of learning and the persistence of past reinforcement provide an adequate behavioral model of habituation? Can we use neuroscientific techniques to understand when and where these various speeds and styles of learning are engaged? Can the relative influence on behavior of past and present experiences be impacted by specific neurological interventions? Can neuroeconomic research be designed to identify appropriate metrics on the state and action spaces to identify how far reinforcement learning is generalized to subjectively similar settings? How does reinforcement learning operate in the case of delayed rewards? What role do emotions such as anxiety and disappointment play in the process of reinforcement? Are particular choices concerning the future likely to be poorly made, due to the lack of direct reinforcement inputs on the path? Of course, before we can take wing and address these interesting questions, we need first to settle on some ground rules for the underlying neuroeconomic investigations. It is just such rules that our current paper aims to provide.

II.D. Why Neuroeconomics Needs Axioms

The goal in what follows is to specify in a simple, parsimonious, and nonparametric way the properties that the dopamine system must have in order to be characterized as encoding a reward prediction error. This is intended to enable us to design a definitive experimental test of the hypothesis. Our axiomatic approach is intended also to reduce barriers between economic and neuroscientific disciplines. In economics, reward concepts are inferred from observed choices, whereas neuroscientists interpret them in relation to intuitions concerning the flow of experience (e.g., a squirt of juice is assumed to be rewarding to a thirsty monkey). Samuelson (1938) pioneered the use of axiomatic techniques within economics in an effort to rid utility theory of unobservable constructs. The revealed preference program that resulted from this effort constitutes one of the major methodological achievements in social science. The idea is to define constructs such as utility directly in terms of what they imply for observables.

The axiomatic approach has the added virtue of pinning down the range of available transformations of these constructs that are equivalent in terms of the observables. In fact, our central result justifies only an ordinal version of the DRPE hypothesis.

Although anticipated and realized rewards are well defined in this characterization, the literal "reward prediction error," or arithmetic difference between experienced and predicted rewards, is not. The reward function in our DRPE representation is defined only up to strictly increasing monotone transformations. Hence our basic representation does not allow one to treat dopamine as an invariant measure of the reward difference or "error," just as the notion of marginal utility is ill-defined for standard ordinal utility functions. We provide additional assumptions that do animate this prediction error yet know of no neuroscientific research that validates these assumptions. This calls into question the most prevalent current methodology for testing the DRPE hypothesis, which is to correlate an estimated error signal from a calibrated reinforcement learning model with brain activity (see, for example, Montague and Berns [2002]; Haruno et al. [2004]; O'Doherty et al. [2003, 2004]; Bayer and Glimcher [2005]). Not only is there little theoretical support for the assumed functional forms, but also the calculated reward prediction error is very highly correlated with, and could therefore be similarly explained by, other relevant measures such as the magnitude of the reward itself and surprise. Although one might use statistical methods to discriminate among the various models, these will at best produce a ranking of the considered alternatives, rather than a global statement on model validity. Furthermore, these tests require parametric specifications of reward and learning. In contrast, our axiomatic approach characterizes the DRPE hypothesis in a parsimonious, nonparametric, and easily testable manner and guides the design of protocols that effectively differentiate among these various theories.

The broader point is that neuroeconomic research needs to follow the lead of utility theory and internalize the perils of treating an ordinal reward function as if it were cardinal. Just as the ordinal revolution in economics reshaped understanding of marginal utility, so the axiomatic approach to neuroeconomics may reshape the understanding of reward prediction errors. Current neuroscientific testing protocols are vulnerable to criticism on these grounds because cardinality plays an unjustified and largely unjustifiable role. Because of this, the axiomatic method is needed to guide design of definitive protocols for establishing the validity of the underlying model. Such definitive protocols are particularly important, as the DRPE hypothesis is not the only active theory of dopamine function. Though largely discredited due to the experiments of Schultz, the "dopamine as hedonia" theory still has its adherents. The "incentive salience" hypothesis of Berridge and Robinson (1998) holds that dopamine responds to how "wanted" a stimulus is, which is separate from how much a stimulus is "liked." Redgrave and Gurney (2006) suggest that dopamine plays a role in "switching attention" between different activities. Moreover, the qualitative fMRI studies of Zink et al. (2003), Delgado et al. (2005), and Knutson and Peterson (2005) suggest that dopamine responses may be modulated not only by a reward prediction error but also by less reward-specific "surprise."

III. The Model

III.A. The DRPE Hypothesis

We develop the DRPE model in the simplest environment in which the concept of a reward prediction error makes sense. The agent either is endowed with or chooses a lottery from which a prize is realized. We observe both any initial act of choice among lotteries, and the dopaminergic response when each possible prize z is realized from lottery p, as measured by the *dopamine release function* (DRF). Many of the subtleties of the theory that follow derive from the fact one cannot observe dopaminergic responses to prizes that are *ex ante* impossible.

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

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

The function $\delta(z, p)$ defined on $M = \{(z, p) | z \in Z, p \in \Lambda(z)\}$ identifies the dopamine release function, $\delta : M \to \mathbb{R}$.[4]

The DRPE hypothesis hinges on the existence of some definition of "predicted reward" for lotteries and "experienced reward" for prizes that captures all the information necessary

---

[4] We endow Λ with the metric

$$d_\Lambda(p, q) = \sum_{z \in \mathrm{Supp}(p) \cup \mathrm{Supp}(q)} (p(z) - q(z))^2$$

and $Z \times \Lambda$ with the product metric.

to determine dopamine output. In this paper, we make the basic rationality assumption that the expected reward of a degenerate lottery is equal to its experienced reward as a prize.[5] Hence the function $r : \Lambda \to \mathbb{R}$, which defines the expected reward associated with each lottery, simultaneously induces the reward function on prizes $z \in Z$ as $r(e_z)$. We define $r(Z)$ as the set of values taken by the function r across degenerate lotteries,

$$r(Z) = \{r(p) \in \mathbb{R} | p = e_z, z \in Z\}.$$

Our first basic requirement for the DRPE hypothesis is that there exist some reward function containing all information relevant to dopamine release. We say that the reward function fully summarizes the DRF if this is the case. Our second requirement is that the dopaminergic response be strictly *higher* for a more rewarding prize than for a less rewarding one. Furthermore, a given prize should lead to a *higher* dopamine response when obtained from a lottery with *lower* predicted reward. Our third and final requirement is that, if expectations are met, the dopaminergic response does not depend on what was expected. If one is told that they will get any prize for sure, and is then given that prize, there is no “reward prediction error,” as one has just received what was expected. We refer to this property as “no-surprise constancy.” These requirements are formalized in the following definition.[6]

DEFINITION 2. A DRF $\delta : M \to \mathbb{R}$ admits a *dopaminergic reward prediction error* (DRPE) representation if there exist a reward function $r : \Lambda \to \mathbb{R}$ and aggregator function $E : r(Z) \times r(\Lambda) \to \mathbb{R}$ that

1. *Represent the DRF*: given $(z, p) \in M$,

$$\delta(z, p) = E(r(e_z), r(p)).$$

2. *Respect dopaminergic dominance*: E is strictly increasing in its first argument and strictly decreasing in its second argument.

---

[5] Dean (2007) allows the reward function to differentiate between realized prizes and the lotteries that yield them with certainty.

[6] Note that although our axioms are stated in terms of dopaminergic measurements, precisely the same axiom system would characterize any neuroscientific measure of a reward prediction error, whether or not solely dopaminergic in nature. Note moreover that one possible refinement of the theory involves dopamine itself carrying only positive reward surprises, as has been suggested by Bayer and Glimcher (2005). See Dean (2007) for an axiomatization of this case.

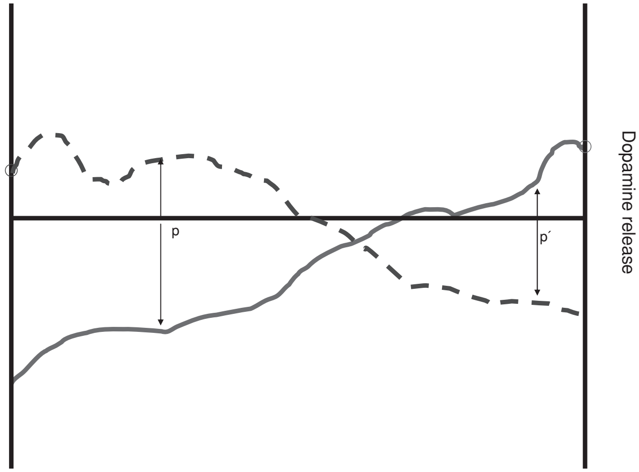

FIGURE I
A Violation of A1

Note. 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.

3. *Satisfy no-surprise constancy*: given $x, y \in r(Z)$,

$$E(x, x) = E(y, y).$$

III.B. The Basic Result

We introduce three critical axioms for $\delta : M \to \mathbb{R}$ to admit a DRPE and illustrate them graphically in Figures I–III for the two-prize case, in which 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 dashed line indicates the amount of dopamine released when prize 1 is obtained from each of these lotteries (i.e., $\delta(z_1, p)$), whereas the solid line represents the amount of dopamine released when prize 2 is obtained from each lottery (i.e., $\delta(z_2, p)$). Note that there are no observations at $\delta(z_1, 0)$ and $\delta(z_2, 1)$, as prize 1 is not in the support of the former, while prize 2 is not in the support of the latter.

Our first axiom demands that the order on the prize space induced by the DRF be independent of the lottery that the prizes are obtained from. In terms of the graph in Figure I, if dopaminergic

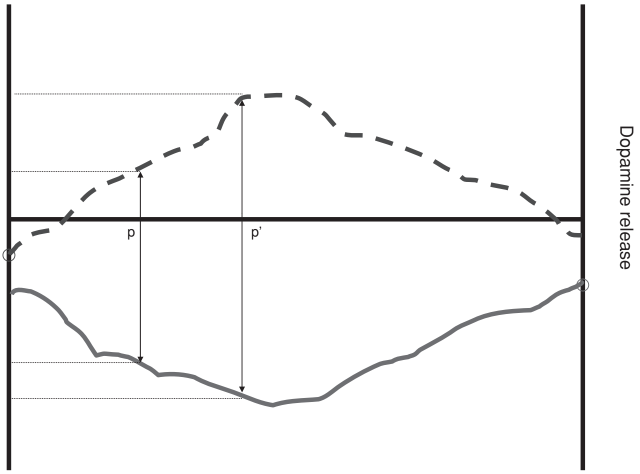

FIGURE II
A Violation of A2

Note. Looking at prize 1, more dopamine is released when this prize is obtained from p' than when it is obtained from p, suggesting that p has a higher predicted reward than p'. The reverse is true for prize 2, making a DRPE representation impossible.

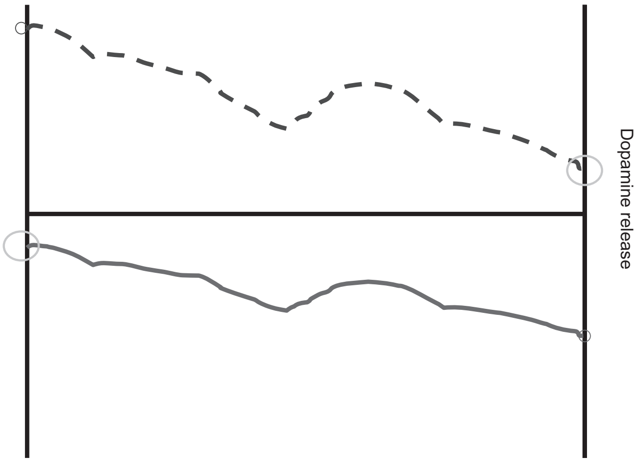

FIGURE III
A Violation of A3

Note. The dopamine released when prize 1 is obtained from its sure thing lottery is higher that that when prize 2 is obtained from its sure thing lottery.

release based on lottery p suggests that prize 1 has a higher experienced reward than prize 2, there should be no lottery p' to which dopaminergic release suggests that prize 2 has a higher experienced reward than prize 1. Figure I shows a violation of such *coherent prize dominance*. It is intuitive that all such violations must be ruled out for a DRPE to be admitted. Our second axiom ensures that the ordering of lotteries by dopamine release is independent of the obtained prize. Figure II shows a case that contradicts this, in which more dopamine is released 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. Our final axiom deals directly with equivalence among situations in which there is no surprise, a violation of which is recorded in Figure III, in which more dopamine is released when prize 2 is obtained from its degenerate lottery than when prize 1 is obtained from its degenerate lottery.

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

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

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

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

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

$$\delta(z', e_{z'}) = \delta(z, e_z).$$

That A1–A3 are necessary for a DRPE representation is demonstrated in Proposition 1, proven along with all other results in the Appendix. Yet there are examples that satisfy A1–A3 in which there is no DRPE representation. Theorem 1 establishes that all such examples derive from the fact that the domain of the dopamine function is not a Cartesian product of the prize and lottery space: distinct prizes $z, z' \in Z$ are associated with distinct sets $\Lambda(z) \neq \Lambda(z') \subset \Lambda$.

PROPOSITION 1. A DRF $\delta : M \to \mathbb{R}$ must satisfy A1–A3 in order to admit a DRPE representation.

THEOREM 1. A DRF $\delta : M \to \mathbb{R}$ permits a DRPE representation if and only if there exists a function $\bar{\delta} : Z \times \Lambda \to \mathbb{R}$ that (a) is an extension of $\delta : M \to \mathbb{R}$ and (b) satisfies A1*–A3* on this extended domain:

A1*: Given $z, z' \in Z$ and $p, p' \in \Lambda$, $\bar{\delta}(z, p) > \bar{\delta}(z', p) \Rightarrow \bar{\delta}(z, p') > \bar{\delta}(z', p')$;

A2*: Given $z, z' \in Z$ and $p, p' \in \Lambda$, $\bar{\delta}(z, p) > \bar{\delta}(z, p') \Rightarrow \bar{\delta}(z', p) > \bar{\delta}(z', p')$;

A3*: Given $z, z' \in Z$, $\bar{\delta}(z', e_{z'}) = \bar{\delta}(z, e_z)$

According to Theorem 1, existence of a DRPE representation is equivalent to the function $\delta : M \to \mathbb{R}$ being extendible to the domain $Z \times \Lambda$ in a manner that retains A1–A3. Lemma 1 in the Appendix implies that $\Lambda \subset \mathrm{cl}(\Lambda(z))$ for all $z \in Z$, suggesting that suitable continuity conditions will allow us to extend δ appropriately to the domain $Z \times \Lambda$. In addition to continuity axiom A4, an additional "separation" axiom is required for technical reasons. Theorem 2 establishes that A1–A5 allow us to construct a unique, continuous extension of δ to $Z \times \Lambda$ satisfying A1*–A3*, which we denote as $\bar{\delta}^C$ (A4 and A5 will be assumed in what follows, so that we can call upon this function $\bar{\delta}^C$ in stating axioms and deriving proofs). The central existence result of the paper is obtained by combining Theorems 1 and 2.

AXIOM 4 (A4: Uniform Continuity). The function $\delta : M \to \mathbb{R}$ is uniformly continuous.

AXIOM 5 (A5: Separation). Given $(z, p), (z', p) \in M$ with $\delta(z, p) \neq \delta(z', p)$,

$$\inf_{\{p' \in \Lambda \mid (z, p'), (z', p') \in M\}} |\delta(z, p') - \delta(z', p')| > 0.$$

THEOREM 2. Given that the DRF $\delta : M \to \mathbb{R}$ satisfies A1–A5, there exists a unique uniformly continuous function $\bar{\delta}^C : Z \times \Lambda \to \mathbb{R}$ that is an extension of δ and satisfies A1*–A3*.

THEOREM 3. A DRF $\delta : M \to \mathbb{R}$ that satisfies A4 and A5 admits a DRPE representation if and only if it satisfies A1–A3.

Although A1–A3 form crucial underpinnings for the DRPE hypothesis, they appear inconsistent with alternative hypotheses relating to salience and to experienced reward. Consider two prizes, z and z', and two lotteries, p, which give a 1% chance of winning z and a 99% chance of winning z', and p', which reverses these two probabilities. It is intuitive that receiving z from p would be a very "salient" or surprising event, whereas receiving z' would be very unsurprising. Thus a system responding to salience should give higher readings when z is obtained from p than when

z' is obtained from p. However, this situation is reversed when the two prizes are obtained from p'. Thus we would expect A1 to fail if dopamine responded to salience. A similar argument shows that A2 would also fail, while A3 would hold, as the salience of getting a prize from a sure thing lottery should be the same in all cases. With regard to the older and somewhat discredited theory that dopamine responds only to "experienced reward," this would lead A3 to be violated—different prizes with different reward values would give rise to different dopaminergic responses, even when received from degenerate lotteries.

III.C. Dopamine and Choice

Although the approach we are taking is intellectually aligned with the revealed preference approach to choice theory, it will be of little interest to economists unless the reward function is somehow related to choice behavior. One such relation would be if choices among lotteries could be modeled as deriving from maximization of the DRPE reward function. It is this simplest of cases that we characterize below. While this case is of obvious interest to economists, it represents an extreme form of the DRPE hypothesis. A more standard scenario involves dopamine as simply one component of a richer overall process of learning and of choice.

DEFINITION 3. The choice correspondence C is defined on Q, the set of all nonempty compact subsets of Λ, with $C(X) \subseteq X \in Q$ denoting the nonempty set of lotteries chosen from X. A DRF $\delta : M \to \mathbb{R}$ and a choice correspondence C admit a *choice-consistent* DRPE representation if there exist $r : \Lambda \to \mathbb{R}$, $E : r(Z) \times r(\Lambda) \longrightarrow \mathbb{R}$ that form a DRPE of $\delta : M \to \mathbb{R}$, and for all $X \in Q$,

$$C(X) = \arg\max_{p \in X} r(p).$$

To make the connection between dopamine and the standard theory of utility maximization requires us first to ensure that this theory applies, by invoking the weak axiom of revealed preference. We then ensure that what the axiom says is chosen in any given pair accords with the dopaminergic responses.

AXIOM 6 (A6: WARP). The choice correspondence C satisfies the weak axiom of revealed preference.

AXIOM 7 (A7: Dopamine and Choice Consistency). Given $p, q \in \Lambda$ and $z \in Z$,

$$\bar{\delta}^C(z, p) \leq \bar{\delta}^C(z, q) \Longleftrightarrow p \in C(\{p, q\}).$$

THEOREM 4. A DRF $\delta : M \to \mathbb{R}$ that satisfies A4 and A5 admits a choice-consistent DRPE representation, with $r : \Lambda \longrightarrow \mathbb{R}$ continuous, if and only if it satisfies axioms A1–A3, A6, and A7.

III.D. The Reward Error

Lemma 2 in the Appendix shows that if $r : \Lambda \to \mathbb{R}$ forms part of a DRPE representation of a DRF $\delta : M \to \mathbb{R}$, then so does any function $r^* : \Lambda \to \mathbb{R}$ that is a strictly increasing monotone transform of r. We develop a more restrictive additive formulation that represents the minimum requirement for using dopaminergic response to animate the notion of reward difference/error.

DEFINITION 4. A DRF $\delta : M \to \mathbb{R}$ admits a *dopaminergic additive reward prediction error* (DARPE) representation if there exist a function $r : \Lambda \to \mathbb{R}$ and a strictly increasing function $G : r(Z) - r(\Lambda) \to \mathbb{R}$ such that, given $(z, p) \in M$,

$$\delta(z, p) = G(r(e_z) - r(p)).$$

In the DARPE representation, the term "error" is interpreted additively—the reward prediction error is literally the arithmetic difference between experienced and predicted reward. It is this version of the theory that (implicitly) forms the basis of many of the models currently in use within neuroscience. The conditions required for the DRPE are equivalent to those required for an additive representation of the ordering generated by the DRF $\delta : M \to \mathbb{R}$ on M.

PROPOSITION 2. The DRF $\delta : M \to \mathbb{R}$ admits a DARPE representation if it satisfies A3 and there exist functions $u : Z \to \mathbb{R}$, $v : \Lambda \to \mathbb{R}$ such that, given $z, z' \in Z$ and $p, p' \in \Lambda$,

$$\bar{\delta}(z, p) \geq \bar{\delta}(z', p') \text{ if and only if } u(z) + v(p) \geq u(z') + v(p').$$

The conditions under which a binary relation can be represented additively have been extensively studied both in economics and in mathematical psychology. Necessary and sufficient condition for the case of countable domains have been provided by Tversky (1964), Scott (1964), Adams (1965), and Fishburn (1970) and for uncountable domains by Debreu (1960), Luce and Tukey (1964), Jaffray (1974), Wakker (1989), and Gonzales (2000). We make a technical assumption to rule out trivial problems and to ensure that we can establish a DARPE representation based on the existing conditions for additivity due to Wakker (1989).

AXIOM 8 (A8: No Triviality). There exist $z, z' \in Z$ and $p, p' \in \Lambda$ such that $\bar{\delta}^C(z, p) \neq \bar{\delta}^C(z', p)$ and $\bar{\delta}^C(z, p) \neq \bar{\delta}^C(z, p')$.

Although our technical approach to additivity derives from a prior literature, the supporting axioms in that literature, such as the "hexagon condition," have little intuitive meaning. We provide a characterization with a more intuitive interpretation. Suppose that lotteries p and p' display dopaminergic responses to some prize z that are precisely equivalent to those that q and q' display with respect to some prize w. If reward differences are to be well defined, we would like to conclude from this that "the reward difference between lottery p and p' is equivalent to that between q and q'." Hence if there are any two other prizes z' and w' such that the dopaminergic response to realization of z' from p' is equivalent to that of w' to q', then it must equally be the case that the dopaminergic response to realization of z' from p is equivalent to that of w' to q. Analogously, if prizes z and z' display dopaminergic responses from lottery p' that are precisely equivalent to those that w and w' display from lottery q', we would like to conclude that "the reward difference between lottery p and p' is equivalent to that between q and q'." Hence, if there are any two other lotteries p and q such that the dopaminergic response to realization of z from p is equivalent to that of w to q, it must be the case that the dopaminergic response to realization of z from p' is equivalent to that of w to q'. Despite the apparent conceptual difference between the reward difference statements above as applied to lotteries and to prizes, a careful reading shows them to reduce to the following common assumption.

AXIOM 9 (A9: Equivalent Differences). The extended DRF $\bar{\delta}^C : Z \times \Lambda \to \mathbb{R}$ is such that, given $z, z', w, w' \in Z$ and $p, p', q, q' \in \Lambda$ with $\bar{\delta}^C(z, p) = \bar{\delta}^C(w, q)$, $\bar{\delta}^C(z', p') = \bar{\delta}^C(w', q')$ and $\bar{\delta}^C(z', p) = \bar{\delta}^C(w', q)$, it follows that

$$\bar{\delta}^C(z, p') = \bar{\delta}^C(w, q').$$

THEOREM 5. Let Z be a connected metric space and assume that the DRF $\delta : M \to \mathbb{R}$ satisfies A4 and A5. Then the DRF satisfies A1–A3, A8, and A9 if and only if $\exists$ functions $r : \Lambda \to \mathbb{R}$ and $G : r(Z) - r(\Lambda) \longrightarrow \mathbb{R}$, with r nonconstant and continuous on both Z and Λ and with G continuous and strictly increasing, that form a DARPE representation of the DRF.

Whereas for the DRPE representation we could say that any strictly monotone transformation of a given reward function $r : \Lambda \to \mathbb{R}$ would also form part of a DRPE, this is not the case for the DARPE representation. Here we can only guarantee that any positive affine transformation will preserve the DARPE structure.

PROPOSITION 3. If r, G form a DARPE representation of $\delta : M \to \mathbb{R}$, and $r^* : \Lambda \to \mathbb{R}$ is a positive affine transformation of r, then there exists $G^* : r^*(Z) - r^*(\Lambda) \to \mathbb{R}$ such that r^*, G^* define also a DARPE representation. Moreover, if Z is connected and A4, A5, and A8 are satisfied, the class of r functions that additively represent the extended function $\bar{\delta}^C$ is unique up to a positive affine transformation.

III.E. *Expected Reward Prediction Error*

An extreme strengthening of the DRPE hypothesis is to demand that the predicted reward be computable as the mathematical expectation of the rewards associated with these prizes. Given the strong analogy with expected utility theory, we refer to this as an expected reward representation. Note again that this is a specialization of the DRPE representation. However, it neither implies nor is implied by the DARPE representation.

DEFINITION 5. A DRF $\delta : M \to \mathbb{R}$ admits a *dopaminergic expected reward prediction error* (DERPE) representation if there exist $r : \Lambda \to \mathbb{R}$, $E : r(Z) \times r(\Lambda) \to \mathbb{R}$ that form a DRPE representation of the DRF in which

$$r(p) \equiv \mu_p[u] \text{ all } p \in \Lambda,$$

for some function $u : Z \to \mathbb{R}$, where $\mu_p[u]$ denotes the expected value of $u : Z \to \mathbb{R}$ with respect to the lottery p.

As might be expected, given the close relationship between the DEPRE representation and standard expected utility, the condition we require is very similar to the independence axiom from choice theory.

AXIOM 10 (A10 : Independence). Given $z \in Z$, $p, p' \in \Lambda(z)$, $q \in \Lambda$ and $\lambda \in (0, 1]$,

$$\begin{aligned} &\bar{\delta}^C(z, p) > \bar{\delta}^C(z, p') \\ &\qquad \Rightarrow \bar{\delta}^C(z, \lambda p + (1-\lambda)q) > \bar{\delta}^C(z, \lambda p' + (1-\lambda)q). \end{aligned}$$

THEOREM 6. A DRF $\delta : M \to \mathbb{R}$ that satisfies A4 and A5 admits a DERPE representation if and only if it satisfies axioms A1–A3 and A10.

IV. THREE APPLICATIONS

In this section we outline three possible economic applications of the DRPE built upon our axiomatic foundations. The first is a direct application of the information inherent in the DRPE, whereas the second and third represent more ambitious agendas that involve building upon the current foundations.

IV.A. Belief Elicitation

Belief formation lies at the center of many game-theoretic solution concepts. Furthermore, many behavioral models of game play also place large explanatory burdens on beliefs that subjects may construct through introspection, experience, or both (Stahl and Wilson 1995; Cheung and Friedman 1997; Fudenberg and Levine 1998). One method for understanding beliefs is to infer them based solely on observed actions of players and an appropriate structural econometric model of the updating processes and decisions (e.g., Cheung and Friedman [1997]). Yet Nyarko and Schotter (2002) showed that they can explain play in various games far better using beliefs estimated from an incentive-compatible mechanism that directly elicit subjects' beliefs about partner play during the course of game.

Although this represents a powerful new form of evidence on beliefs, stated beliefs have their own potentially serious drawback. In the context of a repeated asymmetric matching pennies game, Rutström and Wilcox (2006) provide an example in which model-estimated beliefs predict game play better than stated beliefs, contrary to the results of Nyarko and Schotter (2002). They provide evidence suggesting that one of the reasons for this is that the act of belief elicitation itself can alter paths of play. Belief elicitation procedures require nontrivial instruction of subjects and interrupt the typical flow of subject attention and game play in a potentially significant way. Belief elicitation procedures could move such subjects toward belief-based thinking and play, away from naturalistic play of the form suggested by such belief-free models as that of reinforcement learning (Erev and Roth 1998; Sarin and Vahid 2001). Rutström and Wilcox conjecture that belief elicitation moves subjects away from relatively automatic emotional or "affective" predispositions that favor particular choices to more conscious ones where beliefs are estimated. They argue that cognitive processes that construct and update effective beliefs, as well as processes that combine them with payoff or value information to determine choice probabilities, may not be wholly conscious ones. For many subjects, the normal conscious product of any unconscious processing may be mostly an inclination toward a particular choice, rather than belief formation per se.

Given the continuing difficulties in evaluating subjective beliefs, the dopaminergic measurement techniques implicit in the DRPE hypothesis are potentially very useful tools. In particular, a subjectivist interpretation of our framework would offer researchers access to a method of belief elicitation that would not require deliberative interruption. To give a simple example, in a world with only one good and one bad prize, if one can validate the DERPE representation, the dopaminergic response will be higher for one lottery than for another if and only if the subjective belief in receiving the better prize is stronger. This new form of evidence on beliefs not only would help in understanding play in games, but also would provide vital insights into learning behavior.

IV.B. Toward a Neuroeconomic Theory of Learning

One of the reasons the DRPE hypothesis has gained so much attention is its link to computational models of learning. Schultz, Dayan, and Montague (1997) noted that a "prediction error" signal is precisely what is needed in reinforcement learning algorithms designed by computer scientists to approximate standard dynamic programming value functions (Barto and Sutton 1982). This has led many researchers to conclude that this similarity is no coincidence and that dopamine does indeed measure a reward prediction error that is used to update an evolving value function.

Within computer science, reinforcement learning refers to a class of algorithms designed to solve a specific type of problem, in which an agent tries to choose actions in order to maximize a sequence of rewards discounted at rate β. Let $\omega \in \Omega$ be a set of states, $A : \Omega \Rightarrow \mathcal{A}$ a set of actions available in each state, and $h : \mathcal{A} \times \Omega \to \mathbb{R}$ the instantaneous reward of taking a particular action in a particular state. States change in a possibly stochastic way, depending both on past states and on past actions. One example of a reinforcement learning algorithm is the "Q-learning algorithm," which recursively calculates the value of taking a particular action in a particular state: $\bar{Q}^t(a, \omega)$ is the estimated value associated in period t with taking action $a \in A(\omega)$ in state $\omega \in \Omega$. These value estimates are updated using the following algorithm:

$$\begin{aligned} \bar{Q}^t(a_{t-1}, \omega_{t-1}) &= \bar{Q}^{t-1}(a_{t-1}, \omega_{t-1}) + \alpha \Delta(a_{t-1}, \omega_{t-1}, \omega_t) \\ \Delta(a_{t-1}, \omega_{t-1}, \omega_t) &= \left[h(a_{t-1}, \omega_{t-1}) + \beta \max_{a \in A(\omega_t)} \bar{Q}^{t-1}(a, \omega_t) \right] \\ &\quad - \bar{Q}^{t-1}(a_{t-1}, \omega_{t-1}). \end{aligned} \tag{1}$$

Estimates of the value of all other action/state pairs remains unchanged. The estimate $\bar{Q}^{t-1}(a_{t-1}, \omega_{t-1})$ is updated by adjusting it according to $\Delta(a_{t-1}, \omega_{t-1}, \omega_t)$, the difference between $\bar{Q}^{t-1}(a_{t-1}, \omega_{t-1})$ (how good action a_{t-1} was expected to be in state ω_{t-1}) and $[h(a_{t-1}, \omega_{t-1}) + \beta \max_{a \in A(\omega_t)} \bar{Q}^{t-1}(a, \omega_t)]$ (how good it actually turned out to be, according to current estimates). Such algorithms can, in the right circumstances, well approximate the optimal solution to this class of problems (see Sutton and Barto [1998]). Schultz, Dayan, and Montague (1997) suggested that dopaminergic firing rates encode $\Delta(a_{t-1}, \omega_{t-1}, \omega_t)$, which is then used by other parts of the brain to update value functions associated with different states and actions. As we have discussed, some support for this identification has been found by the work of Bayer and Glimcher (2005), as well as the correlation studies of, for example, Daw et al. (2006).

Important as it is to deepen understanding of reinforcement learning, it is clear that realistic learning behaviors are multifaceted. Camerer and Ho (1999) capture deviations from reinforcement learning relating to payoffs on hypothetical acts that were not taken. Selten and Stoecker (1986) and Bendor, Mookherjee, and Ray (2001) enrich the basic model by incorporating aspirations. Shor (2004) finds experimentally that when payoffs are determined by a hidden state that changes over time, agents adapt rapidly to the unsignaled change in the environment. Charness and Levine (2003) also show that changes in the environment impact whether or not experimental subjects exhibit reinforcement learning. Neuroeconomic research following up on the DRPE hypothesis has the potential for providing economists with much relevant information. Indeed Balleine, Killcross, and Dickinson

(2003) and Balleine (2005) detail a series of provocative studies that are consistent with the idea that there may be multiple modes of learning. Balleine (2005) is currently investigating the development of "habits" as opposed to more flexible responses, and an associated "supervisory function" that determines the extent to which habitual behavior is called into play in any given situation. Ideally, development of dynamic versions of the DRPE hypothesis will be a prelude to a more complete neuroeconomic theory of learning, possibly building on such multiple process ideas.

IV.C. Application 3: Addiction

As pointed out by Bernheim and Rangel (2004), substance addiction is a massive concern in the United States. They highlight a growing consensus in neuroscience and psychology that dopaminergically intermediated learning effects, rather than "hedonic" effects, provide the key to understanding addictive behavior (see Redish [2004]). The key observation underlying this theory is that many addictive substances share an ability to activate the firing of dopamine with much greater intensity and persistence than other substances (e.g., Nestler and Malenka [2004]). The result may be a strong impulse to seek and use the substance, particularly in similar environments that activate latent reward expectations. To understand the role that such dopaminergic stimulation may play, consider again the learning algorithm described by equation (1). The idea is that, by stimulating dopaminergic activity in an unconditional fashion, addictive substances ensure that a high reward prediction error Δ is associated with any environmental cues that are associated with the corresponding activity. This characterizes cues as particularly important, and it is for this reason that high cue-triggered recidivism is seen as a defining feature of addiction.

Another critical aspect of addition that Bernheim and Rangel highlight is that many addicts expend great resources on failed efforts to break their habit. Understanding the neurological mechanisms that underlie addiction may allow interventions to be structured deliberately to reduce cue-induced cravings that frustrate many such efforts. The DRPE-based approach to addiction suggests in particular that well-designed dopaminergic manipulations may be of value to those seeking to change behavior. Although neuroscientists are currently taking the lead in exploring the interaction between dopamine and addiction, we believe that the interaction with economic reasoning is essential given that the ultimate goal is to impact choice. An integrative theory such as ours is a necessary prelude to the required form of interdisciplinary research.

V. CONCLUDING REMARKS

The aim of this paper is to provide an axiomatic representation of the DRPE hypothesis that will both provide parsimonious, nonparametric tests of the model and clear up the language barrier that can plague neuroeconomic research. The obvious next question is whether or not the data support such a model. Unfortunately, existing experiments provide only a rough guide to this question: the data they produce tend not to be of the simple form that our model requires. As we have discussed above, most tests of the DRPE hypothesis take place in environments in which subjects are constantly learning the values of different options. Our approach explicitly abstracts from a learning environment, in order to provide a simple and clean characterization. Thus data from learning environments are not readily interpretable within our framework. Some existing experiments, such as Mirenowicz and Schultz (1994), can to some extent be interpreted within our framework. While these data do not contradict our axioms, they are also not rich enough to provide a thorough test of the existence of a consistent ordering over prizes and beliefs by the dopamine function.

In order to address these shortcomings, we are currently working with the Center for Neural Science at New York University to complete experiments that will allow us to explicitly observe the function δ. These experiments involve taking readings of brain activity using fMRI technology as subjects receive different prizes from different lotteries. By designing our experiments explicitly to test our axioms, we hope to provide a rigorous and complete test of the DRPE hypothesis. This current work is intended to provide firm foundations upon which future experimental and theoretical neuroeconomic research can build.

APPENDIX: PROOFS AND ADDITIONAL RESULTS

A. Additional Results

LEMMA 1. For all $p \in \Lambda$ and finite $X \in 2^Z$, p is a limit point of $\cap_{z \in X} \Lambda(z)$. Furthermore, $Z \times \Lambda \subset \mathrm{cl}(M)$.

Proof. For any $X \subset Z$ such that $|X| < \infty$, every point in Λ is a limit point of $\cap_{z \in X} \Lambda(z)$. To see this, take the sequence $p_n = (1 - \frac{1}{n})p + \frac{1}{n}\bar{p}$, where $\bar{p}$ is the uniform lottery on $X \cup \text{Supp}(p)$, and note that, as $p_n(z) > 0 \ \forall \ z \in X$, it must be the case that $p_n \in \cap_{z \in X} \Lambda(z)$ for all n. As $\lim_{n \to \infty} d_{\Delta(Z)}(p, p_n) = 0$, p is the limit point of $\{p_n\}_{n=1}^{\infty}$. To show that every point of $Z \times \Lambda$ is a limit point of M and thus lies in the closure of M, consider $(z, p) \in Z \times \Lambda$ and construct a sequence $(z, p_n) \in \Lambda(z)$ with $p_n \in \Lambda(z)$ and $\lim p_n = p$, so that by construction $\lim_{n \to \infty}(z_n, p_n) = (z, p)$, completing the proof. ■

LEMMA 2. Let $r : \Lambda \to \mathbb{R}$ and $E : Z \times \Lambda \to \mathbb{R}$ form a DRPE representation of a DRF $\delta : M \to \mathbb{R}$. Then for any function $\bar{r} : \Lambda \to \mathbb{R}$ which is a strictly increasing monotone transform of r, there exists a function $\bar{E} : \bar{r}(Z) \times \bar{r}(\Lambda)$ such that $\bar{r}$, $\bar{E}$ form a DRPE.

Proof. Given that $\bar{r}$ is a strictly positive transform of r, the proof of Theorem 1 tells us that, given $z, z' \in Z$ and $p, p' \in \Lambda$,

$$\bar{r}(e_z) \geq \bar{r}(e_{z'}) \Longleftrightarrow r(e_z) \geq r(e_{z'}) \Longleftrightarrow \bar{\delta}(z, p) \geq \bar{\delta}(z', p); \text{ and}$$
$$\bar{r}(p) \geq \bar{r}(p') \Longleftrightarrow r(p) \geq r(p') \Longleftrightarrow \bar{\delta}(z, p) \leq \bar{\delta}(z, p').$$

Thus the same proof tells us that $\bar{r}$ together with $\bar{E} : \bar{r}(Z) \times \bar{r}(\Lambda)$, defined by

$$a \in \bar{E}(x, y) \text{ iff } \exists \ z \in Z, p \in \Lambda \text{ with } \bar{r}(e_z) = x, \bar{r}(p) = y, \bar{\delta}(z, p) = a,$$

forms a DRPE representation. ■

B. Proofs—Propositions

Proof of Proposition 1. The proposition states that

A DRF $\delta : M \to \mathbb{R}$ must satisfy A1–A3 in order to admit a DRPE representation.

Consider any DRF that admits a DRPE representation, and let r, E correspondingly represent the DRF. Now consider $(z, p), (z', p'), (z', p), (z, p') \in M$. To establish that the DRF satisfies A1, note that, since r and E form a DRPE,

$$\begin{aligned} \delta(z, p) &> \delta(z', p) \\ &\Rightarrow E(r(e_z), r(p)) > E(r(e_{z'}), r(p)) \\ &\Rightarrow r(e_z) > r(e_{z'}) \\ &\Rightarrow E(r(e_z), r(p')) > E(r(e_{z'}), r(p')) \\ &\Rightarrow \delta(z, p') > \delta(z', p'), \end{aligned}$$

where the first and fourth implication follow from the fact that r fully summarizes δ, and the second and third implications follow from the fact that r and E respect dopaminergic dominance.

An analogous argument shows us that A2 holds:

$$\begin{aligned}\delta(z, p) &> \delta(z, p') \\ &\Rightarrow E(r(e_z), r(p)) > E(r(e_z), r(p')) \\ &\Rightarrow r(p) < r(p') \\ &\Rightarrow E(r(e_{z'}), r(p)) > E(r(e_{z'}), r(p')) \\ &\Rightarrow \delta(z', p) > \delta(z', p').\end{aligned}$$

Finally, we combine the fact that E and r represent the DRF with the fact that Definition 4, no-surprise constancy, is satisfied in any DRPE representation, to conclude that given $z, z' \in Z$,

$$\delta(z, e_z) = E(r(e_z), r(e_z)) = E(r(e_{z'}), r(e_{z'})) = \delta(z', e_{z'}).$$

Proof of Proposition 2. The proposition states that

The DRF $\delta : M \rightarrow R$ admits a DARPE representation if it satisfies A3 and there exist functions $u : Z \rightarrow \mathbb{R}$, $v : \Lambda \rightarrow \mathbb{R}$ such that, given $z, z' \in Z$ and $p, p' \in \Lambda$,

$$\bar{\delta}(z, p) \geq \bar{\delta}(z', p') \text{ if and only if } u(z) + v(p) \geq u(z') + v(p').$$

Suppose that there exist functions $u : Z \rightarrow \mathbb{R}$, $v : \Lambda \rightarrow \mathbb{R}$ that satisfy this condition and that A3 is satisfied. We conclude immediately that given $z, z' \in Z$,

$$\begin{aligned}\bar{\delta}(z, e_z) &= \bar{\delta}(z', e_{z'}) \\ &\Rightarrow u(z) + v(e_z) = u(z') + v(e_{z'}) \equiv K.\end{aligned}$$

Hence, given $z, z' \in Z$ and $p, p' \in \Lambda$,

$$\begin{aligned}\bar{\delta}(z, p) &\geq \bar{\delta}(z', p') \\ &\Leftrightarrow u(z) + v(p) \geq u(z') + v(p') \\ &\Leftrightarrow K - v(e_z) + v(p) \geq K - v(e_{z'}) + v(z') \\ &\Leftrightarrow v(p) - v(e_z) \geq v(z') - v(e_{z'}).\end{aligned}$$

This immediately implies the existence of a DARPE representation, defining $r : \Lambda \rightarrow \mathbb{R}$ by $r(p) = -v(p)$, and specifying $G(x)$ on $x \in r(Z) - r(\Lambda)$ by identifying $z \in Z$ and $p \in \Lambda$ with $r(e_z) - r(p) = x$ and setting

$$G(x) = \bar{\delta}(z, p).$$

The fact that this specification $G(x)$ is unambiguous is immediate in light of the definition of $r(p)$ in relation to the additive representation of $\bar{\delta}$.

Proof of Proposition 3. The proposition states that

If r, G form a DARPE representation of $\delta : M \to \mathbb{R}$, and $r^* : \Lambda \to \mathbb{R}$ is a positive affine transformation of r, then there exists $G^* : r^*(Z) - r^*(\Lambda) \to \mathbb{R}$ such that r^*, G^* also define a DARPE representation. Moreover, if Z is connected and A4, A5, and A8 are satisfied, the class of r functions that additively represent the unique continuous extension $\bar{\delta}^C$ is unique up to a positive affine transformation.

Because r^* is a positive affine transformation of r, we can write $r^*(p) = \alpha r(p) + \beta$ with $\alpha, \beta \in \mathbb{R}$ and $\alpha > 0$. Hence, given $z, z' \in Z$ and $p, p' \in \Lambda$,

$$\begin{aligned} r^*(e_z) - r^*(p) \quad &\geq \quad r^*(e_z) - r^*(p') \\ &\Longleftrightarrow \alpha(r(e_z) - r(p)) \geq \alpha(r(e_z) - r(p')) \\ &\Longleftrightarrow \bar{\delta}(z, p) \geq \bar{\delta}(z', p'). \end{aligned}$$

As the proof of Proposition 2 shows, we can thus find a function $G^* : r^*(Z) - r^*(\Lambda) \longrightarrow \mathbb{R}$ such that r^*, G^* form a DRPE representation of $\delta : M \to \mathbb{R}$.

Finally we need to show that, if A4, A5, and A8 are satisfied, and if $r : \Lambda \to \mathbb{R}$ and $s : \Lambda \to \mathbb{R}$ are such that, for $\forall\, z, z' \in Z$ and $p, p' \in \Lambda$,

$$\begin{aligned} r(e_z) - r(p) \geq r(e_{z'}) - r(p') &\Longleftrightarrow s(e_z) - s(p) \geq s(e_{z'}) - s(p') \\ &\Longleftrightarrow \bar{\delta}^C(z, p) \geq \bar{\delta}^C(z', p'), \end{aligned}$$

then $s(p) = \alpha r(p) + \beta$, all $p \in \Lambda$. This follows from the fact that r and s both form additive representations of the binary relation generated by $\bar{\delta}^C(z, p) \geq \bar{\delta}^C(z', p')$ on $Z \times \Lambda$ which (as demonstrated in the proof of Theorem 4), satisfies all the assumptions of Theorem III 6.6 on p. 70 of Wakker (1989). As shown by Observation III 6.6′ on p. 71 of the same reference, the class of such representations are jointly cardinal, meaning that r and s must be positive affine transformations of each other.

C. Proofs—Theorems

Proof of Theorem 1. The theorem states that

A DRF $\delta : M \to \mathbb{R}$ permits a DRPE representation if and only if there exists a function $\bar{\delta} : Z \times \Lambda \to \mathbb{R}$ that (a) is an extension of $\delta : M \to \mathbb{R}$ and (b) satisfies A1*–A3* on this extended domain:

A1*: Given $z, z' \in Z$ and $p, p' \in \Lambda$, $\bar{\delta}(z, p) > \bar{\delta}(z', p) \Rightarrow \bar{\delta}(z, p') > \bar{\delta}(z', p')$;

A2*: Given $z, z' \in Z$ and $p, p' \in \Lambda$, $\bar{\delta}(z, p) > \bar{\delta}(z, p') \Rightarrow \bar{\delta}(z', p) > \bar{\delta}(z', p')$;

A3*: Given $z, z' \in Z$, $\bar{\delta}(z', e_{z'}) = \bar{\delta}(z, e_z)$.

Sufficiency: Suppose that a DRF $\delta : M \to \mathbb{R}$ permits a DRPE representation, and let $r : \Lambda \to \mathbb{R}$ and $E : Z \times \Lambda \to \mathbb{R}$ correspondingly represent the DRF. Now define the function $\bar{\delta} : Z \times \Lambda \to \mathbb{R}$ by

$$\bar{\delta}(z, p) = E(r(e_z), r(p)).$$

By construction, this is an extension of the DRF. Consider now $z, z' \in Z$ and $p, p' \in \Lambda$, and note that, as in Proposition 1, the fact that A1* and A2* hold is established by noting their respective equivalence to the following statements. For all $z, z' \in Z$ and $p, p' \in \Lambda$,

$$E(r(e_z), r(p)) > E(r(e_{z'}), r(p)) \Longrightarrow E(r(e_z), r(p')) > E(r(e_{z'}), r(p'));$$
$$E(r(e_z), r(p')) > E(r(e_z), r(p)) \Longrightarrow E(r(e_{z'}), r(p')) > E(r(e_{z'}), r(p)).$$

The truth of both statements is a direct result of the dopaminergic dominance property of $r : \Lambda \to \mathbb{R}$ and $E : r(Z) \times r(\Lambda) \to \mathbb{R}$. Finally, noting that $\forall\, z \in Z$, $\{z, e_z\} \in M$, direct application of Definition 4, no-surprise constancy, implies that for any $z, z' \in Z$,

$$\bar{\delta}(z, e_z) = \delta(z, e_z) = \delta(z', e_{z'}) = \bar{\delta}(z', e_{z'}),$$

confirming that A3* holds.

Necessity: Assume that there exists a function $\bar{\delta} : Z \times \Lambda \to \mathbb{R}$ that (a) is an extension of $\delta : M \to \mathbb{R}$; and (b) satisfies A1*–A3* on this domain. Now pick an arbitrary $\bar{z} \in Z$ and define a function $r : \Lambda \to \mathbb{R}$ by $r(p) = -\bar{\delta}(\bar{z}, p)$, and the correspondence $E : r(Z) \times r(\Lambda) \to 2^{\mathbb{R}}$ by

$$a \in E(x, y) \text{ iff } \exists\, z \in Z, p \in \Lambda \text{ with } r(e_z) = x, r(p) = y, \text{ and } \bar{\delta}(z,p) = a.$$

The first step in the proof that these definitions form the basis of a DRPE representation is to note that $r(e_z) \geq r(e_{z'}) \Longleftrightarrow \bar{\delta}(z, p) \geq \bar{\delta}(z', p)$. To see this, note that by definition and application of A2*, for any $p \in \Lambda$ and $z, z' \in Z$,

$$r(e_z) \geq r(e_{z'}) \Longleftrightarrow \bar{\delta}(\bar{z}, e_z) \leq \bar{\delta}(\bar{z}, e_{z'}) \Longleftrightarrow \bar{\delta}(z, e_z) \leq \bar{\delta}(z, e_{z'}).$$

Substitution of $\bar{\delta}(z, e_z) = \bar{\delta}(z', e_{z'})$ in light of A3* followed by application of A1* then yields the desired conclusion,

$$\bar{\delta}(z, e_z) \leq \bar{\delta}(z, e_{z'}) \Longleftrightarrow \bar{\delta}(z', e_{z'}) \leq \bar{\delta}(z, e_{z'}) \Longleftrightarrow \bar{\delta}(z, p) \geq \bar{\delta}(z', p).$$

Furthermore, given $p, p' \in \Lambda$ and $z \in Z$, A2* implies

$$r(p) \geq r(p') \Longleftrightarrow \bar{\delta}(\bar{z}, p) \leq \bar{\delta}(\bar{z}, p') \Longleftrightarrow \bar{\delta}(z, p) \leq \bar{\delta}(z, p').$$

Given these connections between the functions r and $\bar{\delta}$, it is immediate that E is a function, because given $p, p' \in \Lambda$ and $z, z' \in Z$ with $r(e_z) = r(e_{z'})$ and $r(p) = r(p')$,

$$\bar{\delta}(z, p) = \bar{\delta}(z', p) = \bar{\delta}(z', p').$$

It follows not only that E, r represent the DRF, but also that they respect dopaminergic dominance. To see that E is increasing in its first argument, note that, given any $p \in \Lambda$ and $z, z' \in Z$ with $r(e_z) = x$ and $r(e_{z'}) = y$,

$$x > y \Longleftrightarrow \bar{\delta}(z, p) > \bar{\delta}(z', p).$$

Thus, by definition, $E(x, v) > E(y, v)$ for any $v \in r(\Lambda)$. A similar argument shows that E is strictly decreasing in its second argument, because given any $z \in Z$ and $p, p' \in \Lambda$ with $r(p) = x$ and $r(p') = y$,

$$x > y \Longleftrightarrow \bar{\delta}(z, p) < \bar{\delta}(z, p').$$

Hence r and E respect dopaminergic dominance. Finally, A3* directly implies that no-surprise constancy is satisfied: given $z, z' \in Z$ with $r(e_z) = x$ and $r(e_{z'}) = y$,

$$\begin{aligned} E(x, x) &= \bar{\delta}(z, e_z) \\ &= \bar{\delta}(z, e_z) \text{ (by A3)} \\ &= E(y, y). \end{aligned}$$

Thus r and E as defined above form a DRPE representation of $\delta : M \to \mathbb{R}$, completing the proof.

Proof of Theorem 2. The theorem states that

Given that the DRF $\delta : M \to \mathbb{R}$ satisfies A1–A5, there exists a unique uniformly continuous function $\bar{\delta}^C : Z \times \Lambda \to \mathbb{R}$ that is an extension of δ and satisfies A1*–A3*.

Lemma 1 establishes that, for all $p \in \Lambda$ and finite $X \in 2^Z$, p is a limit point of $\cap_{z \in X} \Lambda(z)$. Furthermore $Z \times \Lambda \subset \text{cl}(M)$. Hence under $A4$, the uniformly continuous function $\delta : M \to \mathbb{R}$ has a unique continuous extension $\bar{\delta}^C : Z \times \Lambda \to \mathbb{R}$. We show now that A1*–A3* hold for this function.

To confirm that A1* holds for $\bar{\delta}^C : Z \times \Lambda \to \mathbb{R}$, consider $z, z' \in Z$ and $p, p' \in \Lambda$ such that $\bar{\delta}^C(z, p) > \bar{\delta}^C(z', p)$. By the continuity of $\bar{\delta}^C$ and the fact that p is a limit point of $\Lambda(z) \cap \Lambda(z')$, we can find $q \in \Lambda(z) \cap \Lambda(z')$ such that

$$\delta(z, q) = \bar{\delta}^C(z, q) > \bar{\delta}^C(z', q) = \delta(z', q).$$

Now take sequence $p'_n \in \Lambda(z) \cap \Lambda(z')$ such that $\lim_{n\to\infty} p'_n = p'$. By A2 we know that, for all n,

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

With this, we know from axiom A5 that

$$\lim_{n\to\infty} (\delta(z, p'_n) - \delta(z', p'_n)) > 0.$$

Applying continuity once more, we conclude that $\bar{\delta}^C(z, p') > \bar{\delta}^C(z', p')$.

To confirm that A2* holds for $\bar{\delta}^C : Z \times \Lambda \to \mathbb{R}$, consider $z, z' \in Z$ and $p, p' \in \Lambda$ such that $\bar{\delta}^C(z, p) > \bar{\delta}^C(z, p')$. We need to show that as a result $\bar{\delta}^C(z', p) > \bar{\delta}^C(z', p')$. The first step is to note that, since $\bar{\delta}^C$ is continuous, any point in Λ is a limit point of $\Lambda(z) \cap \Lambda(z')$, and $\bar{\delta}^C(z, p) > \bar{\delta}^C(z, p')$, we can find $r, s \in \Lambda(z) \cap \Lambda(z')$ such that

$$\bar{\delta}^C(z, p) > \bar{\delta}^C(z, r) = \delta(z, r) > \delta(z, s) = \bar{\delta}^C(z, s) > \bar{\delta}^C(z, p').$$

Our proof of A2* derives from showing that a similar, though not identical, string of inequalities applies to the prize argument $z' \in Z$,

$$\bar{\delta}^C(z', p) \geq \bar{\delta}^C(z', r) = \delta(z', r) > \delta(z', s) = \bar{\delta}^C(z', s) \geq \bar{\delta}^C(z', p').$$

The inner strict inequality, $\delta(z', r) > \delta(z', s)$, follows from application of A2 to $\delta : M \to \mathbb{R}$ in light of the fact that $\delta(z, r) > \delta(z, s)$. The first weak inequality, $\bar{\delta}^C(z', p) \geq \bar{\delta}^C(z', r)$, is established by using the continuity of $\bar{\delta}^C$ and the fact that p is a limit point of

$\Lambda(z) \cap \Lambda(z')$ to construct a sequences $p_n \to p$ with $(z, p_n), (z', p_n) \in M$ and

$$\delta(z, p_n) > \delta(z, r),$$

whereupon application of A2 to $\delta : M \to \mathbb{R}$ ensures that $\delta(z', p_n) > \delta(z', r)$, while the continuity of $\bar{\delta}^C$ ensures that this survives at least weakly in the limit, $\bar{\delta}^C(z', p) \geq \bar{\delta}^C(z', r)$, as required. The second weak inequality $\bar{\delta}^C(z', s) \geq \bar{\delta}^C(z', p')$ is established in identical fashion, confirming that A2* does indeed hold for $\bar{\delta}^C : Z \times \Lambda \to \mathbb{R}$. Finally, confirmation of A3* for $\bar{\delta}^C : Z \times \Lambda \to \mathbb{R}$ follows directly from the fact this property holds for $\delta : M \to \mathbb{R}$.

Proof of Theorem 3. The theorem states that

A DRF $\delta : M \to \mathbb{R}$ that satisfies A4 and A5 admits a DRPE representation if and only if it satisfies A1–A3.

Proposition 1 establishes that A1–A3 are necessary for a DRPE representation, whereas given that the DRF satisfies A4 and A5, Theorems 1 and 2 establish that they are sufficient.

Proof of Theorem 4. The theorem states that

A DRF $\delta : M \to \mathbb{R}$ that satisfies A4 and A5 admits a choice-consistent DRPE representation, with $r : \Lambda \to \mathbb{R}$ continuous, if and only if it satisfies axioms A1–A3, A6, and A7.

Axioms Imply Representation: Suppose that $\delta : M \to \mathbb{R}$ satisfies A1–A5, A6, and A7. Now construct the function $r : \Lambda \to \mathbb{R}$ as in the proof of Theorem 1. From the proof of this theorem we know both that such an $r : \Lambda \to \mathbb{R}$ is continuous and that there exists an E such that r and E form a DRPE. Note that, with A7,

$$r(p) \geq r(q) \iff \bar{\delta}^C(z, p) \leq \bar{\delta}^C(z, q) \iff p \in C(\{p, q\}).$$

Since r is continuous and any set $X \in Q$ is compact, $\arg\max_{s \in X} r(s)$ is always nonempty. Now we show that it must be the case that $C(X) = \arg\max_{x \in X} r(x)$. To show that $C(X) \subseteq \arg\max_{x \in X} r(x)$, note that $p \in C(X)$ and $q \in X$ such that $r(q) > r(p)$ would imply by A7 that $C(\{p, q\}) = q$, so that choosing p from the larger set X would contradict Axiom A6, WARP. To show that $C(X) \supseteq \arg\max_{x \in X} r(x)$, suppose to the contrary that $p \in \arg\max_{x \in X} r(x)$, but $p \notin C(X)$. Now consider some $q \in C(X)$ and note that $r(p) \geq r(q)$, so that $p \in C(\{p, q\})$ by A7. This again is a violation of Axiom A6, WARP. Overall, we conclude that $C(X) = \arg\max_{x \in X} r(x)$, as required.

Representation Implies Axioms: Suppose that a DRF $\delta : M \to \mathbb{R}$ satisfies A4 and A5 and admits a choice-consistent DRPE representation, with $r : \Lambda \to \mathbb{R}$ continuous. We know from Proposition 1 that the existence of a DRPE representation implies A1–A3. The fact that C is representable as the maximization of a binary relation implies that A6 holds. All that is left is to confirm A7, whereby given $p, q \in \Lambda$ and $z \in Z$,

$$\bar{\delta}^C(z, p) \leq \bar{\delta}^C(z, q) \Longleftrightarrow p \in C(\{p, q\}).$$

Given that the DRPE is choice consistent, we know that $p \in C(\{p, q\}) \Longleftrightarrow r(p) \geq r(q)$, hence what we need to show is that,

$$\bar{\delta}^C(z, p) \leq \bar{\delta}^C(z, q) \Longleftrightarrow r(p) \geq r(q).$$

To prove this, first consider $p, q \in \Lambda, z \in Z$ with $r(p) < r(q)$. By the continuity of r, this implies that we will be able to find $s, t \in \Lambda(z)$ such that

$$r(p) < r(s) < r(t) < r(q).$$

By definition of the DRPE, given that $(z, s), (z, t) \in M$,

$$r(s) < r(t) \Longrightarrow \bar{\delta}^C(z, s) > \bar{\delta}^C(z, t).$$

Furthermore, we will be able to find a sequence p_n within $\Lambda(z)$ such that $\lim_{n\to\infty} p_n = p$, so that $r(p_n) < r(s)$ for high enough n, whereupon the DRPE property again implies that $\bar{\delta}^C(z, p_n) > \bar{\delta}^C(z, s)$. By continuity we conclude that $\bar{\delta}^C(z, p) = \lim_{n\to\infty} \bar{\delta}^C(z, p_n) \geq \bar{\delta}^C(z, s)$. A similar argument shows that $\bar{\delta}^C(z, t) \geq \bar{\delta}^C(z, q)$. Finally we conclude that

$$r(p) < r(q) \Longrightarrow \bar{\delta}^C(z, p) > \bar{\delta}^C(z, q).$$

An analogous argument shows that, given $p, q \in \Lambda, z \in Z$,

$$r(p) \geq r(q) \Longrightarrow \bar{\delta}^C(z, p) \leq \bar{\delta}^C(z, q).$$

Proof of Theorem 5. The theorem states:

Let Z be a connected metric space and assume that the DRF $\delta : M \to \mathbb{R}$ satisfies A4 and A5. Then the DRF satisfies A1–A3, A8, and A9 if and only if $\exists$ functions $r : \Lambda \to \mathbb{R}$ and $G : r(Z) - r(\Lambda) \to \mathbb{R}$, with r nonconstant and continuous on both Z and Λ and G continuous and strictly increasing, that form a DARPE representation of the DRF.

Axioms Imply Representation: This proof relies on Theorem *III*.6.6 on p. 70 of Wakker (1989) to show that we have sufficient conditions for the additive representation of the extension of the binary relation $\trianglerighteq$ to $Z \times \Lambda$ defined as $(z, p) \trianglerighteq (z', p')$ if and only if $\bar{\delta}^C(z, p) \geq \bar{\delta}^C(z', p')$. To do so, we need to show that the following conditions hold:

1. *Z and Λ are connected topological spaces*: Z is connected by assumption, whereas Λ is convex with the obvious definitions, hence connected.
2. *$\trianglerighteq$ is a continuous weak order*: $\trianglerighteq$ is clearly a weak order, as it is represented by the function $\bar{\delta}^C$. Moreover, as $\bar{\delta}^C$ is continuous, $\trianglerighteq$ is also continuous.
3. $(z, p) \trianglerighteq (z', p) \Longleftrightarrow (z, p') \trianglerighteq (z', p')$ *and* $(z, p) \trianglerighteq (z, p') \Longleftrightarrow (z', p) \trianglerighteq (z', p') \forall z', z \in Z$ *and* $p, p' \in \Lambda$. These equivalences follow directly from the fact, established in theorem 1, that $\bar{\delta}^C$ satisfies A1*–A3*.
4. *$\trianglerighteq$ satisfies the hexagon condition*: The hexagon condition follows from Assumption A9. This assumption states that given $z, z', w, w' \in Z$ and $p, p', q, q' \in \Lambda$ with $\bar{\delta}^C(z, p) = \bar{\delta}^C(w, q), \bar{\delta}^C(z', p') = \bar{\delta}^C(w', q')$ and $\bar{\delta}^C(z, p') = \bar{\delta}^C(w, q')$, it follows that

   $$\bar{\delta}^C(z', p) = \bar{\delta}^C(w', q).$$

   To establish the result we make the following identifications:

   $$p = l''; p' = q = l'; q' = l;$$
   $$z = a; z' = w = a'; w' = a''.$$

   This equivalent difference asserts that if $\bar{\delta}^C(a, l'') = \bar{\delta}^C(a', l'), \bar{\delta}^C(a', l') = \bar{\delta}^C(a'', l)$ and $\bar{\delta}^C(a, l') = \bar{\delta}^C(a', l)$, then $\bar{\delta}^C(a', l'') = \bar{\delta}^C(a'', l')$, which is precisely the hexagon condition: given $a, a', a'' \in Z$ and $l, l', l'' \in \Lambda$, if

   $$\bar{\delta}^C(a, l') = \bar{\delta}^C(a', l),$$
   $$\bar{\delta}^C(a, l'') = \bar{\delta}^C(a', l') = \bar{\delta}^C(a'', l),$$

   then $\bar{\delta}^C(a', l'') = \bar{\delta}^C(a'', l')$.
5. $\exists p, p' \in \Lambda$ *and* $z, z' \in Z$ *such that* $(z, p) \triangleright (z', p)$ *and* $(z, p) \triangleright (z, p')$: This follows directly from Assumption A8.

The conditions of the theorem are therefore met (separability of Z and Λ are not necessary; see remark *III*.7.1). Thus we know that $\exists\, u : Z \to \mathbb{R}$ and $v : \Lambda \to \mathbb{R}$ such that

$$\begin{aligned} u(z) + v(p) \geq u(z') + v(p') &\iff (z, p) \trianglerighteq (z', p') \\ &\iff \bar{\delta}^C(z, p') \geq \bar{\delta}^C(z', p'). \end{aligned}$$

Thus, A3 holds, and by Proposition 2 there exists an additive DRPE representation.

Finally, we have to show that r and G are continuous. To see this, first note that, as u and v are continuous and nonconstant above, it is clear that the r function constructed according to the proof of Proposition 2 will be continuous (and nonconstant) on both Z and Λ. Thus we only have to show that G is continuous. To see this, let X be the range of $\bar{\delta}^C$ and $Y = r(Z) - r(\Lambda)$. Note that X and Y are intervals. To see this, note that $Z \times \Lambda$ is connected, that both $\bar{\delta}^C$ and r are continuous functions, and that the range of any continuous function on a connected domain is connected, and therefore they constitute intervals in $\mathbb{R}$. Next, note that G is a strictly increasing function from X to Y, and that it is onto. To see this, note that for any $x \in X$, $\exists\, (z, p) \in Z \times \Lambda$ such that $\bar{\delta}^C(z, p) = x$. Let $\bar{y} = \bar{r}(z) - r(p)$. By definition $G(\bar{y}) = G(\bar{r}(z) - r(p)) = \bar{\delta}^C(z, p) = x$. Finally note that any strictly increasing, onto function mapping between intervals in $\mathbb{R}$ is continuous.

Representation Implies Axioms: Suppose that $\exists$ functions $r : \Lambda \to \mathbb{R}$ and $G : r(Z) - r(\Lambda) \to \mathbb{R}$, with r nonconstant and continuous on both Z and Λ and G continuous and strictly increasing, that form a DARPE representation of the DRF. As the DARPE representation is a special case of a DRPE representation we know that A1–A3 must hold by Proposition 1. A8 is implied by the fact that the function r in the representation is nonconstant on both Z and Λ. Finally, to prove A9 define the function $G^* : Z \times \Lambda \to \mathbb{R}$ as $G^*(z, p) = G(r(z) - r(p))$. This function is by definition continuous, and, as G, r form a DARPE, we know that, for any $(z, p) \in M$,

$$G^*(z, p) = G(r(z) - r(p)) = \delta(z, p).$$

Thus G^* is a continuous extension of δ to $Z \times \Lambda$. As such an extension is unique, it must be the case that $\forall\, z \in Z$ and $p \in \Lambda$,

$$G(r(z) - r(p)) = G^*(z, p) = \bar{\delta}^C(z, p).$$

This allows us to prove A9 based on the implied relationship between $\bar{\delta}^C$ and r. Given $z, z', w, w' \in Z$ and $p, p', q, q' \in \Lambda$,

$$\begin{aligned}\bar{\delta}^C(z, p) = \bar{\delta}^C(w, q) &\Longrightarrow r(e_z) - r(p) = r(e_w) - r(q);\\ \bar{\delta}^C(z', p') = \bar{\delta}^C(w', q') &\Longrightarrow r(e_{z'}) - r(p') = r(e_{w'}) - r(q');\\ \bar{\delta}^C(z, p') = \bar{\delta}^C(w, q') &\Longrightarrow r(e_z) - r(p') = r(e_w) - r(q').\end{aligned}$$

This immediately implies that

$$r(e_{z'}) - r(p) = r(e_{w'}) - r(q) \Longrightarrow \bar{\delta}^C(z', p) = \bar{\delta}^C(w', q),$$

completing the proof.

Proof of Theorem 6. The theorem states that

A DRF $\delta : M \to \mathbb{R}$ that satisfies A4 and A5 admits a DERPE representation if and only if it satisfies axioms A1–A3 and A10.

Axioms Imply Representation: Suppose that the DRF $\delta : M \to \mathbb{R}$ satisfies A1–A5 and A10. Now pick an arbitrary prize $\bar{z} \in Z$, and define the binary relation $\gtrsim_D$ on $\Lambda \times \Lambda$ by

$$p \gtrsim_D p' \Longleftrightarrow \bar{\delta}^C(\bar{z}, p) \leq \bar{\delta}^C(\bar{z}, p').$$

Note that this binary relation satisfies the standard axioms of EU theory: it is immediate that $\gtrsim_D$ is complete and transitive; A10 establishes that the substitution axiom is satisfied; and continuity of $\bar{\delta}^C$ implies that given $p, p', p'' \in \Lambda$ with $p \succ_D p' \succ_D p''$, $\exists\ a, b \in (0, 1)$ such that

$$ap + (1 - a)p'' \succ_D p' \succ_D p + (1 - b)p'',$$

in satisfaction of the Archimedean axiom. We conclude that there exists $u : Z \to \mathbb{R}$ such that

$$p \succ_D p' \Longleftrightarrow \mu_p[u] \equiv \sum_{z \in \mathrm{Supp}(p)} u(z)p(z) > \mu_{p'}[u].$$

Moreover, as axioms A1–A3 are satisfied, we know that there exist functions $s : \Lambda \to \mathbb{R}$ and $\bar{E} : s(Z) \times s(\Lambda) \to \mathbb{R}$, which form a DRPE representation of δ. Finally, as

$$\mu_p[u] > \mu_{p'}[u] \Longleftrightarrow \bar{\delta}^C(\bar{z}, p) < \bar{\delta}^C(\bar{z}, p') \Longleftrightarrow s(p) > s(p'),$$

the function $r(p) = \sum_{z \in \mathrm{Supp}(p)} u(z)p(z)$ is a strictly increasing transform of s, and as shown in Proposition 2, this implies that there exists a function E such that E and r form an RPE representation.

Representation Implies Axioms: Suppose that $\delta : M \to \mathbb{R}$ permits a DERPE representation, and let r, E form such a representation. Existence of the DRPE alone implies that A1–A3 are satisfied. To see that A10 holds, note that, given $z \in Z$, p, $p' \in \Lambda(z)$, $q \in \Lambda$, and $\lambda \in (0, 1]$,

$$\begin{aligned}
\delta(z, p) &> \delta(z, p') \\
&\iff E(r(e_z), r(p)) > E(r(e_z), r(p')) \\
&\iff r(p) < r(p') \\
&\iff \sum_{z \in \mathrm{Supp}(p)} u(z)p(z) < \sum_{z \in \mathrm{Supp}(p')} u(z)p'(z) \\
&\iff \sum_{z \in \mathrm{Supp}(p) \cup \mathrm{Supp}(q)} \left[\lambda p(z) + (1-\lambda)q(z)\right] u(z) \\
&\quad < \sum_{z \in \mathrm{Supp}(p') \cup \mathrm{Supp}(q)} \left[\lambda p'(z) + (1-\lambda)q(z)\right] u(z) \\
&\iff r(\lambda p + (1-\lambda)q)) < r(\lambda p' + (1-\lambda)q) \\
&\iff E(r(e_z), r(\lambda p + (1-\lambda)q)) > E(r(e_z), r(\lambda p' + (1-\lambda)q)) \\
&\iff \delta(z, \lambda p + (1-\lambda)q) > \delta(z, \lambda p' + (1-\lambda)q).
\end{aligned}$$

The proof that this result extends, under A4, to the function $\bar{\delta}^C$, is straightforward but lengthy and is included in Dean (2007).

CENTER FOR EXPERIMENTAL SOCIAL SCIENCE, DEPARTMENT OF ECONOMICS
NEW YORK UNIVERSITY
CENTER FOR EXPERIMENTAL SOCIAL SCIENCE, DEPARTMENT OF ECONOMICS
NEW YORK UNIVERSITY

REFERENCES

Adams, E. W., "Elements of a Theory of Inexact Measurement," *Philosophy of Science*, 32 (1965), 205–228.

Apicella, P., T. Ljungberg, E. Scarnati, and W. Schultz, "Responses to Reward in Monkey Dorsal and Ventral Striatum," *Experimental Brain Research*, 85 (1991), 491–511.

Balleine, Bernard W., "Neural Bases of Food-Seeking: Affect, Arousal, and Reward in Corticostriatolimbic Circuits," *Physiology and Behavior*, 86(5) (2005), 717–730.

Balleine, Bernard W., A. Simon Killcross, and Anthony Dickinson, "The Effect of Lesions of the Basolateral Amygdala on Instrumental Conditioning," *Journal of Neuroscience*, 23 (2003), 666–675.

Barto, Andrew, and Richard Sutton, "Simulation of Anticipatory Responses in Classical Conditioning by a Neuron-like Adaptive Element," *Behavioral Brain Research*, 4 (1982), 221–235.

Bayer, H., and P. Glimcher, "Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal," *Neuron*, 47 (2005), 129–141.

Bayer, H., B. Lau, and P. Glimcher, "Statistics of Midbrain Dopamine Neuron Spike Trains in the Awake Primate," *Journal of Neurophysiology*, 98 (2007), 1428–1439.

Bendor, Jonathan, Dilip Mookherjeee, and Debraj Ray, "Aspiration-Based Reinforcement Learning in Repeated Interaction Games; An Overview," *International Game Theory Review*, 3 (2001), 159–174.

Bernheim, B. Douglas, and Antonio Rangel, "Addiction and Cue-Triggered Decision Processes," *American Economic Review*, 94 (2004), 1558–1590.

Berridge, K., "Pleasure, Pain, Desire, and Dread: Hidden Core Processes of Emotion," in *Well-Being: The Foundations of Hedonic Psychology*, D. Kaheman, E. Diener, and N. Schwarz, eds. (New York: Russell Sage Foundation, 1999).

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.

Bush, R. R., and F. Mosteller, "A Mathematical Model for Simple Learning," *Psychological Review*, 58 (1951), 313–323.

Camerer, Colin, and Teck Hua Ho, "Experience-Weighted Attraction Learning in Games: A Unifying Approach," *Econometrica*, 67 (1999), 827–874.

Camerer, Colin, George Loewenstein, and Drazen Prelec, "Neuroeconomics: How Neuroscience Can Inform Economics," *Journal of Economic Literature*, 43 (2005), 9–64.

Caplin, Andrew, "Economic Theory and Psychological Data: Bridging the Divide," in *The Foundations of Positive and Normative Economics: A Handbook*, A. Caplin and A. Schotter, eds. (New York: Oxford University Press, forthcoming 2008).

Caplin, Andrew, and Mark Dean, "Search, Choice, and Revealed Preference," mimeo, New York University, 2008.

Caplin, Andrew, Mark Dean, P. Glimcher, and R. Rutledge, "Measuring Anticipated and Realized Rewards: A Neuroeconomic Approach," mimeo, New York University, 2008.

Charness, G., and D. Levine, "When Optimal Choices Feel Wrong: A Laboratory Study of Bayesian Updating, Complexity, and Affect," mimeo, University of California at Santa Barbara, 2003.

Cheung, Y. W., and D. Friedman, "Individual Learning in Normal Form Games: Some Laboratory Results," *Games and Economic Behavior*, 19 (1997), 46–76.

Daw, N., J. P. O'Doherty, P. Dayan, B. Seymour, and R. J. Dolan, "Polar Exploration: Cortical Substrates for Exploratory Decisions in Humans," *Nature*, 441 (2006), 876–879.

Dean, Mark, "Generalized Reward Prediction Error Models," mimeo, New York University, 2007.

Debreu, G., "Topological Methods in Cardinal Utility Theory," in *Mathematical Methods in the Social Sciences*, K. J. Arrow, S. Karlin, and P. Suppes, eds. (Stanford, CA: Stanford University Press, 1960).

Delgado, M. R., M. M. Miller, S. Inati, and E. A. Phelps, "An fMRI Study of Reward-Related Probability Learning," *NeuroImage*, 24 (2005), 862–873.

Erev, Ido, and Al Roth, "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *American Economic Review*, 88 (1998), 848–881.

Evans, G. W., and S. Honkapohia, *Learning and Expectations in Macroeconomics*, (Princeton, NJ: Princeton University Press, 2001).

Fiorino, D. F., A. Coury, and A. G. Phillips. "Dynamic Changes in Nucleus Accumbens Dopamine Efflux During the Coolidge Effect in Male Rats," *Journal of Neuroscience*, 17 (1997), 4849–4855.

Fishburn, P. C., *Utility Theory for Decision Making* (New York: Wiley, 1970).

Fudenberg, Drew, and David K. Levine, *Theory of Learning in Games* (Cambridge, MA: MIT Press, 1998).

Gallistel, C. R., P. Shizgal, and J. S. Yeomans, "A Portrait of the Substrate for Self Stimulation," *Psychological Review*, 88(3) (1981), 228–273.

Gardner, Eliot, and James David, "The Neurobiology of Chemical Addiction," in *Getting Hooked: Rationality and Addiction*, Jon Elster and Ole-Jorgen Skog, eds. (Cambridge, MA: Cambridge University Press, 1999).

Glimcher, Paul, *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics* (Cambridge, MA/London: MIT Press, 2003).

Gonzales, C., "Two Factor Additive Conjoint Measurement with One Solvable Component," *Journal of Mathematical Psychology*, 44(2) (2000), 285–309.

Grace, A. A., and B. S. Bunney, "Intracellular and Extracellular Electrophysiology of Nigral Dopamine Neurons. Identification and Characterization," *Neuroscience*, 10 (1983), 301–315.

Haruno, Masahiko, Tomoe Kuroda, Kenji Doya, Keisuke Toyama, Minoru Kimura, Kazuyuki Samejima, Hiroshi Imamizu, and Mitsuo Kawato, "A Neural Correlate of Reward-Based Behavioral Learning in Caudate Nucleus: A Functional Magnetic Resonance Imaging Study of a Stochastic Decision Task," *Journal of Neuroscience*, 24(7) (2004), 1660–1665.

Jaffray, J-Y., "On the Extension of Additive Utility to Infinite Sets," *Journals of Mathematical Psychology*, 11 (1974), 431–452.

Kiyatkin, E. A., and A. Gratton, "Electrochemical Monitoring of Extracellular Dopamine in Nucleus Accumbens of Rats Lever-Pressing for Food," *Brain Research*, 652 (1994), 225–234.

Knutson, B., and R. Peterson, "Neurally Reconstructing Expected Utility," *Games and Economic Behavior*, 52 (2005), 305–315.

Li, J., S. M. McClure, B. King-Casas, and P. R. Montague, "Policy Adjustment In A Dynamic Economic Game," *PlosONE*, forthcoming, 2007.

Luce, R. D., and J. W. Tukey, "Simultaneously Conjoined Measurement: A New Type of Fundamental Measurement," *Journal of Mathematical Psychology*, 1 (1964), 1–27.

Marcet, A., and T. Sargent, "Convergence of Least Squares Learning in Mechanisms in Self-Referential Linear Stochastic Models," *Journal of Economic Theory*, 48(2) (1989), 337–368.

Mirenowicz, J., and W. Schultz, "Importance of Unpredictability for Reward Responses in Primate Dopamine Neurons," *Journal of Neurophysiololgy*, 72(2) (1994), 1024–1027.

Montague, P. R., and G. S. Berns, "Neural Economics and the Biological Substrates of Valuation," *Neuron*, 36 (2002), 265–284.

Montague, P. R., Steven Hyman, and Jonathan D. Cohen, "Computational Roles for Dopamine in Behavioural Control," *Nature*, 431 (2004), 760–767.

Nestler, E. J., and R. C. Malenka, "The Addicted Brain," *Scientific American*, 290 (2004), 78–85.

Nyarko, Yaw, and Andrew Schotter, "An Experimental Study of Belief Learning Using Real Beliefs," *Econometrica*, 70 (2002), 971–1005.

O'Doherty, J., P. Dayan, K. J. Friston, H. D. Critchley, and R. J. Dolan, "Temporal Difference Models Account and Reward-Related Learning in the Human Brain," *Neuron*, 38 (2003), 329–337.

O'Doherty, J., T. W. Buchanan, B. Seymour, and R. Dolan, "Predictive Neural Coding of Reward Preferences Involves Dissociable Responses in Human Ventral Midbrain and Ventral Striatum," *Neuron*, 49 (2006), 157–166.

O'Doherty, J., P. Dayan, J. Schultz, R. Deichmann, K. Friston, and R. J. Dolan, "Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning," *Science*, 304 (2004), 452–454.

Olds, J., and P. Milner, "Positive Reinforcement Produced by Electrical Stimulation of Septal Area and Other Regions of Rat Brain," *Journal of Comparative and Physiological Psychology*, 47 (1954), 419–427.

Pessiglione, Mathias, Ben Seymour, Guillaume Flandin, Raymond J. Dolan, and Chris D. Frith, "Dopamine-Dependent Prediction Errors Underpin Reward-Seeking Behavior in Humans," *Nature Online Letters* (2006), 1–4.

Redgrave, P., and K. N. Gurney, "The Short-Latency Dopamine Signal: A Role in Discovering Novel Actions?" *Nature Reviews Neuroscience*, 7 (2006), 967–975.

Redish, A. D., "Addiction as a Computational Process Gone Awry," *Science*, 306 (2004), 1944–1947.

Rescorla, R. A., and A. R. Wagner, "A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement," in *Classical Conditioning II*, A. H. Black and W. F. Prokasy, eds. (New York: Appleton–Century–Crofts, 1972).

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, P., "A Note on the Pure Theory of Consumer's Behavior," *Economica*, 5 (1938), 61–71.

Sarin, R., and F. Vahid, "Predicting How People Play Games: A Simple Dynamic Model of Choice," *Games and Economic Behavior*, 34 (2001), 104–122.

Schultz, Wolfram, "Getting Formal with Dopamine and Reward," *Neuron*, 36 (2002), 241–263.

——, "Getting Formal with Dopamine and Reward," *Neuron*, 36 (2002), 241–263.

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.

Scott, D., "Measurement Structures and Linear and Linear Inequalities," *Journal of Mathematical Psychology*, 1 (1964), 233–247.

Selten, R., and R. Stoecker, "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames," *Journal of Economic Behavior and Organization*, 7 (1986), 47–70.

Shor, Mikhail, "Learning to Respond: The Use of Heuristics in Dynamic Games," mimeo, Vanderbilt University, 2004.

Stahl, D., and P. Wilson, "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10 (1995), 218–254.

Sutton, R. S., and A. G. Barto, *Introduction to Reinforcement Learning* (Cambridge, MA: MIT Press, 1998).

Tversky, A., "Finite Additive Structures," Michigan Mathematical Psychology Program, University of Michigan 64–66, 1964.

Wakker, P. P., *Additive Representations of Preferences: A New Foundation of Decision Analysis* (Dordrecht, The Netherlands: Kluwer Academic, 1989).

Wise, R. A., "Dopamine, Learning, and Motivation," *Nature Reviews: Neuroscience*, 5 (2004), 1–12.

Zink, C. F., G. Pagnoni, M. E. Martin, M. Dhamala, and G. Berns, "Human Striatal Response to Salient Nonrewarding Stimuli," *Journal of Neuroscience*, 23 (2003), 8092–8097.